The Relationship between Cardiometabolic Disorders and Schizophrenia: From Early-Life Origins to the Development of a Cardiometabolic Risk Prediction Algorithm for Young People with Psychosis

Benjamin Ian Perry, BSc, MBBS, MSc, MRCPsych

Clare College

Department of Psychiatry

University of Cambridge

This dissertation is submitted for the degree of Doctor of Philosophy

10 May 2021

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee.

The Relationship between Cardiometabolic Disorders and Schizophrenia: From Early-Life Origins to the Development of a Cardiometabolic Risk Prediction Algorithm for Young People with Psychosis

Benjamin Ian Perry

Thesis Summary

My thesis considers the theme of comorbidity between cardiometabolic disorders and schizophrenia by focussing on three key aspects: the nature of association between cardiometabolic disorders and schizophrenia; the potential for common underlying biological mechanisms for the comorbidity; and the prediction of cardiometabolic risk in young adults with psychosis.

On the nature of association between cardiometabolic disorders and schizophrenia, using longitudinal repeat measure data from a large birth cohort, I found that disruption to glucose-insulin homeostasis through childhood/adolescence is associated with increased risk of psychosis in early-adulthood; may not be fully explained by common sociodemographic and lifestyle factors; and may be specific to it.

On the mechanisms of association between cardiometabolic disorders and schizophrenia, I used a range of genetic and observational epidemiological methods to examine whether inflammation and shared genetic liability may be common underlying biological mechanisms for the comorbidity. Using birth cohort data, I show that genetic risk for type 2 diabetes is associated with psychosis-risk in adulthood, and *vice versa*. I also show that genetic risk for type 2 diabetes may influence psychosis risk by increasing systemic inflammation. Using summary data from large genome-wide association studies (GWAS), I show a thread of evidence for shared genetic overlap between schizophrenia, cardiometabolic and inflammatory traits. Finally, using Mendelian randomization, I show evidence supporting that inflammation may be a common cause for insulin resistance and schizophrenia.

On the prediction of cardiometabolic risk in young adults with psychosis, I performed a systematic review of cardiometabolic risk prediction algorithms and explored their predictive performance in a sample of young people at risk of developing psychosis. In doing so, I show that none are likely to be suitable for this population. Then, using patient data, I developed and externally validated the Psychosis Metabolic Risk Calculator (PsyMetRiC), the first cardiometabolic risk prediction algorithm specifically tailored for young people with psychosis.

Together, my work suggests that cardiometabolic disorders and schizophrenia share aetiologic mechanisms, namely inflammation and shared genetic liability. I have shown that it is possible to accurately predict cardiometabolic risk in young people with psychosis using a tool tailored for the population. Such tools can in future become valuable resources for clinicians to reduce the risk of long-term cardiometabolic morbidity and mortality in people with schizophrenia.

Dedicated to my daughter, Nina, who was conceived and born during this work, and who teaches me more about life and about myself each day. Watching you grow, develop, and interact with the world is a precious reminder that we must always remain curious.

Table *of* Contents

List of Tables

List of Figures

List of Abbreviations

1kg CEU	1000 Genomes Project Phase 3 European Sample
5-HT	Serotonin
ALSPAC	Avon Longitudinal Study of Parents and Children
ARMS	Psychosis At Risk Mental States
BDNF	Brain Derived Neurotrophic Factor
BIC	Bayesian Information Criterion
BLRT	Parametric Bootstrap Likelihood Ratio Test
BMI	Body Mass Index
C.I.	Confidence Interval
C4D	Coronary artery disease genetics consortia
CAARMS	Comprehensive Assessment of At Risk Mental State
CAD	Coronary Artery Disease
CAMEO	Cambridgeshire and Peterborough NHS Foundation Trust EIS
CARDIoGRAM	Coronary artery disease genome wide replication and meta-analysis
CHARGE	Cohorts for heart and aging research in genomic epidemiology
CIS-R	Clinical Interview Schedule-Revised
CNS	Central Nervous System
CRP	C-Reactive Protein
CVD	Cardiovascular Disease
DAG	Directed Acyclic Graph
DIAGRAM	Diabetes genetics replication and meta-analyses
DNA	Deoxyribonucleic Acid
DUP	Duration of Untreated Psychosis
EIS	Psychosis Early Intervention Service
EPV	Events Per Variable Ratio
FEP	First Episode Psychosis
FIML	Full-Information Maximum Likelihood Estimation
GABA	Gamma Aminobutyric Acid
GIANT	Genetic investigation of anthropometric traits
GLGC	Global lipids genetics consortium
GMM	Growth Mixture Modelling
GNOVA	Genetic Covariance Analyzer
GWAS	Genome-wide Association Study
h_2	Single Nucleotide Polymorphism Heritability Estimate
HbA1C	Glycated Haemoglobin
HDL	High-Density Lipoprotein
HOMA2	Computerised Homeostatic Model Assessment for Insulin Resistance
HPA	Hypothalamic Pituitary Adrenal
HyPrColoc	Hypothesis Prioritization Multi-Trait Colocalization
ICD	International Classification of Diseases
IL-6	Interleukin-6
IVW	Inverse Variance Weighted Regression
LD	Linkage Disequilibrium
LDL	Low-Density Lipoprotein
LDSC	Linkage Disequilibrium Score Regression
MAF	Minor Allele Frequency
MAGIC	Meta-analyses of glucose and insulin-related traits consortium
MAPK	Mitogen-Activated Protein Kinase
MHC	Major Histocompatibility Complex
MICE	Multiple Imputation Using Chained Equations

MR	Mendelian Randomization
MR-PRESSO	Mendelian randomization residual sum and outlier test
MVMR	Multi-Variable Mendelian Randomization
NHS	National Health Service
NLR	Neutrophil: Lymphocyte Ratio
NMDA	N-methyl-D-aspartate
OR	Odds Ratio
PC	Principal Components
PE	Psychotic Experience
PGC	Psychiatric Genomics Consortium
PI3/AKT	Phosphatidylinositol 3-Kinase/Protein Kinase B
PLIKSi	Psychosis-Like Symptom Interview
PRISMA	Preferred Reporting Items for Systematic reviews and Meta-analyses
PROBAST	Prediction model Risk Of Bias Assessment Tool
PRS	Polygenic Risk Score
PsyMetRiC	The Psychosis Metabolic Risk Calculator
PUFA	Poly-Unsaturated Fatty Acid
QC	Quality Control
RCT	Randomized Controlled Trial
REM	Random Eye Movement
r_g	Genetic Correlation Estimate
SD	Standard Deviation
SDQ	Strength and Difficulties Questionnaire
SE	Standard Error
SLaM	South London and Maudsley NHS Foundation Trust EIS
SLE	Stressful Life Events
SNP	Single Nucleotide Polymorphism
T2D	Type 2 Diabetes
VLMR-LRT	Vuong-Lo-Mendell-Rubin Likelihood Ratio Test
YPAG	McPin Young Persons Advisory Group
ρ-HESS	Heritability Estimation from Summary Statistics

Declaration of Originality of Work

This dissertation is the product of my own work and includes nothing which is the outcome of collaboration, except where specifically indicated in the text.

The ideas for the work presented in this thesis build on previous work I conducted as part of my combined clinical-academic training. For example, I hypothesized a potential mechanistic role for inflammation in a systematic review and meta-analysis of glucose-insulin homeostasis in first-episode psychosis, published in *The Lancet Psychiatry* in 2016 (Perry et al., 2016). I continued this theme of work during the thesis project of my MSc in Health Research at University of Warwick, where I found that insulin resistance was cross-sectionally associated with psychotic symptoms at age 18 in the ALSPAC birth cohort, and that the association may be moderated by childhood inflammation. Those findings were published in *Schizophrenia Bulletin* in 2018 (Perry et al., 2018).

The examination of cardiometabolic developmental trajectories and associations with schizophrenia-spectrum outcomes (Chapter 2); the systematic review (Chapter 6); and the cardiometabolic risk prediction algorithm (Chapter 7), were proposed to the National Institute for Health Research as part of a successful Doctoral Research Fellowship application. The Mendelian randomization study (Chapter 5) was additionally proposed to The Wellcome Trust and Medical Research Council as part of unsuccessful fellowship applications. Ideas for the analyses examining the potential for shared genetic overlap between schizophrenia, cardiometabolic and inflammatory traits (i.e., the prospective ALSPAC analysis presented in Chapter 3, and the analysis of GWAS summary data presented in Chapter 4) were formed and developed during the completion of the work. Thus, one systematic review and five original studies form the basis of my thesis. Throughout this dissertation, I have taken care to refer to the original source for discoveries or views that are not my own. Any omission is completely unintentional.

Statement of Length

This dissertation contains a total of 53,060 words excluding tables, figures, references and appendices. This is within the prescribed word limit.

Acknowledgement

The greatest thanks go to Professors Golam Khandaker and Peter Jones. They not only supervised the work presented here but were instrumental in helping me to secure my PhD training fellowship. I have learned so much and have developed immeasurably from them both. I will also never forget their kindness and support through some extremely challenging personal circumstances which arose during this work.

I am grateful to Professor Rachel Upthegrove from the University of Birmingham, whom I first met when she interviewed me for my first junior academic psychiatry training post. Rachel has been highly supportive of me ever since. Rachel helped me secure my fellowship, provided insightful comments on all the analyses in this work and with data for PsyMetRiC (Chapter 7).

I am grateful to Professors Nicholas Wareham and Claudia Langenberg from the Institute of Metabolic Science for their belief in me and in the value of studying cardiometabolic health in psychiatry. Nick and Claudia went out of their way to provide me with learning opportunities within their department, which helped me with many aspects of the work presented in this thesis.

I am grateful to a number of advisors and mentors who have helped me think more deeply about and improve my work. Professors Steven Burgess and David Dunger, Drs Jan Stochl, Amy Mason and Eleanor Winpenny from University of Cambridge; Professor Stan Zammit, Drs Hannah Jones and Tom Richardson from University of Bristol; Professor Oliver Howes from Imperial College London; and Professor Swaran Singh from University of Warwick.

I am grateful to a number of colleagues who have supported me. Dr Nick Bowker from the Institute of Metabolic Science helped me with codes for the colocalization analysis (Chapter 4). Mr Owen Crawford and Miss Soomin Jang from The School of Clinical Medicine helped me duplicate the systematic search (Chapter 6). Dr Emanuele Osimo from Imperial College London assisted me in obtaining and processing data for PsyMetRiC (Chapter 7).

I am grateful to the National Institute for Health Research for the fellowship that made this work possible. I am also thankful to The Wellcome Trust and The Medical Research Council, who rejected my fellowship applications, for thickening my skin and driving fire into my belly. I would like to thank Dr Thomas Kabir and Miss Rachel Temple from The McPin Foundation, and all the participants of the Young Persons Advisory Group, who helped to facilitate and provide valuable insights during the development of PsyMetRiC (Chapter 7).

I am grateful to Mum, Dad, and my brother Jack, who help keep me grounded and are always on the end of the phone to remind me of home. Last but not least, I am especially grateful to my loving wife, Michelle, for her unwavering support, patience, and belief.

Peer Reviewed Publications Arising from the Work to Date

Perry, B. I., Stochl, J., Upthegrove, R., Zammit, S., Wareham, N., Langenberg, C., Winpenny, E., Dunger, D., Jones, P. B. & Khandaker, G. M. 2021. **Longitudinal Trends in Childhood Insulin Levels and Body Mass Index and Associations With Risks of Psychosis and Depression in Young Adults**. *JAMA Psychiatry,* 78**,** 416-425.

Perry, B. I., Jones, H. J., Richardson, T. G., Zammit, S., Wareham, N. J., Lewis, G., Jones, P. B. & Khandaker, G. M. 2020. **Common mechanisms for type 2 diabetes and psychosis: Findings from a prospective birth cohort**. *Schizophr Research,* 223**,** 227-235.

Perry, B. I., Burgess, S., Jones, H. J., Zammit, S., Upthegrove, R., Mason, A. M., Day, F. R., Langenberg, C., Wareham, N. J., Jones, P. B. & Khandaker, G. M. 2021. **The potential shared role of inflammation in insulin resistance and schizophrenia: A bidirectional two-sample mendelian randomization study**. *PLOS Medicine,* 18**,** e1003455.

Perry, B. I., Upthegrove, R., Crawford, O., Jang, S., Lau, E., Mcgill, I., Carver, E., Jones, P. B. & Khandaker, G. M. 2020. **Cardiometabolic risk prediction algorithms for young people with psychosis: a systematic review and exploratory analysis**. *Acta Psychiatr Scand,* 142**,** 215-232.

Perry, B.I., Osimo, E.F., Upthegrove, R., Mallikarjun, P., Yorke, J., Stochl, J., Zammit, S., Howes, O., Jones, P.B. & Khandaker, G.M. 2021. **Development and External Validation of The Psychosis Metabolic Risk Calculator (PsyMetRiC): A Cardiometabolic Risk Prediction Algorithm for Young People with Psychosis**. *The Lancet Psychiatry.*

SECTION A

BACKGROUND

Chapter 1

General Introduction

During my PhD, I have addressed the theme of comorbidity between cardiometabolic disorders and schizophrenia-spectrum disorders by focussing on three key aspects, which I have presented in three sections of this thesis. First, I have examined the nature of association between cardiometabolic disorders and schizophrenia-spectrum disorders, addressing the key limitations of existing research. Second, I have examined the potential for common mechanisms, namely shared genetic influences and inflammation, which could at least in part explain the comorbidity between cardiometabolic disorders and schizophrenia. Third, I have considered approaches to improving the clinical identification of cardiometabolic risk in young people with psychosis, focussing on the role of prognosis research.

In the proceeding introduction section, I will first briefly introduce schizophrenia-spectrum and cardiometabolic disorders and then summarise the comorbidity between them. I will outline different mechanisms for the comorbidity, beginning with the traditional attributions of sociodemographic, lifestyle and clinical factors, and ending with evidence from historical studies and studies of young adults, which may call into question the traditional attributions as sole explanations for the comorbidity. I will then describe how existing evidence may indicate the possibility of common biological mechanisms for comorbid cardiometabolic disorders and schizophrenia, focussing on the role of inflammation and shared genetic liability. Finally, I will introduce prognosis research, describe efforts to predict cardiometabolic risk in the general population, and consider the usefulness of cardiometabolic risk prediction in young people with psychosis.

1.1 Evidence for the Comorbidity between Cardiometabolic Disorders and Schizophrenia

1.1.1 A Brief Introduction to Schizophrenia-Spectrum and Cardiometabolic Disorders

1.1.1.1 Schizophrenia-Spectrum Disorders: Definitions, Epidemiology and Mechanisms

Psychotic disorders are a group of psychiatric syndromes characterised by combinations of positive, negative, and cognitive symptoms. Positive symptoms include hallucinations, delusions, and disorganised behaviour and speech. Negative symptoms involve disruptions to motivational and emotional function. Cognitive symptoms can affect numerous cognitive domains, including attention, working memory, verbal learning and memory, and executive function (Kahn et al., 2015). However, psychotic disorders may differ in severity, chronicity, pathophysiology, and treatment (Lieberman and First, 2018). The psychotic disorder group includes schizoaffective disorder, schizophreniform disorder, delusional disorder, and substance-induced psychotic disorder, but the cardinal member of the group is schizophrenia (Lieberman and First, 2018).

Schizophrenia is a complex neuropsychiatric illness first classified in the late 19th century as dementia praecox, with the term schizophrenia coined later in 1908 by Eugen Bleuler (Jablensky, 2010). Schizophrenia usually takes the form of a chronic course of episodic acute illness episodes, termed psychosis, followed by periods of either partial to complete recovery or gradual deterioration in social and occupational function over time (see Figure 1) (Thara, 2004). Frank symptoms of schizophrenia usually precipitate between the second and third decades of life, with the peak age of incidence in males, age 22 years, slightly earlier than the peak age of onset in females, age 26 years (Eranti et al., 2013, Castle et al., 1998). In the UK, schizophrenia accounts for around 30% of all spending on adult mental health care in the NHS. More broadly, mental illness costs the UK economy around £77 billion per year, around 4% of gross domestic product (Department of Health., 2014).

At the first clinical presentation of psychosis (first-episode psychosis; FEP), it may not be possible to pinpoint an accurate classification beyond the broad psychotic disorder group owing to, for example, an incomplete history, pathophysiological uncertainty, and aspects of diagnostic criteria such as symptom chronicity which may require more extended clinical observation (Lieberman and First, 2018). Furthermore, there are no diagnostic laboratory tests for schizophrenia; and so, the diagnosis relies on clinical observation and self-report (See Table 1).

Figure 1: Types of Longitudinal Illness Course in Schizophrenia

Prevalence	Description	Illness Course
22%	One Episode No Functional Impairment	
35%	Multiple Episodes Minimal or No Functional Impairment	
8%	Multiple Episodes Functional Impairment After First Episode Which Does Not Progressively Worsen	
38%	Multiple Episodes Progressive Functional Impairment After Each Episode	

Adapted from Ciompi *et al* (Ciompi, 1980) and Shepherd *et al* (Shepherd et al., 1989)

Table 1: ICD-10 Diagnostic Criteria for Schizophrenia

At least one of
Thought insertion/withdrawal/broadcast/echo
Delusions of control, influence, or passivity
Delusional perception
Third-person auditory hallucinations giving a running commentary
Persistent bizarre delusions
Or at least two of
Persistent hallucinations in any modality
Thought disorder
Catatonic behaviour
Negative symptoms
Significant behaviour change
Duration
Symptoms present for one month
Exclusions
Symptoms not attributable to a mood disorder, an organic brain disorder or a substance use disorder

A prodromal phase consisting of a more subtle decline in cognitive and social functioning commonly precedes the first psychotic episode and can begin years before the onset of frank psychotic symptoms (Compton, 2004). A duration of untreated psychosis (DUP) may also precede the first clinical diagnosis of FEP (McGlashan, 1999), whose length is associated with time to symptom remission (Loebel et al., 1992), response to antipsychotics (Karson et al., 2016), symptom severity, and functional outcomes (Marshall et al., 2005). Meta-analytic evidence suggests that the length of DUP

could explain between 2-13% of the variance of outcome in schizophrenia (Penttila et al., 2014). A core aim of psychosis early intervention services (EIS) is to identify psychosis in its early stages and reduce the length of DUP (McGorry et al., 2008). While some associate a long DUP with a prolonged neurotoxic environment (McGorry et al., 2008) or have hypothesised that it may reflect a more severe subtype of schizophrenia (Morgan et al., 2006), more recent research suggests that lead-time bias may confound the association between prolonged DUP and poor outcomes (Jonas et al., 2020). Therefore, the apparent poor outcomes associated with a prolonged DUP may instead relate to a more advanced stage of psychotic illness that is already associated with poorer functioning.

Schizophrenia has a global prevalence that is relatively heterogeneous in its reporting. Meta-analytic evidence suggests estimates of 0.33% and 0.40-0.47% for period and lifetime prevalence respectively (Saha et al., 2005, Simeone et al., 2015), whereas estimates derived from single samples can be slightly higher. For example, one general population survey reported a lifetime prevalence of 0.87% for schizophrenia, and over 3% for broader categories of psychotic disorders (Perala et al., 2007). Along with differences in study methodology, the heterogeneity in prevalence estimates suggests the importance of both individual- and population-level factors which may influence schizophrenia risk. For example, schizophrenia is more common amongst first- and second-generation migrant groups than people who do not have a personal or family history of migration (Cantor-Graae and Selten, 2005). Schizophrenia is also more common in people who live in disadvantaged areas of inner cities (Kirkbride et al., 2007), and areas with low social cohesion (Boydell et al., 2001).

Additionally, the influential neurodevelopmental hypothesis of schizophrenia posits that early-life environmental disruption can lead to neuronal circuits primed to generate psychotic symptoms in later life, often in the context of heightened biological or psychological stress (Nour and Howes, 2015, Fatemi and Folsom, 2009). For example, babies born in late winter and spring are slightly over-represented among patients with schizophrenia, possibly due to an increased risk of intrauterine infection or maternal vitamin D deficiency during the winter months (see Section 1.2.3). A range of childhood adversities including physical abuse, sexual abuse, maltreatment and bullying, are also associated with an increased risk of developing schizophrenia in adulthood (Stilo and Murray, 2010).

The pathophysiology of schizophrenia is not yet fully understood, but is traditionally considered to result from a final common pathway involving disruption to brain dopaminergic signalling pathways (Howes and Kapur, 2009). The dopamine hypothesis of schizophrenia originated after the discovery of antipsychotics in the 1950s, when animal model studies confirmed that the medications altered dopamine metabolism (Carlsson and Lindqvist, 1963). Research then showed that amphetamine, which increases synaptic dopamine levels, can induce psychotic symptoms (Lieberman et al., 1987). Since those early findings, a wealth of molecular imaging (McGowan et al., 2004, Breier et al., 1997),

post-mortem (Mackay et al., 1982, Howes et al., 2013) and experimental animal model evidence (Featherstone et al., 2007, Moore et al., 2006, Lapiz et al., 2003) has accumulated, indicating abnormal dopamine signalling in areas such as the mesocortical, mesolimbic, and nigrostriatal pathways in schizophrenia. However, more recently, with an understanding that not all patients respond to antipsychotic medications which aim to target the dopamine-D2 receptor, there is a growing understanding that not all patients with schizophrenia show abnormalities in dopamine signalling (Demjaha et al., 2012). Instead, schizophrenia may consist of hyperdopaminergic and normodopaminergic subtypes, with the latter subtype characterised by abnormalities in other neurotransmitter pathways (Howes and Kapur, 2014).

Indeed, a range of other neurotransmitter pathways are associated with schizophrenia, such as glutamate (Hui et al., 2009), acetylcholine (Tani et al., 2015) and gamma-aminobutyric acid (GABA) (Blum and Mann, 2002). Disruption to these pathways may occur in combination and relate to different aspects of symptomatology and illness course (Howes and Kapur, 2009).

Schizophrenia risk has a strong genetic component. The genetic underpinnings of schizophrenia gained prominence initially from early studies showing familial clustering of schizophrenia (Rudin, 1916). Studies of monozygotic twins show concordance rates for a schizophrenia diagnosis are around 30% (Hilker et al., 2018). Studies have shown that the risk of a schizophrenia diagnosis in the offspring of affected and non-affected monozygotic twins is similar, suggesting that even unaffected twins carry a heritable component for schizophrenia without expressing the disease (Kringlen and Cramer, 1989). Such findings suggest that psychosis may lie on a continuum in the population (van Os et al., 2009), with a diagnosis of schizophrenia corresponding to the most extreme end of the spectrum.

Adoption studies, which permit the dissection of genetic from environmental disease risk, help to confirm the importance of genetic risk in schizophrenia. For example, the risk of schizophrenia in offspring of mothers who had the illness was similar whether the biological or an adoptive parent raised them (Tienari et al., 1994, Heston, 1966). Additionally, the offspring of mothers without schizophrenia did not have an increased risk for the illness when raised by parents who had schizophrenia (Wender et al., 1974).

Schizophrenia has a heritability (the amount of phenotypic variance that genetic factors could explain) of up to 80% (Sullivan et al., 2003, Hilker et al., 2018). However, the pattern of genetic influence in schizophrenia, like many complex diseases, is thought to be polygenic rather than Mendelian. Relatively recent genomic advances have helped to illustrate this. For example, over the past 20 years and owing to the breakthroughs of the Human Genome Project (Lander et al., 2001), genome-wide

association studies (GWAS), which involve scanning large sets of genetic variants (single nucleotide polymorphisms; SNPs) across complete sets of deoxyribonucleic acid (DNA), have transformed our understanding of the polygenic component of many complex diseases. The first GWAS of schizophrenia was published in 2007 and, with a sample size of *n*=322, did not report any genome-wide significant findings (Lencz et al., 2007). Contrastingly, a more recently published GWAS for schizophrenia, which included *n*=105,318 participants, reported 145 genetic variants significantly associated with schizophrenia (Pardinas et al., 2018).

However, only around 23% of the current variance of schizophrenia risk can be explained by identified genetic variation, with only 3% arising from GWAS significant SNPs (Woo et al., 2017). Therefore, a large proportion of the genetic contribution to schizophrenia risk is still unknown. One reason for this is that GWAS measure only common genetic variation, and recent evidence from whole-phenome studies suggests that individuals with schizophrenia carry a significant burden of rare, damaging variants that go undetected with standard GWAS methods (Singh et al., 2017).

1.1.1.2 Cardiometabolic Disorders: Definitions, Epidemiology and Mechanisms

Cardiometabolic disorders encompass a constellation of related traits, including cardiovascular diseases (CVD) such as hypertension, atherosclerosis and coronary heart disease, alongside metabolic traits such as type 2 diabetes (T2D) and its predeterminants (insulin resistance and impaired glucose tolerance), dyslipidaemia and obesity. Symptoms of cardiometabolic disorders are broad and range from being subtle or imperceptible, particularly in the earliest stages (e.g., isolated hypercholesterolaemia, insulin resistance or mild hypertension), through to severe pain and loss of consciousness (as in the case of acute myocardial infarction), permanent loss of cognitive or physical function (as in the case of cerebrovascular events) and, at their most severe endpoint, death. Together, cardiometabolic disorders are the number one causes of death worldwide, accounting for 17.9 million lives each year (World Health Organization, 2018). Around 6.8 million adults live with CVD in the UK, costing the NHS around £7.4 billion per year, and the broader economy an estimated £15.8 billion per year (Waterall, 2019).

The traits and features encompassing cardiometabolic disorders are interrelated, interdependent and progressively additive. In the earliest stages, subtle changes to biochemistry may be detectable, e.g., compensated disruption to glucose-insulin homeostasis or disruption to lipid storage (Cohn et al., 2001, Savage et al., 2007). These subtle biochemical changes both predispose to and are predisposed by weight gain and hypertension via mechanisms including inflammation and other intracellular signalling mechanisms, such as the mitogen-activated protein kinase (MAPK) and

Phosphatidylinositol 3-kinase/protein kinase B pathways (PI3/AKT) (de Luca and Olefsky, 2008, Kumphune et al., 2013, Fujishiro et al., 2003).

Over time and left unchecked, these changes can progress to the clustering of components that make up the metabolic syndrome (Gehart et al., 2010). The metabolic syndrome was first coined in 1988 as 'syndrome X' due to an increasing understanding of the links between glucose-insulin homeostasis, blood pressure, lipid storage and adiposity (Reaven, 2001). In the ensuing decades, there was debate about the characterisation and definition of the syndrome (Reaven, 2006, Oda, 2012), and it was renamed metabolic syndrome in 2001 (World Health Organization, 2006). Since that time, there have been numerous diagnostic criteria to define metabolic syndrome (see Table 2). Some have placed precedence on disruption to glucose-insulin homeostasis, others on adiposity, and the most recent harmonised definition taking an agnostic stance and also recognising the importance of ethnic differences in weight deposition.

Regardless of the diagnostic criteria used to define it, the metabolic syndrome has consistently shown a high risk of progression to more distal and chronic phenotypes such as T2D (Shin et al., 2013) and CVD (Wilson et al., 2005), alongside severe disease endpoints such as myocardial infarction (Younis et al., 2016), cerebrovascular events (Boden-Albala et al., 2008) and death (Hildrum et al., 2009). Therefore, the metabolic syndrome is an important marker of past, present, and future cardiometabolic risk.

The global prevalence of cardiometabolic disorders is increasing universally, and they are now a critical global health concern (Saklayen, 2018). For example, in the USA between 1988-2010, average body mass index (BMI) in adults increased by around 0.5% a year (National Center for Health Statistics, 2012). The prevalence of obesity in US adults has now surpassed 40%, and over one in three US adults meets the criteria for metabolic syndrome (National Center for Health Statistics, 2012).

Similarly, in China, the prevalence of adult overweight and obesity has increased from 14.6% to 29% since 1992, and the prevalence of metabolic syndrome is 16% (Delavari et al., 2009). A national survey in Iran reported a metabolic syndrome prevalence in adults of around 35% (Delavari et al., 2009). Further, the global survey of obesity found that the prevalence of overweight and obesity has doubled since 1980 in over half of the 195 countries surveyed, with the most significant increases in nations with a lower socioeconomic index (Afshin et al., 2017). Additionally, the International Diabetes Federation expects the global prevalence of T2D to increase to 10.4% by 2040, with over half of all those diagnosed living in Southeast Asia and the Western Pacific region (Ogurtsova et al., 2017).

Table 2: Diagnostic Criteria for Metabolic Syndrome

	WHO (1998) (Alberti and Zimmet, 1998)	**EGIR (1999)** (Balkau and Charles, 1999)	**IDF (2005)** (Zimmet et al., 2005)	**NCEP ATP III Revised (2005) (Grundy et al., 2005)**	**Harmonized Consensus Definition (2009)** (Alberti et al., 2009)
Required	IGT / IFG / T2D	Plasma FI >75th percentile	Central Obesity (Ethnicity-specific waist circumference)	-	-
Criteria	Above plus two from:	Above plus two from:	Above plus two from:	At least three from:	At least three from:
Obesity	WHR>0.90 (M) / >0.95 (F); *or* BMI>30	WC>94cm (M) / 80cm (F)	-	WC>100cm (M) / >88cm (F)	Central Obesity (Ethnicity-specific waist circumference) *or* BMI >30
Hyperglycaemia	-	-	FPG>5.6mmol/L	FPG >5.6mmol/L; *or* Rx	FPG>5.6mmol/L; *or* Rx
Dyslipidaemia	TG>1.7mmol/L; *or* HDL<0.9mmol/L(M) / <1.0mmol/L (F)	TG>2mmol/L; *or* HDL<1mmol/L	TG>1.7mmol/L; *or* HDL<1.0mmol/L(M) / <1.3mmol/L(F); *or* Rx	TG>1.7mmol/L; *or* HDL<1.0mmol/L(M) / <1.3mmol/L(F); *or* Rx	TG>1.7mmol/L; *or* HDL<1.0mmol/L(M) / <1.3mmol/L (F); *or* Rx
Hypertension	>140/90mmHg	>140/90mmHg; *or* Rx	>130mmHg systolic; *or* >85mmHg diastolic	>130mmHg systolic; *or* >85mmHg diastolic; *or* Rx	>130mmHg systolic; *or* >85mmHg diastolic *or* Rx
Other	Microalbuminuria	-	-	-	-

WHO=World Health Organization; EGIR=European Group for the Study of Insulin Resistance; IDF=International Diabetes Federation; NCEP=National Cholesterol Education Program; ATP III=Adult Treatment Panel III; IGT=Impaired Glucose Tolerance; IFG=Impaired Fasting Glucose; T2D=Type 2 Diabetes; FI=Fasting Insulin; WHR=Waist: Hip Ratio; M=Male; F=Female; WC=Waist Circumference; BMI=Body Mass Index; FPG=Fasting Plasma Glucose; Rx=Prescribed Treatment; TG=Triglycerides; HDL=High Density Lipoprotein.

Cardiometabolic disorders have a common set of malleable and non-malleable risk factors. Non-malleable risk factors include sex, ethnicity, and age. For example, there are well-known sex differences in the epidemiology, aetiology, biology and clinical expression of cardiometabolic disorders (Pradhan, 2014). Before the menopause, increased adiposity is more commonly precipitated in females than males (Kuk and Ardern, 2010), whereas hypertension and disrupted biochemical indices are more common in males (Kim and Reaven, 2013), possibly due to a metabolically-active effect of oestrogen (Gupte et al., 2015). Longer-term cardiovascular outcomes such as CVD affect both sexes but also show differences in presentation and clinical course (Beale et al., 2018).

Ethnicity is also an important cardiometabolic risk factor, and non-White ethnicity is an important risk factor for cardiometabolic disorders (Deboer, 2011). For example, a UK population-based study reported that South Asian ethnicity carried the highest risk for cardiometabolic disorders, followed by Black/African-Caribbean ethnicity, followed by White European ethnicity (Tillin et al., 2005). East Asian ethnicity has also shown to confer a significant risk for the development of cardiometabolic disorders (Nestel et al., 2007).

Age is an important cardiometabolic risk factor, and the risk of all cardiometabolic disorders increases with increasing age (Dhingra and Vasan, 2012). Age is also likely to interact with many non-malleable risk factors because most contribute a cumulative risk over time (Reinikainen et al., 2015). Thus, age becomes increasingly relevant as one gets older.

In the general population, the most important malleable cardiometabolic risk factors include smoking, physical inactivity, sedentariness, and an unhealthy diet. All are thought to be dose-dependent, conveying greater risk with increased length and amount of exposure.

Smoking is strongly associated with adverse cardiometabolic and cardiovascular outcomes (Banks et al., 2019) and remains the leading cause of death in developed nations (Lariscy, 2019). However, whilst a prolonged smoking history increases cardiometabolic risk compared with 'never smoked' (Duncan et al., 2019), some research suggests that smoking cessation in young people can reduce this risk to baseline in as little as five years (Lloyd-Jones et al., 2017).

Physical inactivity is the next most important cause of death in developed nations (McGinnis and Foege, 1993). Findings from the UK analysis of the Global Burden of Diseases Injuries and Risk Factors Study suggest that physical inactivity contributed to 10% of premature deaths from coronary heart disease and 35% of all-cause deaths (Allender et al., 2007). Sedentariness has a global impact on mortality comparable with smoking (Lee et al., 2012) and is distinct from physical inactivity (Salman et al., 2019). For example, high volumes of high-intensity physical activity only partly attenuate the cardiometabolic risk associated with sedentariness (Ekelund et al., 2016). Replacing sedentariness with even light physical activity leads to improvements in insulin sensitivity and lipid profiles not replicated by combining intensive physical activity and a sedentary lifestyle (Duvivier et al., 2018). In increasing recognition of the importance of physical inactivity and sedentariness on cardiometabolic risk, in 2019, the UK Chief Medical Officer published combined guidelines on increasing physical activity levels and reducing sedentariness (Department of Health and Social Care., 2019).

Diet is an important cardiometabolic risk factor, and dietary habits have changed considerably in recent decades, imparting considerable cardiometabolic risk (Anand et al., 2015). For example, snacking and snack foods have risen to prominence (Popkin and Duffey, 2010), eating frequency has increased (Monteiro et al., 2010), and a dietary increase in fried, processed and sugary foods is a global phenomenon (Monteiro et al., 2013). Data from meta-analyses and large cohort studies suggest that such diets are intricately related to cardiometabolic disorders such as T2D (Schwingshackl et al., 2017), obesity (Askari et al., 2020), and CVD (Srour et al., 2019).

Genetic variation conveys a key influence on cardiometabolic risk. Similarly to schizophrenia, family studies have shown clustering of cardiometabolic disorders within families (Slack and Evans, 1966), and twin (Zdravkovic et al., 2002) and adoption (Sundquist et al., 2011) studies have confirmed the genetic contribution to cardiometabolic disorders. More recently, while there has been success in elucidating monogenic causes of rare cardiometabolic disorders such as familial

hypercholesterolaemia via loss-of-function mutations in the low-density lipoprotein (LDL) receptor gene *LDLR* (Goldstein and Brown, 1974), rare forms of cardiometabolic disorders do not significantly impact population-level prevalence (Cambien and Tiret, 2007). Conversely, discoveries of mutations in other genes such as apolipoprotein E are much more common in the general population. While they contribute a weak effect on individual-level disease risk, their common frequency results in a more significant impact at the population level (Cambien and Tiret, 2007).

Recent GWAS of cardiometabolic traits such as BMI have been conducted on samples of close to 700,000 adults, identifying over 300 significant variants (Pulit et al., 2019). Similarly, GWAS of cardiometabolic disorders such as T2D have been conducted in over 400,000 adults, identifying over 150 significant variants (Mahajan et al., 2018). However, while the heritability of most cardiometabolic is predicted to be high, the variance explained by identified genetic variants is but a fraction of this (Elks et al., 2012). Therefore, rarer variants may together play a significant polygenic role in the genetic influence of cardiometabolic risk.

1.1.2 The Burden of Cardiometabolic Comorbidity of Schizophrenia

Schizophrenia is a life-shortening illness (McGrath et al., 2008), and people with schizophrenia live on average 10-15 years less than the general population (Plana-Ripoll et al., 2019). Moreover, while mortality rates in the general population are decreasing, the same reductions in mortality rates have not been observed to the same extent in people with schizophrenia, so the mortality gap is widening (Hayes et al., 2017). Mortality rates for people with schizophrenia are now over 2.5 times higher than the general population, irrespective of sex and socioeconomic status (Saha et al., 2007).

Unnatural causes such as accidents and suicide account for only a small portion of the increased mortality, with more than two-thirds explained by physical illnesses such as a significantly higher prevalence of cardiometabolic disorders (Saha et al., 2007). For example, the prevalence of obesity in older adults with chronic schizophrenia is twice as high as in the general population at 55%, the prevalence of dyslipidaemia is five times as high as in the general population at 70%, and the prevalence of hypertension is three times as high as in the general population at 60% (De Hert et al., 2011). These cardiometabolic phenotypes result in a higher prevalence of metabolic syndrome, which is five times as common as in the general population, and T2D, which is twice as common as in the general population (De Hert et al., 2011). Together, the higher prevalence of cardiometabolic disorders in people with schizophrenia contributes to a three-fold higher risk of death from myocardial infarction and cerebrovascular events than the general population, even after adjusting for factors such as sex, ethnicity, and social class (Correll et al., 2017).

The significant burden of comorbid schizophrenia and cardiometabolic disorders is not just felt by the individual but by the whole of society. Comorbid schizophrenia and cardiometabolic disorders lead to increased health service use through emergency hospital admissions, GP consultations, and prolonged lengths of hospital stay (Hochlehnert et al., 2011). In already stretched healthcare systems, this can contribute toward increased waiting times and poorer care standards universally. Increased use of health services translates into substantial additional healthcare costs. For example, studies have shown that comorbid schizophrenia and cardiometabolic disorders can increase direct healthcare costs by up to 45-75%, even after controlling for severity of physical illness and not including the costs associated with the treatment of schizophrenia (Naylor, 2012).

Together, comorbid physical and psychiatric disorders account for up to 18% of all expenditure on long-term health conditions in the UK (Naylor, 2012). Moreover, in addition to increasing health service costs, comorbid schizophrenia and cardiometabolic disorders can have broader economic implications, such as higher levels of unemployment (Hutter et al., 2010), higher workplace absence due to sickness (Von Korff et al., 2005), and increased use of the benefits system (Naylor, 2012). In the UK, the yearly total societal costs of comorbid schizophrenia and cardiometabolic disorders are £700m higher than the costs of treating schizophrenia and cardiometabolic disorders separately (McDaid, 2015).

1.1.3. Traditional Attributions for the Cardiometabolic Comorbidity of Schizophrenia

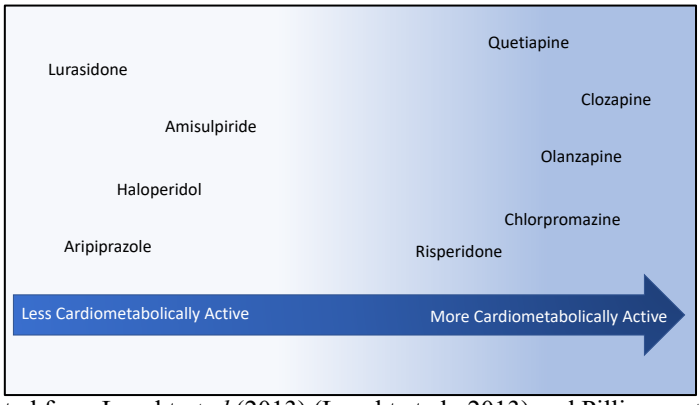
1.1.3.1 The Adverse Effects of Antipsychotic Medication

In 1952, the accidental discovery that chlorpromazine, an anaesthetic medication, may be effective as a calming agent (Laborit et al., 1952) led to its first investigation as a psychiatric treatment (Delay et al., 1952), and later its widespread introduction as the first licensed antipsychotic medication (Ban, 2007). Chlorpromazine catalysed the fledging period of 'deinstitutionalisation', involving the large-scale transfer of psychiatric patients from inpatient units to community care. The introduction of chlorpromazine coincided with the culmination of a wider socio-political movement to provide improved freedoms to psychiatric patients (Niles, 2013). This medication transformed our understanding of schizophrenia pathophysiology (Howes and Kapur, 2009), helped to instil patients with civil liberties, and stimulated the field of psychopharmacology toward the discovery of an array of antipsychotic medications commonly used in modern psychiatry. Whilst none could doubt the transformational improvements these pharmacological developments instigated, they have nonetheless added complexity in examining associations between cardiometabolic disorders and schizophrenia due to an increased risk of confounding.

But a few years after the introduction of chlorpromazine into clinical psychiatric practice, studies began to be published highlighting the potential adverse effects of the medication on cardiometabolic indices. For example, early meta-analytic evidence reported that the more recent 'second-generation' antipsychotics, developed initially to combat the common adverse effects of movement disorders in earlier antipsychotics, exerted more influence on cardiometabolic indices than the earlier 'first-generation' antipsychotics (Bergman and Ader, 2005, Smith et al., 2008).

However, more recently, the consideration that newer 'second-generation' or 'atypical' antipsychotics have greater adverse cardiometabolic effects than older 'first-generation' or 'typical' antipsychotics has been called into question. Newer meta-analyses have shown that the differential cardiometabolic effects of antipsychotics do not necessarily abide by these distinctions (Leucht et al., 2013, Pillinger et al., 2020). For example, aripiprazole conveys relatively little adverse cardiometabolic risk, yet olanzapine conveys significant adverse cardiometabolic risk, and both are second-generation antipsychotics. Similarly, chlorpromazine conveys significant cardiometabolic risk, yet haloperidol does not, and both are typical antipsychotics. It is now generally understood that the metabolically-active nature of different antipsychotics lies on a continuum rather than across a dichotomy (See Figure 2), and the cardiometabolic impact of such medications can precipitate relatively quickly after initiation (Spertus et al., 2018).

Figure 2: Comparative Cardiometabolic Impact of Different Commonly Prescribed Antipsychotic Medications



Adapted from Leucht *et al* (2013) (Leucht et al., 2013) and Pillinger *et al* (2020) (Pillinger et al., 2020)

There is biological plausibility for the cardiometabolic impact of antipsychotic medications. While all antipsychotic medications target the dopamine D2 receptor, none are specific to it and have differing affinities for a wide array of other receptors in the central nervous system and the periphery. For example, antipsychotics bind to histamine-1 (H1), serotonin-2c (5-HT2c) and adrenergic receptors (a2 and b3) in the brain (Starrenburg and Bogers, 2009). Each is important in regulating food intake, and animal model studies have shown that knockouts of the genes coding these receptors cause rats to become obese (Kroeze et al., 2003, Jackson et al., 1997, Leibowitz, 1984).

Additionally, antipsychotics may disrupt glucose-insulin homeostasis at the level of the pancreatic beta-cell in the periphery through a decrease in insulin sensitivity and a resultant increase in insulin secretion (Starrenburg and Bogers, 2009). More metabolically-active antipsychotics such as clozapine, olanzapine, and quetiapine also show relatively high affinities for serotonin-1a (5HT-1a), muscarinic-3 (m3), and a2 receptors, which are expressed on pancreatic beta-cells (DeFronzo and Ferrannini, 1991, Reaven, 1988, Shulman, 2000). Conversely, less metabolically-active antipsychotics such as aripiprazole and ziprasidone show considerably less affinity to these receptors (Leucht et al., 2013).

1.1.3.2 The Effects of Lifestyle Factors

Schizophrenia exerts substantial impacts upon all aspects of the lives of people who are diagnosed with it. Whilst frank positive psychotic symptoms are perhaps the most instantly recognisable features of the illness, they are also the features of the illness that respond the most quickly with antipsychotic treatment. Negative and cognitive symptoms of schizophrenia can be harder to identify due to their subtle and insidious nature and typically respond less actively to antipsychotic treatment (Harvey et al., 2016). Negative and cognitive symptoms are common in schizophrenia and account for much of the long-term morbidity and poor functional outcome associated with it (Austin et al., 2013). Negative symptoms such as amotivation, decreased sociability and decreased spontaneity can be pervasive. Such changes can have wide-ranging impacts upon the health of the sufferer (Kirkpatrick et al., 2001) and may predispose to lifestyle factors that could increase cardiometabolic risk.

1.1.3.2.1 Diet

Multiple studies (McCreadie et al., 1998, McCreadie and Scottish Schizophrenia Lifestyle, 2003, Heald et al., 2017) have shown that the diets of people with schizophrenia may be less healthy than the general population. A recent systematic review of observational and interventional studies on diet in schizophrenia found consistent associations between having the illness and consuming a diet higher in refined sugars and saturated fats and lower in fibre (Aucoin et al., 2020). Such a pattern is typical of the 'Western diet' and is associated with adverse cardiometabolic outcomes (Fung et al., 2001).

Furthermore, cross-sectional studies have shown that people with schizophrenia, on average, may consume lower than recommended levels of ω-3 polyunsaturated fatty acids (PUFAs) (Aucoin et al., 2020). Similarly, a large prospective cohort study of women also found that lower consumption of ω-3 PUFAs was associated with an increased risk of psychotic symptoms (Hedelin et al., 2010). ω-3

PUFAs are associated with a favourable cardiometabolic profile of lower cholesterol levels, lower blood pressure and lower levels of systemic inflammation (Natto et al., 2019, Cabo et al., 2012).

1.1.3.2.2 Smoking

People with schizophrenia are over three times more likely to smoke than the general population (de Leon and Diaz, 2005). While the prevalence of smoking in the general population has declined over the past two decades (Windsor-Shellard, 2020), the prevalence of smoking in schizophrenia remains high (Ziaaddini et al., 2009). Some attribute the higher prevalence of smoking in schizophrenia to symptom amelioration since nicotine may have short-term cognitive-enhancing effects (Freedman, 2014). Also, smoking may ameliorate perturbations in dopaminergic, glutamatergic and GABAergic pathways observed in schizophrenia (Lucatch et al., 2018).

However, more recent research is beginning to call the symptom-amelioration hypothesis into question, finding that cigarette smoking in schizophrenia was associated with impairments in memory (Stramecki et al., 2018) and even increased suicidality (Dickerson, 2019). A systematic review of longitudinal studies also found that adolescent exposure to smoking was associated with a higher risk of developing schizophrenia (Gurillo et al., 2015), and recent evidence suggests a potential bi-directional association of smoking with risk of schizophrenia (Wootton et al., 2020).

1.1.3.2.3 Physical Inactivity and Sedentariness

A meta-analysis of 69 case-control studies reported that, compared with the general population, people with schizophrenia on average spend more time sedentary and are less likely to meet recommended physical activity guidelines (Vancampfort et al., 2017). In addition, people with schizophrenia may overestimate their physical activity levels. For example, a large-scale population-based cohort study found that while people with schizophrenia self-reported similar physical activity levels to the general population, objective accelerometer data suggested that they engaged in much less physical activity than the general population, and so over-estimated their activity levels (Firth et al., 2018).

1.1.3.2.4 Alcohol

While some studies have found that light alcohol consumption with meals is associated with a lower risk of incident cardiometabolic disorders (Zhang et al., 2014), possibly due to beneficial effects on inflammation (Piano, 2017), an abundance of research suggests that heavy alcohol use is associated

with increased risk of metabolic syndrome (Vieira et al., 2016), hypertension (Bermudez et al., 2015), T2D and CVD (Roerecke and Rehm, 2014). These associations may be explained by the effects of alcohol itself on endothelial function and myocardial function (Goncalves et al., 2015), apoptosis (Fernandez-Sola et al., 2011), oxidative stress (Piano and Phillips, 2014) and haemostatic factors (Salem and Laposata, 2005), or mediated via comorbid poor diet or smoking (Sluik et al., 2016).

People with schizophrenia have up to a three-times higher prevalence of alcohol use disorders than the general population (Hartz et al., 2014). A recent meta-analysis found that over one in four patients with schizophrenia would meet the criteria for an alcohol use disorder (Hunt et al., 2018), and alcohol use disorders could be present before the onset of psychosis in young adults (Brunette et al., 2018).

A combination of genetic and environmental factors could explain the comorbidity between alcohol use disorders and schizophrenia. Regarding genetic factors, alcohol use disorders and schizophrenia may share genetic liability and genetic overlap (Walters et al., 2018). For example, genetic variants associated with brain-derived neurotrophic factor (BDNF) correlate with comorbid schizophrenia and alcohol use disorders but not with alcohol use disorders alone (Cheah et al., 2014). Regarding environmental factors, alcohol use in schizophrenia may lead to symptom reduction (Hjorthoj et al., 2015) or decrease antipsychotic side effects (Khantzian, 1997). Alcohol use disorders in schizophrenia are associated with poor adherence to treatment, an increased frequency of psychosis relapse, longer duration of inpatient stays, and poor functional outcomes (Kerner, 2015, Archibald et al., 2019).

1.1.3.2.5 Sleep

Inadequate sleep quantity and quality is associated with adverse cardiometabolic outcomes. In both children and adults, short sleep duration is associated with an increased risk of obesity in meta-analyses of cross-sectional and longitudinal studies (Miller et al., 2018). Poor sleep quality and quantity are also longitudinally associated with the development of hypertension (Knutson et al., 2009), T2D (Cappuccio et al., 2010), coronary heart disease (Cappuccio et al., 2011) and cerebrovascular events (Leng et al., 2015). Proposed mechanisms include alterations to circadian rhythms involving cortisol which may disrupt glucose-insulin homeostasis; increases in appetite-increasing grehlins; increases in systemic inflammation; and hypothalamic-pituitary-adrenal (HPA) axis alterations leading to weight gain (Cappuccio and Miller, 2017).

Disturbed sleep is common in schizophrenia and is self-reported in 30-80% of patients, depending on the severity of symptomatology (Yang and Winkelman, 2006, Royuela, 2002, Kato et al., 1999). Meta-analyses of studies examining objective measures of sleep such as polysomnography have

shown changes in sleep latency, sleep efficiency, awake time, slow-wave sleep and random eye movement (REM) sleep in schizophrenia (Yang and Winkelman, 2006, Chouinard et al., 2004, Krystal et al., 2008). Antipsychotics may also affect sleep, depending on the level of histamine-1 receptor (H1) antagonism (Kane and Sharif, 2008). Histamine receptors may be involved in regulating circadian rhythms and the sleep-wake cycle, and H1-receptor blockade can lead to an increase in somnolence and sedation as well as changes in sleep architecture (Monti and Monti, 2004).

Metabolically-active antipsychotics such as clozapine and quetiapine show a high affinity to the H1 receptor and have pronounced effects on sleep induction and total sleep time. Comparatively, less metabolically-active antipsychotics such as aripiprazole and risperidone show low affinity to the H1 receptor (Monti and Monti, 2004). Additionally, antipsychotics differ in affinities to adrenergic, 5HT-2 and cholinergic receptors, which all have roles in sedation (Cohrs, 2008).

1.1.3.3 Healthcare Inequalities

Healthcare inequalities are another important cause for the increased risk of cardiometabolic disorders in schizophrenia. Research has shown that people with schizophrenia may be less likely to attend their general practitioner for physical health concerns than the general population (Goldman, 1999, Brown et al., 2000). When they do attend, they may be less likely to be diagnosed with physical health problems than the general population (Goldman, 1999, Jeste et al., 1996). Research from whole-population studies suggests that following a first hospital admission for CVD, people with schizophrenia are more likely to die and die sooner than the general population (Westman et al., 2018).

Other research has found that amongst hospital admission for ischemic heart disease, people with schizophrenia were half as likely to be recommended for surgical intervention (Lawrence and Kisely, 2010), twice as likely to suffer from hospital-acquired infections, and have a longer length of stay than the general population (Daumit et al., 2006). Among people with a diagnosis of T2D, people with comorbid schizophrenia were half as likely to be offered a referral for specialist care (Jones et al., 2008).

One potential contributor to the discrepancies in healthcare access for people with schizophrenia may be ‘diagnostic overshadowing’, which is defined as the attribution of clinical symptoms and behaviours by clinicians to the mental disorder rather than a physical illness, leading to inadequate assessment and delayed treatment (Jones et al., 2008). Surveys of liaison psychiatrists have cited concerns about stigmatising attitudes by general healthcare staff toward people with schizophrenia and a lack of understanding of complex presentations and challenging behaviours. Similarly,

qualitative research of patients with mental disorders, including schizophrenia, has reported common themes such as feeling stigmatised against by healthcare staff, and barriers to healthcare access due to perceived social isolation (Kemp, 2014).

Furthermore, fragmentation of physical healthcare across primary and secondary mental health services may contribute to healthcare inequalities in schizophrenia (Crawford et al., 2014). For example, primary care staff may not always be confident working with patients with mental disorders, including schizophrenia (Blythe and White, 2012), and reciprocally, secondary mental health staff report low confidence in identifying and managing physical health problems (Happell et al., 2012).

1.1.4 Evidence That the Association between Cardiometabolic Disorders and Schizophrenia May Not Be Fully Explained By Lifestyle Factors and Adverse Treatment Effects

Antipsychotic medications, lifestyle factors and healthcare inequalities are key contributors to the comorbidity between cardiometabolic disorders and schizophrenia. However, a growing body of observational evidence is beginning to question the notion that the aforementioned traditional attributions are likely to be the full explanation for the comorbidity. This observational research can be divided into historical research that predates antipsychotic medication; research conducted on young adults with FEP; and research conducted on adolescents/young adults who are at risk of developing psychosis.

1.1.4.1 Historical Evidence Predating The Use of Antipsychotic Medication

The scientific literature has recognised the cardiometabolic comorbidity of schizophrenia since the beginning of the 20th century (Kohen, 2004), long before the discovery of antipsychotic medication. Indeed, Henry Maudsley once referred to T2D as "a disease which often shows itself in families in which insanity prevails" (Maudsley, 1895). The first observational research in the field was published in 1919 and consisted of a small cross-sectional study of 10 participants showing the commonality of hyperglycaemia in schizophrenia (Kooy, 1919). Two years later, a cross-sectional study of 22 participants with schizophrenia found common abnormalities in glucose tolerance, using an early form of the oral glucose tolerance test (Lorenz, 1922). In 1944, the first case-control study was published examining for differences in glucose tolerance between returning war-time soldiers with psychiatric diagnoses and healthy controls, finding higher rates of impaired glucose tolerance in the case compared with the control group (Drury, 1921).

While these historical studies had some methodological shortcomings, for example, small sample sizes and an inability to consider the direction of association, they demonstrate a thread of consistent evidence for the association of disrupted glucose-insulin homeostasis in schizophrenia in antipsychotic-naïve individuals. Therefore, this suggests that confounding by antipsychotic medication may not fully explain the cardiometabolic comorbidity of schizophrenia.

Additionally, other important features of historical studies suggest that confounding by chronic lifestyle factors may not fully explain the cardiometabolic comorbidity of schizophrenia. While the average UK life expectancy has increased from 50 to 80 years since 1910 (Raleigh, 2020), several negative influences on population health have increased over the last century to offset those gains. For example, participants recruited into historical studies may have been less affected by the adverse impacts of the modern 'Western diet' as outlined in Section 1.1.1.2, including food over-supply and over-consumption, along with consumption of higher proportions of high-calorie and nutrient-deficient foods. For example, McDonalds opened its first restaurant in 1937, the first KFC franchise opened in 1952, and the first Burger King franchise opened in 1954. Moreover, the menu offerings of such establishments are becoming increasingly unhealthy over time (McCrory et al., 2019).

Sedentary behaviour is also increasing over time (Yang et al., 2019). Contributors to this trend include an increasing amount of time spent across childhood, adolescence and adulthood watching television, and technological advances leading to increasing amounts of leisure time spent on computers (Yang et al., 2019). Furthermore, work roles have become increasingly sedentary over recent decades. Research in the US has shown average decreases in occupation-related energy expenditure of over 100 calories per day in both men and women since the 1960s (Church et al., 2011).

Together, this pattern of change in population health suggests that historical research on the associations between cardiometabolic disorders and schizophrenia may be less affected than modern research by some of the traditional attributions for the cardiometabolic comorbidity of schizophrenia. Therefore, historical findings imply that factors such as metabolically-active antipsychotic medications, a poor diet, and sedentariness may more likely exacerbate rather than cause the cardiometabolic comorbidity of schizophrenia. Nevertheless, there are limitations in interpreting the findings of historical studies in the field. For example, the majority featured small sample sizes, were cross-sectional, included dated definitions and assessment techniques for schizophrenia and cardiometabolic disorders, and may be rated at high risk of bias by modern standards.

1.1.4.2 Evidence of Cardiometabolic Dysfunction in Young Adults with FEP

Modern studies conducted on samples of young adults with FEP provide further evidence that traditional attributions may not fully explain the cardiometabolic comorbidity of schizophrenia. Such studies are essential because most cardiometabolic risk factors such as smoking, diet, and physical activity confer cumulative risk over time (Reinikainen et al., 2015). Therefore, studies conducted on young adults can lessen the confounding impact of these traditional attributions.

Studies conducted in participants presenting with FEP can also help to address the confounder of antipsychotic medication at least partly, since participants enrolled in such studies are likely to have had minimal, if any, prior antipsychotic exposure. Modern studies are also able to address methodological and measurement-related shortcomings of historical studies.

1.1.4.2.1 Evidence for Disruption to Glucose-Insulin Homeostasis in FEP

Consistent findings from recent meta-analyses of case-control studies (Perry et al., 2016, Pillinger et al., 2017a, Greenhalgh et al., 2017) suggest that subtle aberrations in glucose-insulin homeostasis are detectable from the onset of psychosis in young antipsychotic naïve adults compared with healthy controls matched on age, sex, ethnicity and body mass index. For example, compared with controls, FEP cases had a higher prevalence of insulin resistance measured using the updated and computerised homeostatic model assessment (HOMA2) method (Levy et al., 1998); and a higher prevalence of impaired glucose tolerance.

Insulin resistance and impaired glucose tolerance are early forms of disrupted glucose-insulin homeostasis and relate to decreased sensitivity of peripheral cells to insulin (O'Rahilly et al., 1994). This decreased insulin sensitivity results in lower glucose transport into cells. In turn, this leads to a negative feedback loop involving increased insulin secretion to maintain stable plasma glucose levels. Therefore, insulin resistance corresponds to a physiological state where higher circulating insulin levels are required to maintain steady plasma glucose levels (Samuel and Shulman, 2016).

Impaired glucose tolerance represents a state of reduced peripheral insulin sensitivity, progressive loss of beta-cell function, and reduced secretion of glucose-dependent insulinotropic polypeptide (Faerch et al., 2009). This state of disrupted glucose-insulin homeostasis accentuates following oral consumption of a glucose-rich bolus. In impaired glucose tolerance, plasma glucose levels take longer to stabilise secondary to a) decreased sensitivity of cells to insulin; and b) an attenuated response range for insulin to correct plasma glucose imbalance. Therefore, both insulin resistance and impaired glucose tolerance represent states of early glucose-insulin dyshomeostasis (Tabak et al., 2012).

Recent studies have reported consistent findings regarding the higher risk of insulin resistance in FEP, extending upon previous meta-analytic results. For example, cross-sectional research suggests that insulin resistance at FEP may be more strongly associated with negative rather than positive schizophrenia symptomatology (Misiak et al., 2019), may be associated with childhood stressful life events (Tosato et al., 2020), and may not be explained by chronic stress (Steiner et al., 2017). Longitudinal research also suggests that insulin resistance in FEP may be a baseline risk factor for weight gain during the first year after FEP (Keinanen et al., 2015).

Studies of glucose-insulin homeostasis in FEP have not found evidence for abnormalities in fasting plasma glucose or glycated haemoglobin (Perry et al., 2016, Pillinger et al., 2017a), which is unsurprising since such alterations represent more chronic, pronounced and potentially irreversible phenotypes within the realms of a T2D diagnosis. Therefore, insulin resistance is an early marker of a more chronic phenotype in T2D, just as FEP could be considered an early marker of a more chronic phenotype of schizophrenia. Since these early phenotypes appear to precipitate during the same period in the life course, even after accounting for several relevant confounders, shared biological processes may link both phenotypes, in at least a subset of individuals with FEP.

1.1.4.2.2 Evidence for Disruption to Lipid Homeostasis in FEP

Recent meta-analyses of case-control studies have also shown that alterations in lipid homeostasis, such as hypertriglyceridaemia and reduced total and low-density lipoprotein (LDL) cholesterol levels, are also detectable from FEP in young adults (Pillinger et al., 2017b, Misiak et al., 2017). One meta-analysis also reported lower high-density lipoprotein (HDL) levels in FEP cases compared with controls (Misiak et al., 2017). Additionally, findings from subsequent studies have extended upon the meta-analytic results. For example, longitudinal research indicates that triglycerides may be associated with worse psychiatric outcomes at both one and two years (Osimo et al., 2021) after FEP.

Together, this pattern of lipid alteration suggests that the primary cardiometabolic risk-increasing phenotype in early psychosis relates to disruption to glucose-insulin homeostasis rather than lipid dysfunction for two reasons. First, a pattern of low total and LDL cholesterol represents lower cardiometabolic risk. For example, large-scale observational studies have consistently found that LDL and total cholesterol are positively associated with a higher risk of coronary heart disease (Ference et al., 2017, Peters et al., 2016). Second, a pattern of raised triglycerides and low HDL is a hallmark of insulin resistance both in older (McLaughlin et al., 2005) and younger adults (Murguia-Romero et al., 2013). A raised triglyceride:HDL ratio has been suggested as a suitable surrogate marker for insulin resistance when it may not be possible to measure it using the HOMA2 or gold standard hyperinsulinaemic-euglycaemic clamp method (Pantoja-Torres et al., 2019).

1.1.4.2.3 Limitations of Existing Studies of Young Adults with FEP

While the consistency and biological plausibility from studies of cardiometabolic dysfunction at FEP suggest the possibility for primary disruptions to glucose-insulin homeostasis in schizophrenia, current studies are limited primarily for three reasons. First, existing research in the field is mostly either cross-sectional or has included incident cases of FEP at baseline, so it is not possible to consider the direction of association. For example, the first clinical presentation of FEP may not accurately correspond with the actual onset of psychotic symptoms, and a duration of untreated psychosis may precede the first clinical presentation by months or even years (Compton, 2004).

Second, existing studies have primarily included one-off measures of cardiometabolic markers, overlooking the potential for dynamic temporal changes in these markers. Cardiometabolic indices, including measures of glucose-insulin homeostasis (Moebus et al., 2011) are subject to normal fluctuation, which cannot be addressed with a one-off measurement. Alternatively, repeated measurements of glucose-insulin homeostasis over time could provide a more reliable measure of potential underlying biological mechanisms.

Third, whilst meta-analyses have included case-control studies that matched participants for relevant potential confounders, residual confounding remains a possible explanation, as is the case with all observational research. For example, existing studies have mostly not adjusted for alcohol use, smoking, physical activity levels, dietary intake, or sleep problems. As described in Section 1.1.3, these lifestyle factors are associated with schizophrenia and cardiometabolic disorders and are likely to be relevant potential confounders.

1.1.4.3 Evidence for Cardiometabolic Dysfunction in Adolescents/Young Adults at Risk of Developing Psychosis

Studies of cardiometabolic indices in adolescents/young adults who are at risk of developing psychosis provide further evidence for the potential of a primary disruption to cardiometabolic function in schizophrenia. Evidence from such studies potentially casts backwards further in the life-course the cardiometabolic associations of schizophrenia and may point to the suggestion that cardiometabolic dysfunction may precede the onset of psychosis in at least some individuals. Evidence from such studies can also further address the potential for confounding since participants may be even less likely to have been prescribed antipsychotic medications than cases of FEP. Additionally, since participants are generally younger, the risk of confounding by chronic lifestyle factors is further reduced.

A 2016 meta-analysis of 47 studies, which examined the association of cardiometabolic risk factors with ultra-high risk for psychosis (UHR) in young adults, found that none included indices of glucose-insulin or lipid homeostasis (Carney et al., 2016). The review also found no significant difference in BMI between cases of UHR with matched controls (Carney et al., 2016). However, since that time, a consistent thread of evidence has emerged showing an association between insulin resistance and UHR status, for example, in case-control studies matched by factors such as age, sex, ethnicity and BMI levels (Petruzzelli et al., 2018, Cadenhead et al., 2019); and cross-sectional studies (Scott et al., 2019, Perry et al., 2018) including one which adjusted for a range of potential confounders including sex, ethnicity, BMI, social class, smoking and alcohol use (Perry et al., 2018). One longitudinal study found no evidence between childhood insulin levels and psychotic symptoms at age 18 years, although the sample size was relatively small (Perry et al., 2018).

Lipid alterations have also been detected in at-risk mental states in the findings from case-control studies and may be helpful to predict transition to psychosis (Lamichhane et al., 2021, Dickens et al., 2021). One longitudinal study found an association between childhood alteration in lipid profiles with psychotic symptoms at age 18 years (Madrid-Gambin et al., 2019). Paradoxically, longitudinal research suggests that lower BMI in childhood and adolescence (Zammit et al., 2007, Weiser et al., 2004, Sorensen et al., 2006) is associated with a higher risk for developing schizophrenia in adulthood.

1.1.4.3.1 Limitations of Existing Studies of Adolescents/Young Adults at Risk of Developing Psychosis

Together, existing evidence suggests that alterations to cardiometabolic indices may occur before the development of psychosis. However, the primary limitation of existing research on younger participants at risk of developing psychosis is the heterogeneity of at-risk mental states. At present, there are no accurate means to distinguish who will and who won't develop psychosis from a baseline of being at risk of developing it. For example, psychotic symptoms in adolescence are also strongly associated with other mental disorders, including anxiety and depression (Varghese et al., 2011), and only around 30% of people classified as at risk for psychosis develop FEP within three years (Fusar-Poli et al., 2012). Other limitations of existing research mirror those of research in FEP (See Section 1.1.4.2.3), for example, the paucity of adequately powered longitudinal studies, the lack of appropriate confounding adjustment, and the inclusion of single point-measures of cardiometabolic indices.

1.2 Existing Evidence for Common Biological Mechanisms for Comorbid Cardiometabolic Disorders and Schizophrenia

1.2.1 Evidence for Disruption to Glucose-Insulin Homeostasis as a Potential Cause for Comorbid Cardiometabolic Disorders and Schizophrenia

Recently developed epidemiological approaches can examine for evidence of potential causality between an exposure and an outcome. For example, Mendelian randomisation (MR) is an epidemiological approach that uses genetic variants (single nucleotide polymorphisms or SNPs) as proxies for a putative risk factor to untangle the problems of reverse causation and unmeasured confounding. This is because genetic variants are fixed at conception; hence genetically-predicted levels of risk factors must precede any event, and genetic variants are often specific in their associations with risk factors (Smith and Ebrahim, 2003).

MR studies examining the associations between genetically-predicted levels of cardiometabolic indices and schizophrenia are relatively scarce, have included a limited set of cardiometabolic exposures, and have reported mixed findings. For example, one previous MR study reported only weak evidence of an association between genetically-predicted insulin resistance schizophrenia (Polimanti et al., 2017). Another reported an association between genetically-predicted fasting insulin and schizophrenia, which attenuated to the null after adjusting for BMI (Li et al., 2018). Furthermore, previous MR studies have included ethnically heterogeneous samples, which increases the risk of population stratification bias (Brumpton et al., 2020). Finally, since the previous MR studies were published, larger GWAS have been conducted, which could increase the statistical power of MR research.

There is biological plausibility for the potential causal association between disruption to glucose-insulin homeostasis and schizophrenia. For example, CNS insulin can regulate striatal dopamine and glutamate levels (Nash, 2017, Caravaggio et al., 2015), and, reciprocally, both CNS insulin and striatal dopamine can regulate peripheral glucose-insulin homeostasis (Berndt et al., 2013). Peripheral insulin can also cross into the CNS via cannabinoid and N-methyl-D-aspartate (NMDA) receptors, and is actively transported into the CNS via the blood-brain barrier (Dodd and Tiganis, 2017). Insulin receptors are widely expressed in the brain, with notable concentrations in regions of the brain known to be associated with schizophrenia, such as the hypothalamus, midbrain and dopaminergic neurons (Figlewicz et al., 2003), striatum, prefrontal cortex, amygdala, and hippocampus (Unger et al., 1991). Brain insulin is associated with the regulation of neuronal growth and neuronal plasticity (Schulingkamp et al., 2000, Ferrario and Reagan, 2018), and is associated with memory and cognition (Grillo et al., 2015).

1.2.2 Evidence for Genetic Liability as a Common Biological Mechanism for Comorbid Cardiometabolic Disorders and Schizophrenia

Shared genetic liability may also be a common biological mechanism for the cardiometabolic comorbidity of schizophrenia, as first mooted in perspectives articles in the early 2000s (Lin and Shuldiner, 2010, Gough and O'Donovan, 2005). Since that time, due to improving analysis methods and larger genetic samples, evidence has begun to accumulate to suggest that shared genetic liability may at least partly explain the common comorbidity between schizophrenia and cardiometabolic disorders. This evidence can be divided into prospective observational research; and secondary analyses of large-scale GWAS datasets.

Regarding prospective observational research, studies conducted in relatively small samples have shown that the prevalence of insulin resistance (Chouinard et al., 2019) and impaired glucose tolerance (Ferentinos and Dikeos, 2012) is higher in unaffected relatives of patients with schizophrenia compared with matched controls. These findings suggest that genetic influences on glucose-insulin signalling may co-occur with genetic influences for psychosis, independent of disease expression and treatment effects. Additionally, a prospective GWAS from a relatively small sample has shown that people with comorbid schizophrenia and T2D have a higher genetic predisposition for both disorders than controls (Hackinger et al., 2018). Also, a small study of people with schizophrenia found an association between genetic predisposition for schizophrenia, insulin resistance and antipsychotic treatment response (Tomasik et al., 2019). Conversely, another relatively small study found no evidence of an association between genetic risk for T2D and schizophrenia (Padmanabhan et al., 2016). The main limitations of existing evidence are that studies remain relatively scarce and are likely underpowered owing to relatively small sample sizes.

Regarding the secondary analysis of GWAS data, genomic methods have been developed to examine for genetic similarity between traits. The most well-known and commonly used method is linkage-disequilibrium (LD) score regression (LDSC) (Bulik-Sullivan et al., 2015a), which examines for genetic correlation between traits by comparing the association between test statistics of genetic variants of each trait on their LD scores. Where LD is defined as the non-random association of alleles at different loci, the LD score of a genetic variant is the sum of LD r^2 measured with all other SNPs, and can be calculated in a reference sample of the same ethnicity when individual genotype data are not available for the GWAS sample (Ni et al., 2018). Previous studies have predominantly used LDSC to estimate whole-genome correlation between schizophrenia and cardiometabolic traits, with one recent study reporting evidence of partial genetic similarity between schizophrenia and BMI (Bahrami et al., 2020). However, there is limited evidence for other cardiometabolic traits (Bulik-Sullivan et al., 2015a).

Nevertheless, the LDSC approach may have limitations. First, LDSC could be susceptible to the 'missing heritability' problem, where subtle population stratification may bias the effects of relatively lower-frequency variants towards the null (Mathieson and McVean, 2012). Therefore, genetic correlation analysis which considers the relative frequency of variants is required. Second, LDSC estimates may be biased towards the null when opposing mechanisms exist (e.g., regions of positive and negative correlation nullifying each other when averaged (Shi et al., 2017)). Opposing mechanisms are likely to be relevant for a relatively heterogeneous condition like schizophrenia (Wolfers et al., 2018). Therefore, more fine-grained locus level genetic correlation analysis is required to identify genomic regions of interest. Third, while LDSC can provide evidence of overall genomic similarity between traits, it cannot provide information to consider biological plausibility, or infer potential causality. Therefore, methods that can distinguish between correlation and potential causation, and consider biological plausibility, are required.

1.2.3 Evidence for Inflammation as a Common Biological Mechanism for Comorbid Cardiometabolic Disorders and Schizophrenia

Emerging evidence indicates that inflammation could be relevant for the pathogenesis of cardiometabolic disorders and schizophrenia. Higher levels of circulating inflammatory markers are associated with schizophrenia and cardiometabolic disorders (Dandona et al., 2004, Upthegrove et al., 2014, Khandaker et al., 2014). Particularly, schizophrenia and cardiometabolic disorders share similar patterns of association with elevated concentrations of circulating inflammatory markers such as C-reactive protein (CRP) and interleukin-6 (IL-6), both cross-sectionally (Upthegrove et al., 2014, Wang et al., 2013) and longitudinally (Bowker et al., 2020, Khandaker et al., 2014).

Longitudinal research has also reported that inflammation may interact with disruption to glucose-insulin homeostasis to increase the risk of psychotic symptoms in young adults (Perry et al., 2018). Additionally, two independent longitudinal studies of clinical samples have shown that a combination of adverse inflammatory and cardiometabolic indices at baseline, including CRP and triglycerides, were associated with psychosis symptom severity and worse outcomes (Nettis et al., 2019, Osimo et al., 2021).

Longitudinal research conducted on clinical samples has also shown that FEP patients with higher CRP levels at baseline were more likely to develop hypertriglyceridaemia at three-month follow-up (Russell et al., 2015). Finally, MR studies have provided similar evidence suggesting that genetically-predicted levels of IL-6 and CRP could be causally related to cardiometabolic disorders (Georgakis et al., 2020) and schizophrenia (Hartwig et al., 2017) separately.

A potential shared role of inflammation in the pathogenesis of cardiometabolic disorders and schizophrenia is biologically plausible. For example, animal model studies have shown that neuroinflammation in the hypothalamus is associated with impaired peripheral insulin sensitivity (Zhang et al., 2008); and central infusion of tumour necrosis factor can impair the peripheral function of the insulin receptor (Arruda et al., 2011). Finally, MAPK pathways are associated with inflammatory regulation, particularly regarding inflammatory pathways involving IL-6 and CRP (Thalhamer et al., 2008). Activation of c-Jun-N-terminal Kinase, one of the MAPKs, phosphorylates the insulin receptor substrate, thus inhibiting insulin action at the cell membrane (Aguirre et al., 2000) and has been associated with the development of insulin resistance (Aguirre et al., 2002).

Post-mortem brain studies have found abnormal activity of the MAPK pathways in schizophrenia, (Kyosseva et al., 1999) and genetic modelling studies have found that genes implicated in MAPK pathways are associated with schizophrenia (Perez-Santiago et al., 2012). Additionally, animal models of schizophrenia have shown that maternal inflammation is associated with dose-dependent increases in MAPK phosphorylation in the striatum (Deng et al., 2011), and that treatment with antipsychotics can reverse these changes (Farrelly et al., 2015).

Early life factors may contribute to changes in the immune system and inflammatory pathways leading simultaneously to increased risk of cardiometabolic disorders and schizophrenia. This idea is consistent with the developmental programming hypothesis first proposed by British epidemiologist David Barker. This hypothesis posits that the early developmental environment can have implications far-ranging and far-reaching on the life-course (Barker et al., 1993).

Barker's early studies (Barker et al., 1993, Barker and Osmond, 1987, Barker et al., 1989) involved ecological comparisons of infant mortality rates in the early 20th century and mortality rates from adult CVD in the latter part of the 20th century in local authority regions of England and Wales. Barker discovered that the most impoverished, polluted, and crowded regions in the early 20th century had the highest infant mortality rates and adult CVD mortality rates decades later despite improvements to living conditions and pollution levels in the intervening period. Barker surmised that the surviving infants in the early 20th century were likely to have been exposed to similar environmental conditions to those who died, and this could help to explain the excess adult mortality decades later. Barker proposed that there may be a critical developmental period in early life that, if disrupted, could predispose to adult disease.

This early work has paved the way for discoveries that disruptions to prenatal and early life conditions are strongly associated with risks of developing obesity (Entringer et al., 2012), hypertension (Ojeda et al., 2008), metabolic syndrome (Rinaudo and Wang, 2012), T2D (Yajnik, 2010) and CVD

(Alexander et al., 2015) in adulthood. These findings have been extended to include various neurodevelopmental conditions, including schizophrenia (Knuesel et al., 2014).

Disruption to the early-life environment may permanently alter the function of the immune system, and this may be a putative mechanism that could link cardiometabolic disorders and schizophrenia. For example, poor intrauterine nutrition such as vitamin D deficiency is associated with a higher risk of schizophrenia (Eyles et al., 2018) and CVD in later life (Sauder et al., 2019). Vitamin D holds important roles in the development and regulation of the immune system, and intrauterine vitamin D deficiency is associated with alterations in immune function in adulthood (Harvey et al., 2010).

Additionally, prenatal infection is associated with an increased risk of schizophrenia and CVD in the offspring (Mazumder et al., 2010, Khandaker et al., 2012, Khandaker et al., 2013), either through direct effects of the pathogen on the foetus after crossing the placenta or indirectly via activation of the maternal immune system (Hsiao and Patterson, 2011). Prenatal infection is also associated with lasting changes to offspring immune function (Pedersen et al., 2019).

Furthermore, prenatal and early-life stressful life events (SLEs) are also associated with an increased risk of developing schizophrenia and CVD in adulthood (Kershaw et al., 2014, Malaspina et al., 2008). Prenatal SLEs are also associated with lasting alterations to the immune system in the offspring (Merlot et al., 2008).

Genetic influences may also lead to permanent alterations of the immune system and an increased risk of cardiometabolic disorders and schizophrenia. For example, genetic variation in the *IL-6R* gene is associated with changes to CRP levels and a higher risk of heart disease in later life (Swerdlow et al., 2012, Georgakis et al., 2020). The same genetic variation can influence the risk of schizophrenia in adulthood (Hartwig et al., 2017). Genetic correlation studies have shown evidence for an overlap between cardiometabolic and inflammatory traits (Wu et al., 2014) . They have also identified common-causal risk genes for immune changes and increased risk of cardiometabolic disorders in adulthood (Nath et al., 2019). Genetic studies have also identified the potential for common genetic variants, which could simultaneously increase the risk for schizophrenia and cardiometabolic disorders. Several are related to the immune system (So et al., 2019).

However, large-scale genetic studies examining the role of inflammation on the simultaneous risk of comorbid cardiometabolic disorders and schizophrenia remain relatively scarce. Therefore, putative mechanisms must at present be extrapolated largely from studies examining either the genetic overlap of inflammation and cardiometabolic disorders, or inflammation and schizophrenia.

In summary, there is evidence from observational, genetic, and animal model studies suggesting a biologically plausible association of inflammation as a potential biological mechanism for comorbid

cardiometabolic disorders and schizophrenia. However, observational studies are limited by the risk of residual confounding, and a scarcity of longitudinal research has prevented an examination of the temporal role that inflammation might play in the comorbidity between cardiometabolic disorders and schizophrenia. Furthermore, while genetic studies involving methods such as MR or LDSC can help to show evidence of potential causality or genetic overlap, current studies have not examined schizophrenia with cardiometabolic and inflammatory traits simultaneously.

1.3 Approaches to Improving The Prediction of Cardiometabolic Comorbidity of Schizophrenia

1.3.1 An Introduction to Prognostic Research

Having discussed the links between cardiometabolic and schizophrenia spectrum disorders, including the potential for commmon biological mechanisms, I now turn to the clinical prediction of cardiometabolic disorders in young people with psychosis.

At its essence, prognostic research can be distinguished from traditional observational research in that it analyses at the individual rather than the group level (Breiman, 2001). Prognostic research deals in estimating the accuracy with which a prediction model, usually a regression equation consisting of weighted sums of predictors, can estimate the probability of an outcome occurring (Moons et al., 2009). Preferably, these estimations are achieved by first fitting the regression equation in a model development sample and then testing the equation in similar unobserved individuals separated by geography, time, or clinical setting (Altman et al., 2009). This external validation step is fundamental for prognostic research since risk prediction models can only be useful if they are generalisable (Altman et al., 2009).

Replication is also central to observational research, since observational studies are usually conducted on population sub-samples. Therefore, replication in observational findings helps to account for inaccuracies due to sampling variability and helps to show consistency (Casella, 2002). However, in prognostic research, the bar for generalisation is raised since it goes beyond replication that amounts to testing the same association twice (Bzdok et al., 2021). For example, showing that an exposure is associated with a disease in a second patient sample does not mean that this same exposure can tell health and disease apart at the individual level (Bzdok and Ioannidis, 2019).

Prognostic research is increasingly relevant for a diverse range of disease states to improve outcomes for those affected (Riley, 2019). For example, more people live with one or more disease or health-impairing conditions than ever before, putting strain on already stretched resources. Consequently, there is increasing interest in prognostic research at the level of the clinician, who is interested in the long-term interests of the patient; the commissioner who is interested in future service planning; and the politician who is interested in international health comparisons (Riley, 2019).

In the UK general population, risk prediction algorithms are commonly used to identify high-risk individuals for tailored interventions from baseline demographic, lifestyle, and clinical information. The UK National Institute for Health and Care Excellence (NICE) first published guidance on the use of risk prediction algorithms for cardiometabolic and cardiovascular risk assessment in March 2003

(National Institute for Health and Care Excellence., 2003). In 2010 the guidance was updated to specify that the QRISK algorithm (Hippisley-Cox et al., 2007), developed to predict the 10-year risk of CVD, should be used preferentially (National Institute for Health and Care Excellence., 2010).

The QRISK family of cardiometabolic risk prediction algorithms are therefore good examples of prognosis research that are integrated into routine clinical practice. However, the positive example set by the QRISK family of algorithms is seemingly rarely replicated. A systematic review of cardiometabolic risk prediction algorithms in 2016 found that the literature is "overwhelmed with models for predicting the risk of cardiovascular outcomes in the general population" (Damen et al., 2016). The review concluded that the reporting quality of most algorithms was poor, very few algorithms were externally validated, and almost none were assessed for their impact or uptake in clinical practice (Damen et al., 2016).

1.3.2 Cardiometabolic Risk Prediction in Young People with Psychosis

As outlined in detail throughout this introduction, young people who have psychosis are at high risk for developing cardiometabolic disorders. Therefore, there is a clear and crucial need for accurate clinical tools to predict cardiometabolic risk in this population, to optimise care and improve long-term outcomes. However, due to antipsychotic medications (see Section 1.1.3.1), a higher prevalence of most traditionally attributed lifestyle factors compared with the general population (see Section 1.1.3.2), and the possibility of intrinsic biological or genetic links (see Section 1.2), there are likely to be tangible differences in the type, balance, and sum of cardiometabolic risk factors which affect young people who have psychosis compared with the general population.

Such differences are likely to result in differences in baseline cardiometabolic risk, the ideal balancing of predictor weights, and in the choice of predictors. For example, the prescription of cardiometabolic risk-increasing antipsychotic medications is likely to be of prime importance in predicting cardiometabolic risk in young people with psychosis, but is unlikely to be important in the general population since such medications are rarely prescribed in that setting.

Therefore, it is unclear whether existing cardiometabolic risk prediction algorithms developed for the general population are likely to be suitable for use in young people who have psychosis. A recent study of a small sample of people with chronic schizophrenia found that commonly used general population cardiometabolic risk prediction algorithms, including QRISK, returned significantly different risk scores when tested on the same participants (Berry et al., 2018). This calls into question the reliability and suitability of such algorithms for relatively older people with chronic schizophrenia, let alone young people with psychosis. Indeed, no studies have sought to examine the

predictive accuracy of existing cardiometabolic risk prediction algorithms in young people who have psychosis, even though current guidance recommends the routine use of the QRISK algorithm in this population (Royal College of Psychiatrists., 2020).

1.4 Aims of the Analytic Work Presented in this Thesis

1.4.1 Section B – Examining the Nature of Association between Cardiometabolic Disorders and Schizophrenia

In Section B of this thesis, I have aimed to establish a more detailed understanding of the nature of association between cardiometabolic disorders and schizophrenia. Specifically, I have sought to test whether cardiometabolic dysfunction could be a cause or consequence of schizophrenia. A more detailed understanding of the nature of association between cardiometabolic disorders and schizophrenia could lead to improvements in the treatment of the cardiometabolic comorbidity of schizophrenia and could contribute toward closing the mortality gap of 10-15 years faced by people who have the illness (Plana-Ripoll et al., 2020).

To summarise, existing research to date has shown that: a) chronic schizophrenia is strongly associated with a range of cardiometabolic disorders including T2D, obesity and CVD, and this leads to a shortened life expectancy; b) FEP is associated more strongly with measures of aberrant glucose-insulin homeostasis than other cardiometabolic traits; c) limited research has shown that at-risk mental states are associated with altered cardiometabolic indices, particularly relating to glucose-insulin homeostasis.

However, existing research is limited for three key reasons. First, current studies have predominantly been cross-sectional or have included prevalent cases of schizophrenia spectrum disorders. Therefore, existing research cannot appropriately test the direction of association between cardiometabolic and psychiatric traits. Second, most existing research in the field has not appropriately addressed the risk of confounding by sociodemographic, lifestyle or treatment factors. Third, current studies have primarily included one-off measurements of cardiometabolic indices, overlooking the potential for dynamic temporal changes in these markers. Longitudinal repeated measurements could provide a more reliable measure of underlying biological mechanisms and could identify population sub-groups.

I have aimed to address each of these limitations using data from the Avon Longitudinal Study of Parents and Children (ALSPAC), a population-representative UK birth cohort. Using ALSPAC data, I aimed to (1) delineate longitudinal trajectories of fasting insulin and BMI from repeated measurements of these indices between ages 1-24 years in the total ALSPAC sample; (2) examine the clinical and biochemical characteristics of the identified trajectories; (3) test associations of cardiometabolic developmental trajectories with psychosis at age 24 years, before and after adjusting for a set of key potential sociodemographic and lifestyle confounders. I chose fasting insulin and BMI since they are markers of distinct pathways and have shown the strongest associations with early

schizophrenia-spectrum phenotypes in young adults (Perry et al., 2018, Zammit et al., 2007). To test the specificity of association, To test specificity of association, I also included depression as an outcome at age 24 years since depression has also shown strong associations with cardiometabolic traits such as T2D (Kan et al., 2016), obesity (Gibson-Smith et al., 2020) and CVD (Khandaker et al., 2019).

1.4.2 Section C – Testing Potential Mechanisms of Association between Cardiometabolic Disorders and Schizophrenia

In Section C of the thesis, I have aimed to examine potential mechanisms by which cardiometabolic traits could be associated with schizophrenia, over and above the traditional attributions of sociodemographic, lifestyle and clinical factors. I have focused on the potential roles of shared genetic liability and inflammation. A more detailed understanding of the mechanisms of association between cardiometabolic disorders and schizophrenia could lead to pathophysiological insights into the cardiometabolic comorbidity of schizophrenia and possibly schizophrenia itself. A more detailed mechanistic understanding could help identify novel therapeutic targets for schizophrenia and its associated cardiometabolic comorbidity.

To summarise, existing research to date has: a) shown some evidence that shared genetic liability may be responsible for the development of comorbid cardiometabolic disorders and schizophrenia; b) reported consistently on the potential biological role of inflammation in the pathogenesis of cardiometabolic disorders and schizophrenia separately; c) reported heterogeneously regarding the potential causal role of genetically-predicted cardiometabolic traits with schizophrenia.

However, existing research on the mechanisms of association between cardiometabolic disorders and schizophrenia is limited for three key reasons. First, existing prospective research examining for genetic overlap between cardiometabolic traits and schizophrenia has included small sample sizes and so may be limited in statistical power. Second, existing secondary studies of GWAS datasets may be limited due to methodological shortcomings and in its consideration of biological plausibility. Third, while a wealth of research has accumulated on the possible pathophysiological role of inflammation in both cardiometabolic disorders and schizophrenia, most studies have not included schizophrenia, cardiometabolic and inflammatory traits simultaneously to test this hypothesis.

I have aimed to address each of these limitations across three studies. In the first study, I aimed to use data from the relatively large ALSPAC birth cohort to examine whether (1) genetic predisposition for schizophrenia was associated with insulin resistance at age 18 years, before and after adjusting for relevant confounders; (2) genetic predisposition for T2D was associated with risk of psychosis at

age 18 years, before and after adjusting for relevant confounders; (3) these associations may be mediated by genetic influences on childhood inflammatory markers. In the second study, I performed an analysis of summary data from large-scale GWAS to rigorously examine for the potential of a common genetic basis for schizophrenia, cardiometabolic and inflammatory traits, using a range of complementary genomic approaches which can address the key methodological shortcomings of previous research. In the third study, I performed bidirectional and multi-variable two-sample MR analysis of summary GWAS data to examine whether: (1) insulin resistance-related cardiometabolic traits could have a potentially causal role in schizophrenia pathogenesis or *vice versa*; (2) inflammation could be a common mechanism linking insulin resistance and schizophrenia.

1.4.3 Section D: Improving the Prediction of Cardiometabolic Risk in Schizophrenia

In Section D of the thesis, I turned to the clinical significance central to Sections A&B, which is the prediction of cardiometabolic risk in young people with psychosis. Accurate prediction of cardiometabolic risk in young people with psychosis is a fundamental step toward reducing its significant short, medium, and long-term impact upon the lives of people who have schizophrenia.

To summarise existing research in the field, it is long established that the best way to address cardiometabolic disorders is with primary prevention and intervening early to slow or prevent progression to more distal, chronic and deadly disease endpoints (Chrysant, 2011). Given the cardiometabolic associations of schizophrenia that I have described in the introduction and have tested in Sections B & C of this thesis, this means intervening at the earliest possible opportunity in young people at the onset of psychotic illness. In the general population, risk prediction algorithms have been developed to predict an individual's probability of developing an adverse cardiometabolic outcome years in advance. Healthcare professionals can then use the risk estimates to tailor interventions in the intervening period to reduce the probability of adverse cardiometabolic outcomes occurring.

Given the tangible differences in baseline cardiometabolic risk and the differences in lifestyle and treatment factors between people who have schizophrenia and the general population, it is unlikely that tools developed for the general population will be suitable for the schizophrenia population. Indeed, research has shown that existing tools developed for the general population return extremely variable risk estimates when tested on older adults with chronic schizophrenia (Berry et al., 2018), let alone young people at the onset of their psychotic illness.

Therefore, I first aimed to gain a comprehensive understanding of the current field of prognostic research for cardiometabolic disorders. I have done this by performing a systematic review of existing

cardiometabolic risk prediction algorithms developed either for the general or psychiatric populations and assessing whether any existing algorithm is likely to be suitable for young people who have psychosis. Furthermore, I aimed to quantify the predictive performance of potentially suitable algorithms identified from the systematic review by testing their predictive performance in a sample of young people with or at risk of developing psychosis, using ALSPAC data.

Next, using patient data from three psychosis early intervention services (EIS), I aimed to develop and externally validate The Psychosis Metabolic Risk Calculator (PsyMetRiC), a cardiometabolic risk prediction algorithm developed and tailored specifically for young people with psychosis.

SECTION B

EXAMINING THE NATURE OF ASSOCIATION BETWEEN CARDIOMETABOLIC DISORDERS AND SCHIZOPHRENIA

Section B Summary

This section addresses the temporality of association between cardiometabolic disorders and schizophrenia by analysing prospective data from the ALSPAC birth cohort (Chapter 2). This study examined: a) whether disrupted cardiometabolic indices could be a cause or consequence of the mental disorder by testing the direction of association between disrupted cardiometabolic indices and psychosis; b) the specificity of association between disrupted cardiometabolic indices and psychosis; c) whether confounding by sociodemographic or lifestyle factors could explain any associations.

I used a growth mixture modelling approach to delineate developmental trajectories of fasting insulin and BMI from 5,790 and 10,463 ALSPAC participants, respectively. Fasting insulin was measured at four time-points (ages 9, 15, 18, and 24 years), and BMI was measured at twelve (ages 1, 2, 3, 4, 7, 9, 10, 11, 12, 15, 18 and 24 years). I used regression analyses to examine the sociodemographic, biochemical, and clinical characteristics of the identified trajectories. Next, I tested longitudinal associations between the identified trajectories and the risk of schizophrenia-spectrum and depression phenotypes at age 24 years. I included depression as an outcome to test the specificity of association because depression shows strong genetic (Anttila et al., 2018) and phenotypic (Buckley et al., 2009) overlap with schizophrenia and has similar associations with cardiometabolic disorders (Firth et al., 2019). I adjusted for a detailed range of potential confounders, including sex, ethnicity, social class, childhood emotional and behavioural problems, and cumulative scores of sleep problems, average calorie intake, physical activity, smoking, alcohol, and substance use in childhood/adolescence.

To the best of my knowledge, this is the first longitudinal study that is sufficiently able to examine the direction of association between cardiometabolic traits and psychosis and the first study to model the cardiometabolic exposures as repeated measurements through childhood/adolescence.

This study presents evidence that disruptions to glucose-insulin homeostasis may predate the onset of psychosis. The study also presents evidence suggesting that disrupted glucose-insulin homeostasis may be specific to psychosis. The associations persisted after adjusting for a detailed range of potential confounders, suggesting that disrupted glucose-insulin homeostasis could be a risk factor for psychosis.

Findings from this study have been published in *JAMA Psychiatry* (Perry et al., 2021b). See Appendix B for the published manuscript.

Chapter 2

Longitudinal Trends in Insulin Levels and BMI From Childhood and Their Associations with Risks of Psychosis and Depression in Young Adults in the ALSPAC Birth Cohort

2.1 Introduction

Cardiometabolic disorders commonly co-occur with depression and schizophrenia (Firth et al., 2019), leading to a reduced quality of life, increased healthcare costs (Naylor, 2012) and a shortened life expectancy (Laursen et al., 2019, Plana-Ripoll et al., 2019). This comorbidity is usually attributed to chronic lifestyle factors (e.g. physical inactivity or smoking) or the adverse effects of psychotropic medications (Leucht et al., 2013). However, meta-analyses report altered glucose-insulin homeostasis in relatively young drug-naïve first-episode psychosis patients (Perry et al., 2016, Pillinger et al., 2017a). Similarly, reports from population-based longitudinal studies suggest a bidirectional association between depression and CVD (Penninx et al., 2001, van Melle et al., 2004). Together, this evidence indicates that cardiometabolic and psychiatric conditions may share pathophysiologic mechanisms. However, three key issues remain.

First, existing studies have predominantly included prevalent depression or psychosis cases, so cannot appropriately test the direction of association between cardiometabolic and psychiatric phenotypes. Second, most existing research in the field has not appropriately addressed the risk of confounding by sociodemographic, lifestyle or treatment factors. Third, studies have primarily included one-off measures of cardiometabolic indices, overlooking dynamic temporal changes in these markers. Longitudinal repeated measurements could provide a more reliable assessment of underlying homeostatic mechanisms and could identify population sub-groups. For example, aberrant trajectories of childhood BMI are associated with adult cardiometabolic disorders (Buscot et al., 2018). While cardiometabolic function encompasses a broad range of parameters, two pathways, insulin sensitivity and adiposity, are of particular interest regarding psychosis and depression. Previous genetic studies indicate distinct associations of BMI with depression (Tyrrell et al., 2019) and fasting insulin with schizophrenia (Li et al., 2018). However, to the best of my knowledge, no studies have examined whether fasting insulin and BMI trajectories from childhood are associated with adult psychosis and depression.

2.2 Aims and Objectives

Using data from ALSPAC, I aimed to: (1) delineate longitudinal trajectories of fasting insulin and BMI from repeated measurements between 1-24y; (2) examine the characteristics of identified trajectories; (3) test associations with risks of psychosis and depression at age 24 years, in the total sample and two sexes separately. I hypothesised that altered cardiometabolic development from childhood would be associated with increased risks for depression and psychosis in adulthood.

2.3 Methods

2.3.1 Description of cohort and sample

ALSPAC initially recruited 14,541 pregnant women resident in southwest England, with expected delivery dates between 1.4.1991-31.12.1992, resulting in 14,062 live births (Boyd et al., 2013b, Fraser et al., 2013, Northstone et al., 2019). An additional 913 participants were recruited subsequently. See www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/ for a fully searchable data dictionary. Data were collected and managed using REDCap (University of Bristol (Harris et al., 2019, Harris et al., 2009)). ALSPAC Ethics and Law Committee and Local Research Ethics Committees provided ethical approval for the study. All participants provided informed consent. Modelling of fasting insulin and BMI trajectories were based on 5,790 and 10,463 participants, respectively. See Figure 1 of Appendix B for a flow-chart of participants in the study. Missing exposure data was handled using full-information maximum likelihood (FIML) estimation, as FIML estimates parameters directly using all the information that is already contained in the incomplete data set (Dong and Peng, 2013). FIML has demonstrated to produce unbiased estimates (Enders, 2001b) and valid model fit information (Enders, 2001a).

2.3.2 Measurement of Exposures

2.3.2.1 Fasting insulin

Fasting insulin was measured at ages 9 (n=894), 15 (n=3484); 18 (n=3286); and 24 (n=3253) years using the ultrasensitive ELISA (Mercodia, Uppsala, Sweden) automated microparticle enzyme immunoassay, that does not cross-react with proinsulin. Its sensitivity was 0.07 mU/L, and inter- and intra-assay coefficients of variation were <6%. Fasting blood samples were drawn at 0900 after a 10-hour fast, then spun and stored at -80°C. There was no evidence of freeze-thaw cycles during storage.

2.3.2.2 BMI

BMI was measured at ages 1 (n=1236); 2 (n=1036); 3 (n=1050); 4 (n=1018); 7 (n=8200); 9 (n=7633); 10 (n=7465); 11 (n=7100); 12 (n=6704); 15 (n=5415); 18 (n=5061) and 24 (n=3975) years.

2.3.3 Measurement of Psychiatric Outcomes at Age 24

2.3.3.1 Schizophrenia Spectrum Outcomes

2.3.3.1.1 Psychotic Experiences (PEs)

PEs were identified through the semi-structured Psychosis-Like Symptom Interview (PLIKSi) conducted by trained psychology graduates and coded per the definitions in the Schedules for Clinical Assessment in Neuropsychiatry, V2.0. The PLIKSi had good interrater (Intraclass correlation: 0.81; 95% CI, 0.68-0.89) and test-retest (0.9; 95% CI 0.83-0.95) reliability (Sullivan et al., 2020). PEs, occurring in the last six months, covered the three main positive symptom domains: hallucinations, delusions, and thought interference. After cross-questioning, interviewers rated PEs as absent, suspected, or definite. I included cases of definite PEs; the comparator group was suspected/absent PEs.

2.3.3.1.2 Psychosis At Risk Mental State (ARMS)

Cases of ARMS were identified by mapping PLIKSi data to Comprehensive Assessment of At-Risk Mental State (CAARMS) criteria (Yung et al., 2005). Cases were defined as participants meeting CAARMS criteria for attenuated psychosis (symptoms not reaching the psychosis threshold due to intensity or frequency) or brief limited intermittent psychosis (frank psychotic symptoms that resolved spontaneously within one week).

2.3.3.1.3 Psychotic Disorder

Cases of psychotic disorder were defined (Sullivan et al., 2020) as definite PEs that were not attributable to sleep/fever, had occurred >once per month over the previous six months, and were either (i) very distressing, (ii) negatively impactful on social/occupational functioning, (iii) led to professional help-seeking. I also included participants meeting the criteria for CAARMS psychotic disorder (threshold psychotic symptoms for >1 week).

2.3.3.1.4 Negative Psychotic Symptoms Score

Ten questions from the Community Assessment of Psychic Experiences questionnaire (Stefanis et al., 2002) were administered covering interest, motivation, emotional reactivity, pleasure, and sociability. Participants rated each item 0=never; 1=sometimes; 2=often; and 3=always. I recoded the

variables by coding ‘always’ and ‘often’ as 1; ‘never’ and ‘sometimes’ as 0, and then summed giving a total score of 0-10.

2.3.3.2 Depression Outcomes

Depression was measured using the computerised Clinical Interview Schedule–Revised (CIS-R) (Lewis et al., 1992). The CIS-R assesses symptoms of depression occurring in the past week and provides a diagnosis of depressive episode based on the International Statistical Classification of Diseases (ICD), 10th Revision criteria, which I used as a binary outcome (ICD-10 codes F32.0-32.2). I also included a CIS-R depression severity score - comprising scores for mood, thoughts, fatigue, concentration, and sleep - as a continuous outcome.

2.3.4 Assessment of Potential Confounders

2.3.4.1 Sociodemographic Confounders

I included sex at birth, ethnicity, and paternal social class. Sex was recorded at birth (binary variable). Ethnicity was recorded from participant-completed questionnaire data and coded as White vs. non-White. Paternal social class was recorded from participant-completed questionnaire data based on occupation as per the UK Office of National Statistics classification system: I, II, III non-manual, III manual, IV, V).

2.3.4.2 Lifestyle Confounders

I included childhood emotional and behavioural problems and cumulative scores of smoking, physical activity, alcohol use, substance use, sleep problems and average calorie intake. Childhood emotional and behavioural problems were assessed at age 7 years via the Strength and Difficulties Questionnaire (SDQ) (Goodman, 2001), which screens for emotional symptoms, hyperactivity/inattention and peer relationship problems, and summed into a ‘total difficulties score’, which I used as an adjustment variable. However, due to a considerable reduction in the available sample size when the SDQ ‘total difficulties score’ was included as an adjustment variable, I used the *k*-nearest neighbours (Knn) imputation algorithm of the VIM package (Kowarik, 2016) in R (using recommended settings) to replace missing data for the SDQ variable only. The Knn algorithm is sensitive and robust to different data types and performs comparatively well to other imputation methods such as multiple imputation using chained equations (Schmitt, 2015, Liao, 2014). I used Knn imputation in place of multiple

imputation methods since where the former produces a single imputed dataset; the latter produces multiple imputed datasets, which would have led to significant and prohibitive computation burden coupled with the three-step method of analysis.

Smoking (on average >1 cigarette each day) was coded as a binary variable at ages 15, 18 and 24 years from participant-completed questionnaire data and summed to create a cumulative smoking score between ages 15-24 years of 0-3.

Physical activity (averaged over the past year) was recorded from participant-completed questionnaire data at ages 15, 18 and 24 years as 0=never, 1=less than once per month; 2=one to three times per month; 3=one to four times per week; 4= five or more times per week. I summed the three variables creating a cumulative physical activity score between ages 15-24 years of 0-12.

Alcohol use was coded as a binary variable (>1 alcoholic beverage on average each week) at ages 12, 15, 18 and 24 years from participant-completed questionnaire data. I summed the four variables creating a cumulative alcohol use score between ages 12-24 years of 0-4.

Substance use was coded as a binary variable at ages 12, 15, 18 and 24 years. At age 12 years, the self-report questionnaire asked whether the participant had ever taken any illicit substance. At ages 15 and 18 years, the self-report questionnaire asked whether the participant had taken any illicit substance in the past year. At age 24 years, the self-report questionnaire asked how many illicit substances the participant had taken in the past year. I recoded the age 24 variable as a binary variable, with a score of 1 if the participant recorded taking at least one illicit substance in the past year. I summed the four variables creating a cumulative substance use score between ages 12-24 years of 0-4.

Sleep problems were coded as binary variables at ages 7, 8, 9 and 14 years from questionnaire data completed by the primary caregiver and 15 years completed by the participant. At ages 7 and 9 years, the primary caregiver was asked whether the participant had difficulty sleeping in the past year, and at ages 8, 14 and 15 years, the same question was asked with a duration of the past month. I summed the five variables creating a cumulative sleeping difficulties score between ages 7-15 years of 0-5.

Average calorie intake was assessed at ages 7, 10 and 13 years via a food frequency questionnaire, sent to the primary caregiver a week before the child's clinic appointment. The primary caregiver was asked to record everything the child ate or drank for three days, including one weekend day. When they brought the child to the clinic appointment, they were interviewed by a trained member of the nutrition team to ensure the completeness of the record concerning the type of food/drink and the amount consumed. At each age, average daily calorie (kcal) intake was recorded. I standardized (z-

scores) the three variables and summed them together, creating a cumulative average calorie intake score between ages 7-13 years.

2.3.5 Statistical Analysis

2.3.5.1 Delineating trajectories of fasting insulin and BMI

I standardized (z-transformed) fasting insulin and BMI separately in males and females, then combined to delineate trajectories individually for fasting insulin and BMI using growth mixture modelling (GMM) (Ram and Grimm, 2009). I used z-scores to measure relative change in fasting insulin and BMI since BMI increases in all young people during early life.

GMM was run iteratively whilst increasing the number of trajectory classes to fit. Estimates of the Bayesian Information Criterion (BIC), entropy, Vuong-Lo-Mendell-Rubin Likelihood Ratio Test (VLMR-LRT) and Parametric Bootstrap Likelihood Ratio Test (BLRT) were recorded at each iteration, along with a visual inspection of graphical outputs. Once achieving successful convergence, checks were performed to rule out local solutions by replicating the estimation using the same seed values and comparing model parameter estimates for replication. A successfully converged model with no local solutions would have the best loglikelihood values repeated (Jung, 2007). In selecting the optimum class solution, I aimed to choose the solution with the lowest BIC, suitable statistical evidence (p<0.05) in VLMR-LRT and BLRT tests (suggesting the solution with n trajectories is an improvement over the solution with $n-1$ trajectories), and high entropy values (close to 1.0). Also, I aimed to include no less than 1% of the total sample in a particular trajectory (Jung, 2007) to allow suitable statistical power in subsequent analysis.

Since the sample size for fasting insulin at age 9 years was smaller, I repeated GMM without data from that time-point and compared the characteristics of the resultant trajectories. Analyses were conducted using MPlus Version 8 (Muthen, 2017) and R (R Core Team, 2017). P-values were corrected for multiple testing using the Holm-Bonferroni method (Holm, 1979) for the six psychiatric outcomes. I estimated how participants overlapped between fasting insulin and BMI trajectories (the most common and highest-risk) using the phi-coefficient.

2.3.5.2 Associations of Sociodemographic, Lifestyle and Clinical Factors with Trajectories

I used the three-step method (Asparouhov, 2014) to estimate associations of sociodemographic, lifestyle and clinical factors with trajectory membership. The three-step method allows class separation unaffected by auxiliary variables, retains and includes information on class uncertainty, and is robust when entropy is >0.60 (Asparouhov, 2014).

After establishing the optimum number of trajectories, the second step is to calculate classification uncertainty, which is computed as a natural log of the average latent class probabilities for most likely class membership and the number of observations per trajectory class. These logits are used in the third step, including regression on predictors of trajectory class membership (using trajectory class membership as an outcome) or regression of trajectory classes on an outcome (using trajectory class membership as a predictor). Detailed information on the statistical methodology underpinning the three-step method alongside data simulations are available elsewhere (Asparouhov, 2014).

Multinomial regression estimated ORs and 95% confidence intervals (CI) for the associations of sociodemographic/lifestyle factors with fasting insulin and BMI trajectories, compared with the most common trajectory. I considered time-invariant (sex, ethnicity, social class at birth, family history of CVD, gestational age, birthweight, perinatal stressful life events) and time-variant (physical activity and smoking in adolescence/early adulthood) factors.

The coding and description of sex, ethnicity, social class at birth, physical activity and smoking is presented in Section 2.3.4. A positive family history of cardiometabolic disorders was coded from self-report questionnaire data encompassing T2D, hypercholesterolaemia, or CVD. Stressful life events (SLEs) were based on self-report questionnaire data comprising a summed total of up to 42 pre-specified life events affecting the mother at 18- and 36-weeks gestation and the participant at 8-weeks and 6-months postpartum. Examples included loss of a partner or family member, loss of employment, moving-house or financial difficulty. A complete list of the 42 SLEs is reported elsewhere (Kingsbury et al., 2016). I compared the top tertile of summed SLE scores vs the bottom tertile. Birthweight and gestational age were coded as continuous variables derived from questionnaire data. ORs represent the increase in risk of trajectory membership per SD increase in factor.

Next, I examined the clinical phenotype of trajectories at age 24 years, examining mean levels of commonly measured clinical and biochemical factors for participants, grouped by most-likely trajectory membership. I included measures of BMI, waist circumference (cm, assessed during clinical assessment), FPG, HDL, LDL (all mmol/L), fasting insulin (μIU/mL), and CRP (mg/L). All biochemical samples were taken at 0900 during clinic assessment from consenting participants,

following a 10-hour fast (water only). I present mean values for waist circumference separately for males and females since the reference ranges are different.

Next, I used logistic regression to estimate the association of trajectory membership with an age-appropriate cardiometabolic outcome, metabolic syndrome at age 24 years. Metabolic syndrome was defined using the most recent harmonized consensus definition (Alberti et al., 2009) (see Table 2).

2.3.5.3 Associations of Cardiometabolic Trajectories with Risk of Psychiatric Outcomes

Logistic regression was used to estimate ORs and 95% CIs for binary outcomes per trajectory compared with the most common trajectory via the three-step method. Linear regression for continuous outcomes estimated β-coefficients and 95% CIs representing the SD increase in the risk of outcome per trajectory. I tested associations for the total sample and then separately for males and females, before and after adjusting for potential confounders. Regression models for negative symptoms were additionally adjusted for depressive symptoms and *vice versa*.

2.4 Results

2.4.1 Trajectories of Fasting Insulin from Childhood to Young Adulthood

Based on 5,790 participants (45.9% male), the three-trajectory solution was optimum (Table 3), representing 'stable average' (Class 1; 77.8%, *n*=4,939), 'minor increase' (Class 2; 19.0%, *n*=693), and 'persistently high' (Class 3; 3.1%, *n*=158) fasting insulin trajectories between ages 9-24 years (Figure 3A). See Appendix B Figures 2A-C for trajectory means and individual values per developmental trajectory of fasting insulin. The trajectories were similar after excluding age 9 data (Appendix B Figure 3).

Table 3: Growth Mixture Model Fit Indices for Fasting Insulin

***n* Trajectories**	**BIC**	**Entropy**	**VLMR-LRT (*p*-value)**	**BLRT (*p*-value)**
1	76474	-	-	-
2	69389	0.957	0.007	<0.001
3†	66304	0.853	0.034	<0.001
4*	67872	0.750	0.253	0.042
5*	67688	0.836	0.319	0.114
6*	67521	0.729	0.409	0.440

BIC = Bayesian Information Criterion; VLMR-LRT = Vuong-Lo-Mendell-Rubin Likelihood Ratio Test; BLRT = Parametric Bootstrap Likelihood Ratio Test; **Contained one trajectory with <1% of sample;* †Selected for further analysis

2.4.2 Trajectories of BMI from Childhood to Young Adulthood

Based on 10,463 participants (49.0% male), the five-trajectory solution was optimum (Table 4), representing 'stable average' (Class 1; 71.1%, *n*=8,383), 'gradually decreasing' (Class 2; 7.0%, 949), 'puberty-onset minor increase' (Class 3; 14.5%, *n*=668), 'puberty-onset major increase' (Class 4; 1.9%, *n*=174), and 'persistently high' (Class 5; 5.5%, *n*=289) BMI trajectories between ages 1-24 years (Figure 3B). See Appendix B Figure 4A-E for trajectory means and individual values per trajectory of BMI.

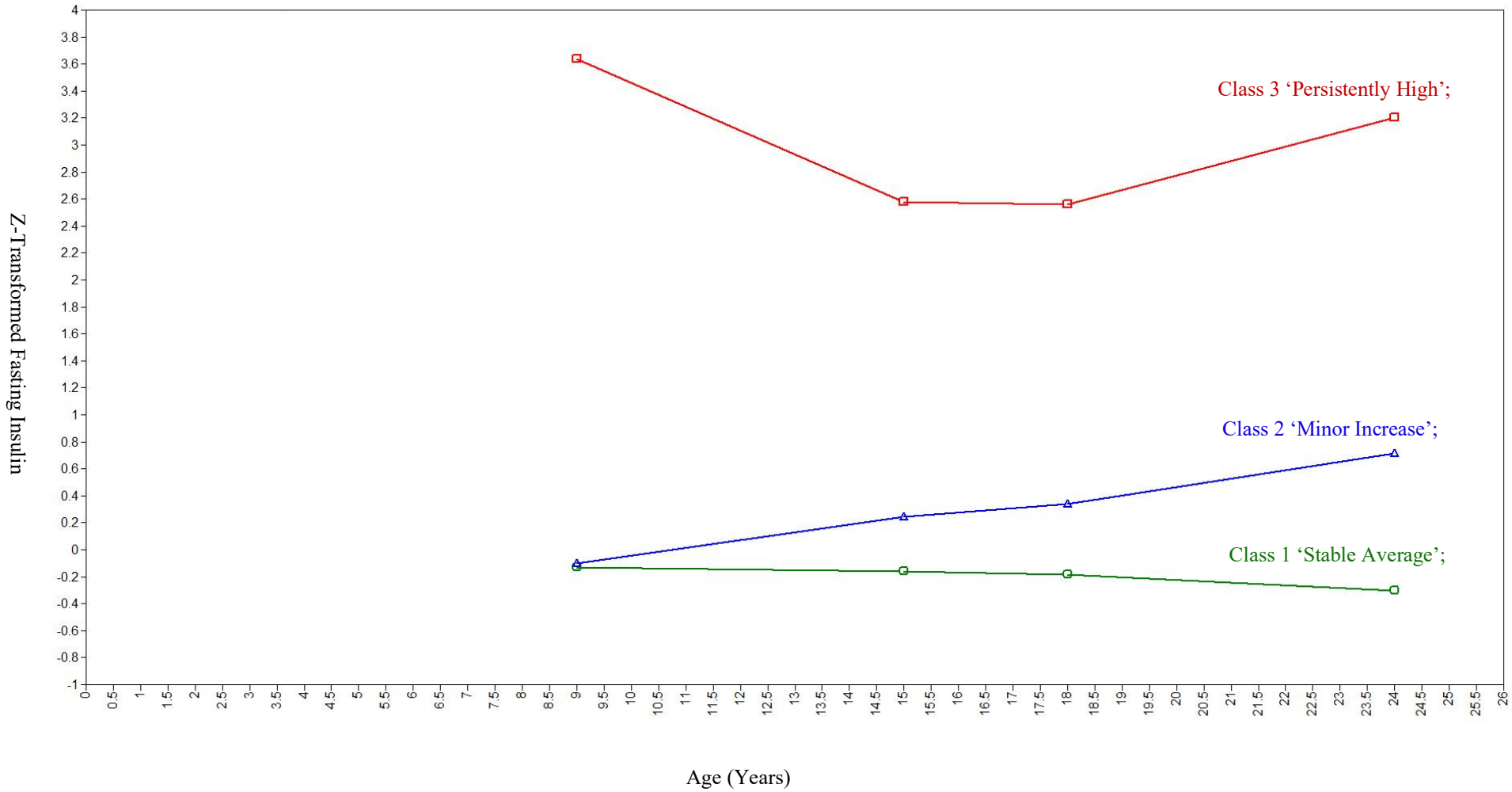
Table 4: Growth Mixture Model Fit Indices for Body Mass Index

***n* Trajectories**	**BIC**	**Entropy**	**VLMR-LRT (*p*-value)**	**BLRT (*p*-value)**
1	223514	-	-	-
2	224574	0.663	<0.001	<0.001
3	222745	0.774	<0.001	<0.001
4	222142	0.768	0.029	<0.001
5†	221575	0.885	0.010	<0.001
6*	221138	0.766	0.102	0.073

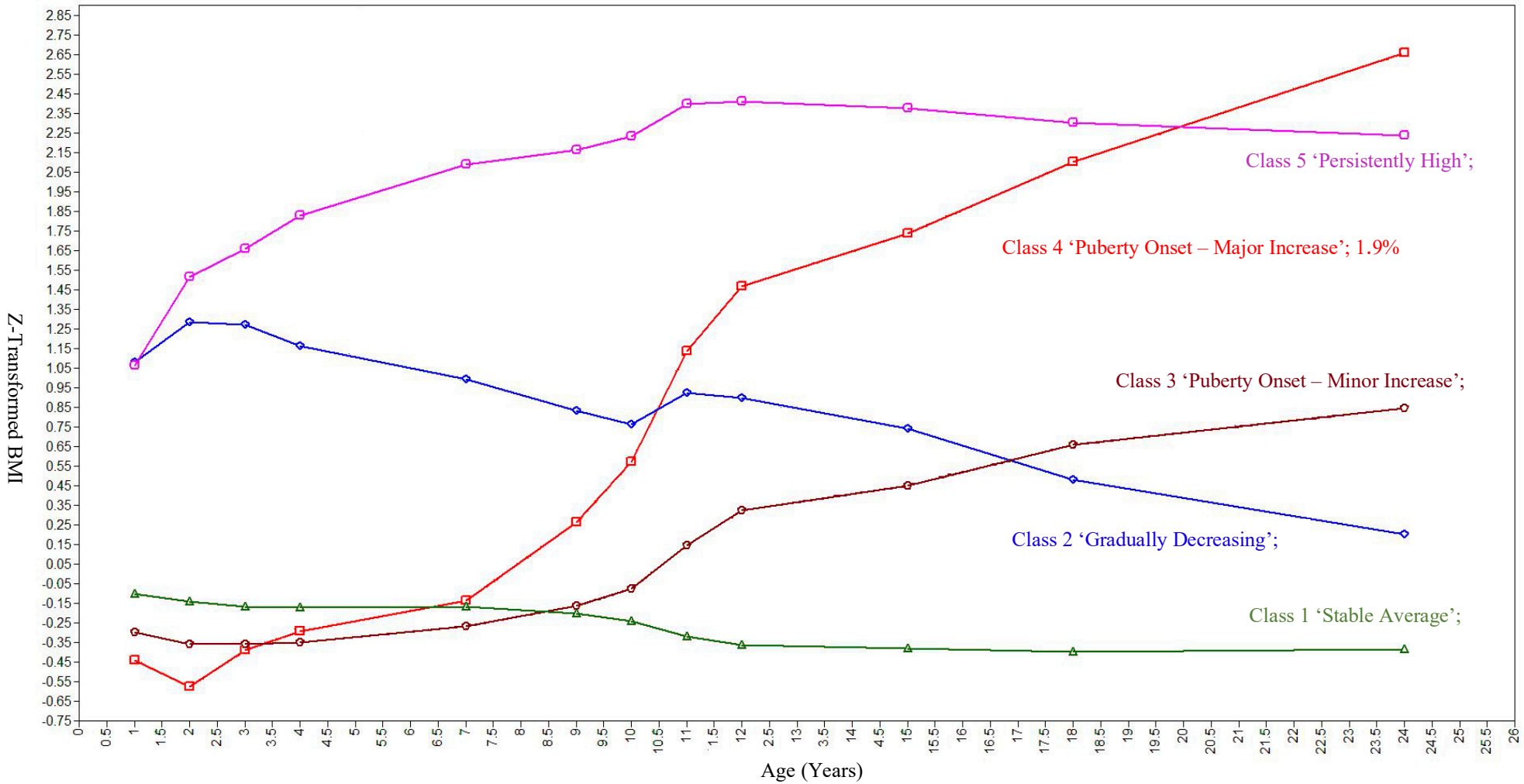
BIC = Bayesian Information Criterion; VLMR-LRT = Vuong-Lo-Mendell-Rubin Likelihood Ratio Test; BLRT = Parametric Bootstrap Likelihood Ratio Test; **Contained one trajectory with <1% of sample;* †Selected for further analysis

Figure 3: Fasting Insulin (Ages 9-24 years) and Body Mass Index (Ages 1-24 years) Trajectories

A: Fasting Insulin



B: BMI



Trajectories were delineated using growth mixture modelling at four time points for Fasting Insulin, and twelve time-points for body mass index. Nodes in the graph represent mean z-scores for fasting insulin or BMI at each time-point for each developmental trajectory

2.4.3 Correlation between Fasting Insulin and BMI Trajectories

The 'stable average' fasting insulin and BMI trajectories were weakly but statistically significantly correlated (r_{φ}=0.233, p<0.001), as were the 'persistently high' trajectories (r_{φ}=0.092, p<0.001).

2.4.4 Associations of Sociodemographic, Lifestyle and Clinical Factors with Trajectories

2.4.4.1 Fasting Insulin

Both deviating fasting insulin trajectories were associated with lower social class, family history of cardiometabolic disease, lower physical activity and smoking in adolescence/early adulthood. Lower birthweight and more perinatal stressful life events were associated with the 'persistently high' trajectory (Table 5), which also had mean fasting insulin, HDL, triglycerides, and CRP levels outside of UK reference ranges at age 24 years (Table 6). Deviating fasting insulin trajectories were associated with metabolic syndrome at age 24 years (adjusted OR for the 'persistently high' trajectory=9.21; 95% C.I., 3.77-20.15) (see Appendix B Table 1).

Table 5: Odds Ratios for Multinomial Logistic Regression Analyses Examining Predictors of Membership of Fasting Insulin Developmental Trajectories

Variable	**Odds Ratio (95% CI)**		
	Class 1[a] 'Stable Average'	**Class 2 'Minor Increase'**	**Class 3 'Persistently High'**
Female Sex	1.00	1.37 (1.10-2.04)	1.10 (0.89-1.23)
Non-White British Ethnicity	1.00	1.22 (0.89-1.62)	1.21 (0.91-1.73)
Lower Social Class	1.00	1.05 (1.00-1.09)	1.89 (1.35-2.50)
FHx Cardiometabolic Disorders	1.00	1.10 (0.92-1.41)	1.66 (1.14-1.69)
Gestational Age	1.00	1.10 (0.95-1.31)	1.21 (0.90-1.44)
Birthweight	1.00	0.89 (0.60-1.10)	0.76 (0.44-0.92)
Stressful Life Events (Top tertile)	1.00	1.21 (0.55-4.32)	2.06 (1.43-4.31)
Low Exercise (age 15)	1.00	1.13 (1.06-1.31)	1.16 (1.02-1.41)
Smoking (age 15)	1.00	1.45 (1.03-1.76)	1.10 (0.86-1.55)
Low Exercise (age 18)	1.00	1.45 (1.14-1.89)	1.54 (1.06-2.22)
Smoking (age 18)	1.00	1.39 (1.07-1.43)	1.40 (1.10-1.78)

[a]Reference group

Table 6: Anthropometric and Biochemical Measures for Different Fasting Insulin Trajectories

Measure, Mean (SD)	Trajectory		
	Class 1 (Stable Average)	Class 2 (Minor Increase)	Class 3 (Persistently High)
Fasting Insulin (μIU/mL)	6.93 (2.70)	8.57 (1.21)*	13.65 (4.32)*
Body Mass Index (kg/m^2)	22.12 (3.76)	26.18 (4.23)*	24.76 (7.74)
Waist Circumference, Males (cm)	84.58 (8.79)	99.56 (15.34)	94.45 (16.72)
Waist Circumference, Females (cm)	75.62 (9.44)	91.41 (14.14)*	89.20 (18.14)*
Fasting Plasma Glucose (mmol/L)	5.24 (0.67)	5.49 (0.65)	5.78 (0.78)
HDL Cholesterol (mmol/L)	1.60 (0.41)	1.32 (0.38)*	1.31 (0.46)*
Triglycerides (mmol/L)	0.89 (0.38)	1.31 (0.88)	1.75 (1.01)*
LDL Cholesterol (mmol/mL)	2.39 (0.73)	2.71 (0.80)	2.73 (0.89)
C-Reactive Protein (mg/L)	1.85 (3.98)	2.19 (3.24)	3.40 (4.21)*

*indicates outside of U.K. reference range: Body Mass Index=18.5-24.9kg/m^2 ; Fasting Insulin=3-8μIU/mL; Waist Circumference (males)=<102cm; Waist Circumference (females)=<88cm; Fasting Plasma Glucose=<7mmol/L; HDL=>1.35mmol/L; Triglycerides=<1.70mmol/L; LDL=<3.36mmol/L; CRP<3mg/L.

2.4.4.2 BMI

Deviating BMI trajectories were associated with lower social class, family history of cardiometabolic disease, more perinatal stressful life-events, lower physical activity and smoking in adolescence/early adulthood. Higher birthweight was associated with the 'gradually decreasing' and 'persistently high' trajectories, whereas lower birthweight was weakly associated with both 'puberty-onset' increase trajectories (Table 7). Deviating BMI trajectories were associated with mean values of fasting insulin, waist circumference, HDL, and CRP outside of UK reference ranges at age 24 years (Table 8). All deviating BMI trajectories were associated with metabolic syndrome at age 24 years (adjusted OR for the 'persistently high' trajectory=10.62; 95% C.I., 5.89-19.13) (see Appendix B Table 1).

Table 7: Odds Ratios for Multinomial Regression Analyses Examining Predictors of Membership of BMI Developmental Trajectories

Variable	Odds Ratio (95% CI)				
	Class 1[a] 'Stable Average'	Class 2 'Gradually Decreasing'	Class 3 'Puberty Onset – Minor Increase'	Class 4 'Puberty Onset – Major Increase'	Class 5 'Persistently High'
Female Sex	1.00	1.10 (0.90-1.21)	1.35 (1.12-1.56)	1.10 (0.91-1.26)	0.89 (0.70-1.13)
Non-White British Ethnicity	1.00	1.76 (1.16-2.65)	1.09 (0.61-1.93)	1.12 (0.35-3.56)	0.62 (0.20-1.96)
Lower Social Class	1.00	1.08 (0.99-1.17)	1.11 (1.01-1.21)	1.13 (1.05-1.22)	1.26 (1.11-1.44)
FHx Cardiometabolic Disorders	1.00	1.19 (0.92-1.55)	1.48 (1.20-1.84)	2.43 (1.35-4.37)	2.69 (1.82-3.98)
Gestational Age	1.00	0.98 (0.92-1.05)	1.00 (0.94-1.07)	1.12 (0.43-2.95)	1.12 (0.87-1.32)
Birthweight	1.00	1.30 (1.18-1.43)	0.99 (0.93-1.07)	0.90 (0.83-1.15)	1.44 (1.25-1.65)
Stressful Life Events	1.00	0.84 (0.68-1.04)	1.11 (0.96-1.30)	1.44 (1.01-2.07)	1.89 (1.44-3.21)
Low Exercise (age 15)	1.00	1.06 (0.84-1.32	1.38 (1.13-1.69)	1.90 (1.08-3.35)	1.36 (0.87-2.12)
Smoking (age 15)	1.00	1.22 (0.75-2.03)	1.62 (1.17-2.25)	1.14 (0.57-3.67)	1.20 (0.72-2.01)
Low Exercise (age 18)	1.00	0.78 (0.56-0.95)	1.31 (1.04-1.65)	1.50 (1.01-2.90)	0.94 (0.65-1.36)
Smoking (age 18)	1.00	1.15 (0.71-1.86)	1.63 (1.12-2.38)	2.37 (0.99-5.72)	1.44 (0.73-2.84)

[a]reference group

Table 8: Anthropometric and Biochemical Characteristics of Participants included in BMI Trajectories

Measure, Mean (SD)	Trajectory				
	Class 1 (Stable Average)	Class 2 (Gradually Decreasing)	Class 3 (Puberty Onset – Minor Increase)	Class 4 (Puberty Onset – Major Increase)	Class 5 (Persistently High)
Body Mass Index	23.60 (3.46)	25.32 (3.85)*	27.25 (4.47)*	33.67 (8.68)*	31.55 (5.66)*
Fasting Insulin (μIU/mL)	6.42 (2.12)	6.45 (3.39)	7.32 (4.44)	8.44 (5.43)*	8.21 (3.19)*
Waist Circumference Males (cm)	83.50 (8.45)	86.40 (10.36)	100.67 (11.60)	121.46 (6.70)*	111.77 (12.29)*
Waist Circumference Females (cm)	75.62 (9.12)	78.80 (10.15)	87.12 (11.71)	99.76 (19.17)*	94.97 (14.26)*
Fasting Plasma Glucose (mmol/L)	5.28 (0.70)	5.24 (0.49)	5.44 (0.59)	5.36 (0.51)	5.49 (0.97)
HDL Cholesterol, (mmol/L)	1.57 (0.42)	1.54 (0.42)	1.45 (0.32)*	1.32 (0.21)*	1.35 (0.46)*
Triglycerides (mmol/L)	0.94 (0.50)	0.93 (0.47)	1.34 (0.82)	1.44 (0.61)	1.29 (0.77)
LDL Cholesterol (mmol/mL)	2.41 (0.75)	2.37 (0.73)	2.48 (0.80)	2.77 (0.62)	2.79 (0.87)
C-Reactive Protein (mg/L)	2.08 (6.93)	2.11 (3.99)	3.01 (4.49)*	4.76 (3.76)*	4.03 (4.20)*

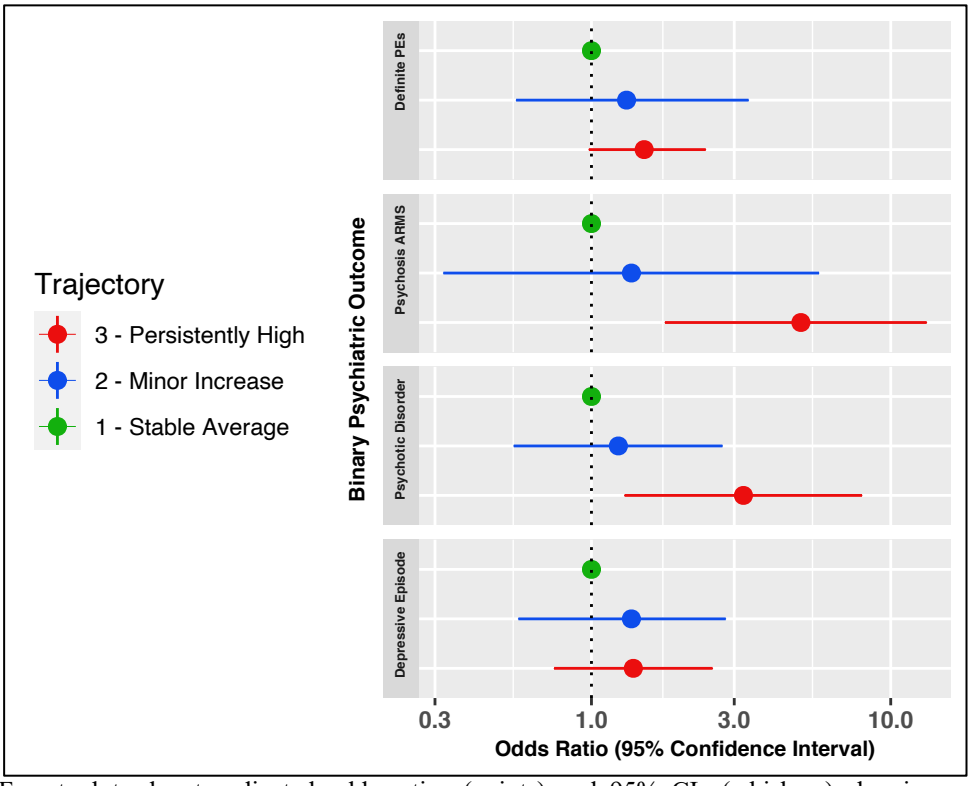
*indicates outside of U.K. reference range: Body Mass Index=18.5-24.9kg/m^2 ; Fasting Insulin=3-8μIU/mL; Waist Circumference (males)=<102cm; Waist Circumference (females)=<88cm; Fasting Plasma Glucose=<7mmol/L; HDL=>1.35mmol/L; Triglycerides=<1.70mmol/L; LDL=<3.36mmol/L; CRP<3mg/L;

2.4.5 Associations of Cardiometabolic Trajectories with Psychiatric Outcomes at Age 24

2.4.5.1 Fasting Insulin

The 'persistently high' fasting insulin trajectory was associated with psychosis ARMS (adjusted OR=5.01; 95% C.I., 1.76-13.19), psychotic disorder (adjusted OR=3.22; 95% C.I., 1.29-8.02), and weakly associated with negative symptoms (adjusted β=0.07; 95% C.I., 0.01-0.13) at age 24 years. Fasting insulin trajectories were not associated with depression (Figure 4; Table 9 & Table 10).

Figure 4: Associations of Fasting Insulin Trajectories with Binary Psychiatric Outcomes at Age 24 Years



Forest plots denote adjusted odds ratios (points) and 95% CIs (whiskers) showing associations of fasting insulin trajectories with risk of binary psychosis and depression outcomes at age 24, after adjusting for sex, ethnicity, social class, childhood emotional and behavioural problems, cumulative smoking, physical activity, alcohol and substance use, sleep problems and calorie intake.

Table 9: Odds Ratios (95% CIs) for Associations of Fasting Insulin Trajectories with Binary Psychiatric Outcomes at Age 24 Years

Trajectory/Outcome	Sample	Odds Ratio (95% C.I.)		*P*-value[a]
		Unadjusted	Adjusted for sex, ethnicity, social class, SDQ (7y), cumulative smoking, physical activity, alcohol and substance use, sleep problems and calorie intake	
Definite PEs at Age 24				
Class 1 – 'Stable Average'	4,939	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Minor Increase'	693	1.48 (0.98-2.24)	1.31 (0.56-3.35)	>0.999
Class 3 – 'Persistently High'	158	1.88 (1.05-3.60)	1.50 (0.98-2.41)	0.329
Psychosis 'At Risk Mental State' at Age 24				
Class 1 – 'Stable Average'	4,939	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Minor Increase'	693	1.59 (0.20-8.02)	1.36 (0.32-5.76)	>0.999
Class 3 – 'Persistently High'	158	6.33 (1.97-20.30)	5.01 (1.76-13.19)	0.006
Psychotic Disorder at Age 24				
Class 1 – 'Stable Average'	4,939	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Minor Increase'	693	1.85 (0.70-4.88)	1.23 (0.55-2.74)	>0.999
Class 3 – 'Persistently High'	158	4.74 (1.67-13.42)	3.22 (1.29-8.02)	0.048
Depressive Episode at Age 24				
Class 1 – 'Stable Average'	4,939	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Minor Increase'	693	1.26 (0.73-2.67)	1.36 (0.57-2.81)	0.883
Class 3 – 'Persistently High'	158	1.31 (0.81-4.32)	1.38 (0.75-2.54)	0.686

[a]*p*-values adjusted for multiple testing using Holm-Bonferroni method

Table 10: Beta Coefficients (95% CIs) for Associations of Fasting Insulin Trajectories with Continuous Psychiatric Outcomes At Age 24 Years

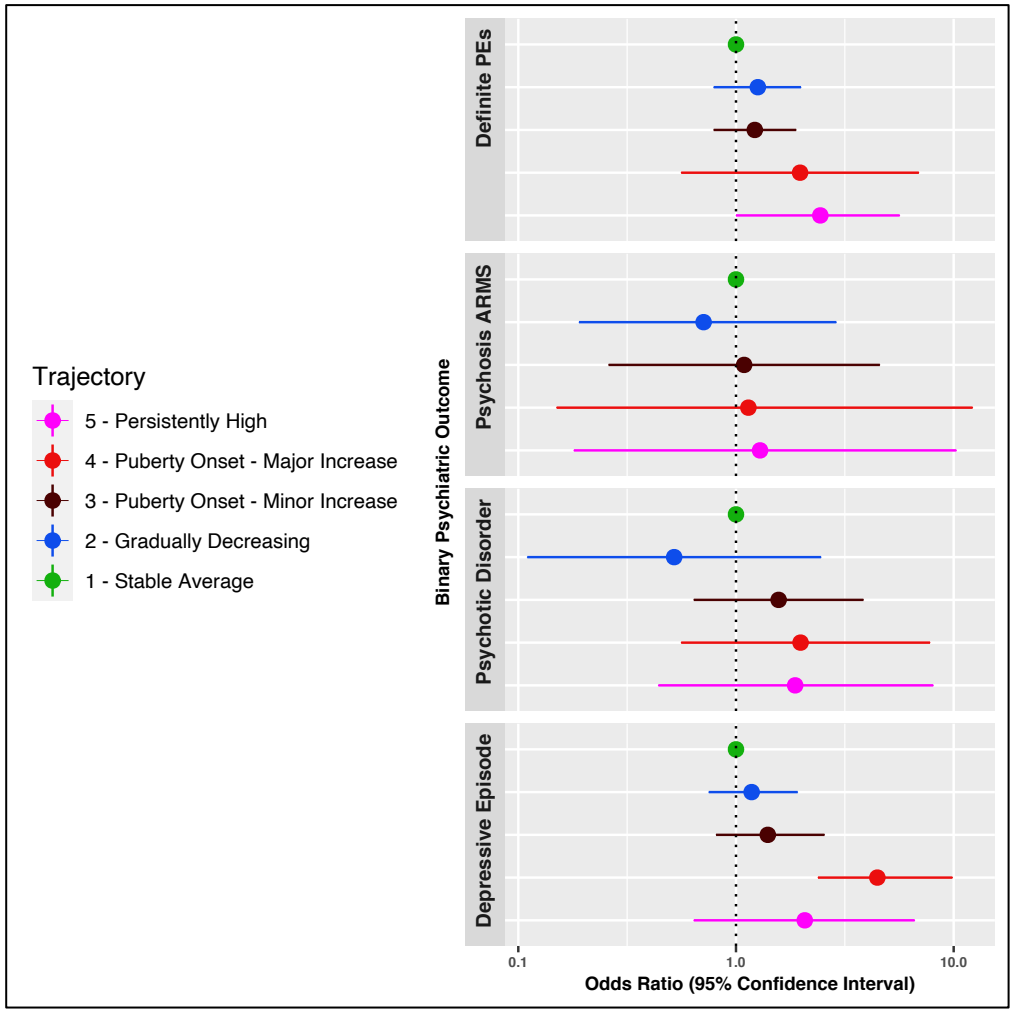
Trajectory	Sample	Beta Coefficient (95% C.I.)		*p*-value[a]
		Unadjusted	Adjusted for sex, ethnicity, social class, SDQ (7y), cumulative smoking, physical activity, alcohol and substance use, sleep, calorie intake, negative/depressive symptoms	
Depressive Symptom Score at Age 24				
Class 1 – 'Stable Average'	4,939	0.00 [reference]	0.00 [reference]	-
Class 2 – 'Minor Increase'	693	0.03 (-0.02, 0.08)	0.02 (-0.04. 0.08)	>0.999
Class 3 – 'Persistently High'	158	0.08 (0.04, 0.13)	0.05 (-0.03,0.13)	0.669
Negative Psychotic Symptom Score at Age 24				
Class 1 – 'Stable Average'	4,939	0.00 [reference]	0.00 [reference]	-
Class 2 – 'Minor Increase'	693	0.08 (-0.01,0.16)	0.05 (0.01,0.09)	0.192
Class 3 – 'Persistently High'	158	0.18 (0.10,0.26)	0.07 (0.01, 0.13)	0.049

[a]*p*-values adjusted for multiple testing using Holm-Bonferroni method

2.4.5.2 BMI

The 'puberty-onset major increase' trajectory was associated with higher risk of depressive episode (adjusted OR=4.46; 95% C.I., 2.38-9.87) and depressive symptoms (adjusted β=0.08; 95% C.I., 0.03-0.14) at age 24 years. The 'puberty-onset minor increase' trajectory was weakly associated with depressive symptoms at 24y (adjusted β=0.06; 95% C.I., 0.01-0.11). BMI trajectories were not associated with psychosis outcomes (Figure 5; Table 11 & Table 12).

Figure 5: Associations of BMI Trajectories with Binary Psychiatric Outcomes At Age 24 Years



Forest plots denote adjusted odds ratios (points) and 95% CIs (whiskers) showing associations of BMI trajectories with risk of binary psychosis and depression outcomes at age 24, after adjusting for sex, ethnicity, social class, childhood emotional and behavioural problems, cumulative smoking, physical activity, alcohol and substance use, sleep problems and calorie intake.

Table 11: Odds Ratios (95% CIs) for Associations of BMI Trajectories with Binary Psychiatric Outcomes at Age 24 Years

Trajectory/ Outcome	Sample	Odds Ratio (95% C.I.)		*P*-value[a]
		Unadjusted	Adjusted for sex, ethnicity, social class, SDQ (7y), cumulative smoking, physical activity, alcohol and substance use, sleep problems and calorie intake	
Definite PEs at Age 24				
Class 1 – 'Stable Average'	8,383	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Gradually Decreasing'	949	1.43 (0.82-1.96)	1.26 (0.79-1.99)	>0.999
Class 3 – 'Puberty Onset – Minor Increase'	668	1.66 (0.87-2.55)	1.22 (0.79-1.89)	>0.999
Class 4 – 'Puberty Onset – Major Increase'	174	3.56 (0.87-11.54)	1.97 (0.56-6.92)	>0.999
Class 5 – 'Persistently High	289	3.21 (1.01-9.11)	2.44 (1.00-5.65)	0.367
Psychosis 'At Risk Mental State'				
Class 1 – 'Stable Average'	8,383	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Gradually Decreasing'	949	0.49 (0.10-3.21)	0.71 (0.19-2.89)	>0.999
Class 3 – 'Puberty Onset – Minor Increase'	668	1.12 (0.23-5.43)	1.09 (0.26-4.58)	>0.999
Class 4 – 'Puberty Onset – Major Increase'	174	1.32 (0.10-13.11)	1.14 (0.15-12.22)	>0.999
Class 5 – 'Persistently High	289	1.55 (0.44-3.21)	1.29 (0.18-10.29)	>0.999
Psychotic Disorder at Age 24				
Class 1 – 'Stable Average'	8,383	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Gradually Decreasing'	949	0.44 (0.21-2.03)	0.52 (0.11-2.46)	>0.999
Class 3 – 'Puberty Onset – Minor Increase'	668	1.97 (0.60-3.46)	1.57 (0.64-3.85)	>0.999
Class 4 – 'Puberty Onset – Major Increase'	174	2.14 (0.65-6.21)	1.98 (0.56-7.79)	>0.999
Class 5 – 'Persistently High	289	3.11 (0.53-13.22)	1.87 (0.44-8.06)	>0.999
Depressive Episode at Age 24				
Class 1 – 'Stable Average'	8,383	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Gradually Decreasing'	949	1.33 (0.77-1.88)	1.18 (0.75-1.92)	>0.999
Class 3 – 'Puberty Onset – Minor Increase'	668	1.69 (0.90-3.21)	1.40 (0.81-2.55)	>0.999
Class 4 – 'Puberty Onset – Major Increase'	174	8.91 (4.21-17.12)	4.46 (2.38-9.87)	0.006
Class 5 – 'Persistently High	289	3.01 (0.91-7.59)	2.07 (0.64-6.62)	>0.999

[a] *p*-values adjusted for multiple testing using Holm-Bonferroni method

Table 12: Beta Coefficients (95% CIs) for Associations of Body Mass Index Trajectories with Continuous Psychiatric Outcomes at Age 24 Years

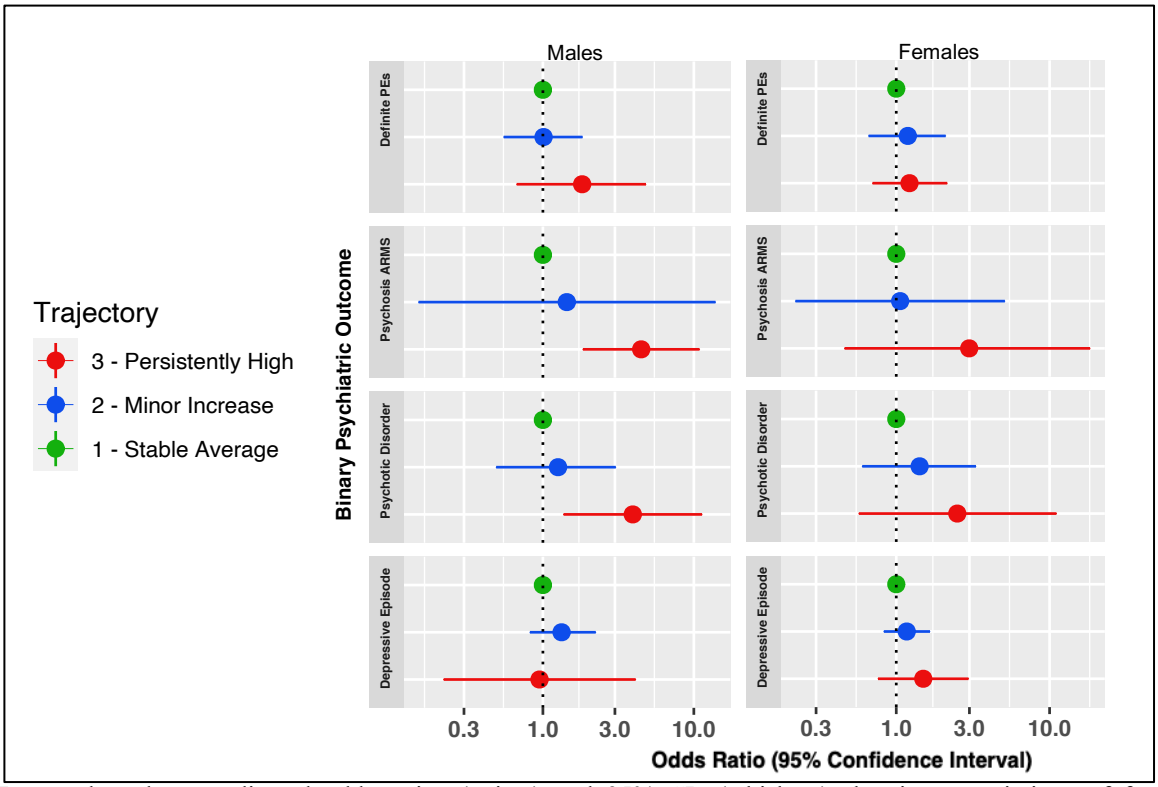
Trajectory	Sample	Beta Coefficient (95% C.I.)		*p*-value[a]
		Unadjusted	Adjusted for sex, ethnicity, social class, SDQ (7y), cumulative smoking, physical activity, alcohol and substance use, sleep, calorie intake negative/depressive symptoms	
Depressive Symptom Score at Age 24				
Class 1 – 'Stable Average'	8,383	0.00 [reference]	0.00 [reference]	-
Class 2 – 'Gradually Decreasing'	949	0.02 (-0.06, 0.10)	0.01 (-0.05, 0.08)	>0.999
Class 3 – 'Puberty Onset – Minor Increase'	668	0.14 (0.08, 0.20)	0.06 (0.01, 0.11)	0.124
Class 4 – 'Puberty Onset – Major Increase'	174	0.20 (0.10, 0.31)	0.08 (0.03, 0.14)	0.033
Class 5 – 'Persistently High	289	0.10 (-0.09, 0.21)	0.02 (-0.08, 0.13)	>0.999
Negative Psychotic Symptom Score at Age 24				
Class 1 – 'Stable Average'	8,383	0.00 [reference]	0.00 [reference]	-
Class 2 – 'Gradually Decreasing'	949	0.07 (-0.03, 0.16)	0.04 (-0.05, 0.13)	>0.999
Class 3 – 'Puberty Onset – Minor Increase'	668	0.11 (0.05, 0.17)	0.03 (-0.05, 0.11)	0.796
Class 4 – 'Puberty Onset – Major Increase'	174	0.18 (0.11, 0.24)	0.06 (-0.03, 0.16)	0.514
Class 5 – 'Persistently High	289	0.13 (0.02, 0.24)	0.09 (-0.04, 0.23)	>0.999

[a]*p*-values adjusted for multiple testing using Holm-Bonferroni method

2.4.6 Sex Stratified Associations of Fasting Insulin and BMI Trajectories with Risks for Psychosis and Depression

For fasting insulin, the pattern of association with risks for psychiatric outcomes in sex stratified analysis was similar to the primary analysis (Figure 6; Table 13; Appendix B Table 2). For BMI, point estimates for depression for both 'puberty-onset' increase trajectories were larger in females. There was no significant association of BMI trajectories with psychosis outcomes (Figure 7; Table 14; Appendix B Table 3).

Figure 6: Sex-Stratified Associations of Fasting Insulin Trajectories with Binary Psychiatric Outcomes at Age 24 Years



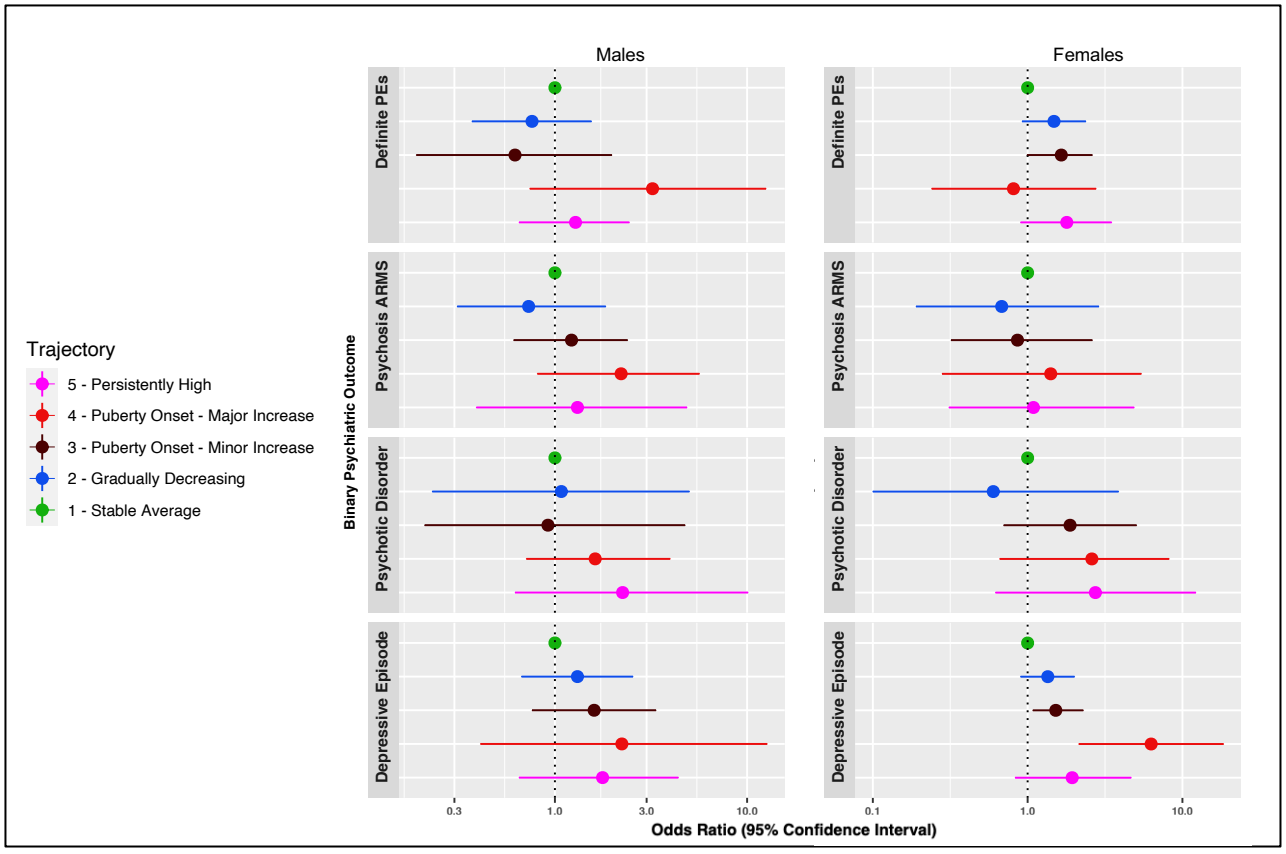
Forest plots denote adjusted odds ratios (points) and 95% CIs (whiskers) showing associations of fasting insulin trajectories with risk of binary psychosis and depression outcomes at age 24 years in males and females separately, after adjusting for sex, ethnicity, social class, childhood emotional and behavioural problems, cumulative smoking, physical activity, alcohol and substance use, sleep problems and calorie intake.

Table 13: Odds Ratios for Sex-Stratified Associations of Fasting Insulin Trajectories with Binary Psychiatric Outcomes at Age 24 Years

Trajectory	Sample	Odds Ratio (95% C.I.)		*p*-value[a]
		Unadjusted	Adjusted for sex, ethnicity, social class, SDQ (7y), cumulative smoking, physical activity, alcohol and substance use, sleep, and calorie intake	
Definite PEs at Age 24 (Males)				
Class 1 – 'Stable Average'	2,319	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Minor Increase'	278	1.51 (0.91-2.54)	1.01 (0.55-1.83)	>0.999
Class 3 – 'Persistently High'	66	1.91 (1.02-5.03)	1.82 (0.67-4.82)	0.472
Definite PEs at Age 24 (Females)				
Class 1 – 'Stable Average'	2,620	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Minor Increase'	415	1.32 (1.11-1.89)	1.19 (0.66-2.10)	>0.999
Class 3 – 'Persistently High'	92	1.65 (1.12-2.01)	1.22 (0.70-2.15)	>0.999
Psychosis At Risk Mental State at Age 24 (Males)				
Class 1 – 'Stable Average'	2,319	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Minor Increase'	278	1.65 (0.42-5.30)	1.44 (0.15-13.92)	>0.999
Class 3 – 'Persistently High'	66	8.32 (3.13-16.49)	4.48 (1.84-10.91)	0.006
Psychosis At Risk Mental State at Age 24 (Females)				
Class 1 – 'Stable Average'	2,620	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Minor Increase'	415	1.39 (0.29-5.57)	1.06 (0.22-5.11)	>0.999
Class 3 – 'Persistently High'	92	4.84 (0.47-31.18)	2.99 (0.46-18.37)	0.842
Psychotic Disorder at Age 24 (Males)				
Class 1 – 'Stable Average'	2,319	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Minor Increase'	278	1.55 (0.61-4.31)	1.26 (0.49-3.04)	>0.999
Class 3 – 'Persistently High'	66	5.79 (1.24-27.09)	3.94 (1.37-11.34)	0.046
Psychotic Disorder at Age 24 (Females)				
Class 1 – 'Stable Average'	2,620	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Minor Increase'	415	1.45 (0.63-3.35)	1.42 (0.60-3.31)	>0.999
Class 3 – 'Persistently High'	92	3.29 (0.53-9.86)	2.50 (0.57-11.09)	>0.999
Depressive Episode at Age 24 (Males)				
Class 1 – 'Stable Average'	2,319	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Minor Increase'	278	1.80 (1.04-3.11)	1.33 (0.82-2.24)	>0.999
Class 3 – 'Persistently High'	66	0.97 (0.23-4.13)	0.95 (0.22-4.12)	>0.999
Depressive Episode at Age 24 (Females)				
Class 1 – 'Stable Average'	2,620	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Minor Increase'	415	1.23 (0.88-1.73)	1.17 (0.83-1.66)	>0.999
Class 3 – 'Persistently High'	92	1.61 (0.82-3.14)	1.50 (0.76-2.96)	>0.999

[a]*p*-values adjusted for multiple testing using Holm-Bonferroni method

Figure 7: Sex-Stratified Associations of BMI Trajectories with Binary Psychiatric Outcomes at Age 24 Years



Forest plots denote adjusted odds ratios (points) and 95% CIs (whiskers) showing associations of BMI trajectories with risk of binary psychosis and depression outcomes at age 24 years in males and females separately, after adjusting for sex, ethnicity, social class, childhood emotional and behavioural problems, cumulative smoking, physical activity, alcohol and substance use, sleep problems and calorie intake.

Table 14: Odds Ratios for Sex-Stratified Associations of Body Mass Index Trajectories with Binary Psychiatric Outcomes at Age 24 Years

Trajectory	Sample	Odds Ratio (95% C.I.)		*p*-value[a]
		Unadjusted	Adjusted for sex, ethnicity, social class, SDQ (7y), cumulative smoking, physical activity, alcohol and substance use, sleep, and calorie intake	
Definite PEs at Age 24 (Males)				
Class 1 – 'Stable Average'	4,164	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Gradually Decreasing'	443	1.22 (0.46-1.87)	0.76 (0.37-1.55)	>0.999
Class 3 – 'Puberty Onset – Minor Increase'	311	1.22 (0.63-2.36)	0.62 (0.19-1.98)	>0.999
Class 4 – 'Puberty Onset – Major Increase'	105	5.87 (0.53-9.21)	3.22 (0.74-12.55)	>0.999
Class 5 – 'Persistently High	107	1.47 (0.43-4.98)	1.28 (0.65-2.44)	>0.999
Definite PEs at Age 24 (Females)				
Class 1 – 'Stable Average'	4,219	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Gradually Decreasing'	506	1.14 (0.74-1.75)	1.48 (0.92-2.38)	0.501
Class 3 – 'Puberty Onset – Minor Increase'	357	1.90 (1.11-3.26)	1.65 (0.99-2.62)	0.328
Class 4 – 'Puberty Onset – Major Increase'	184	1.54 (0.65-3.66)	0.81 (0.24-2.77)	1.000
Class 5 – 'Persistently High	67	2.32 (0.88-6.13)	1.79 (0.90-3.49)	0.182
Psychosis At Risk Mental State at Age 24 (Males)				
Class 1 – 'Stable Average'	4,164	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Gradually Decreasing'	443	0.60 (0.44-2.12)	0.73 (0.31-1.84)	>0.999
Class 3 – 'Puberty Onset – Minor Increase'	311	1.51 (0.55-4.64)	1.22 (0.61-2.39)	>0.999
Class 4 – 'Puberty Onset – Major Increase'	105	3.13 (1.01-5.12)	2.21 (0.81-5.65)	>0.999
Class 5 – 'Persistently High	107	1.69 (0.60-2.01)	1.31 (0.39-4.87)	>0.999
Psychosis At Risk Mental State at Age 24 (Females)				
Class 1 – 'Stable Average'	4,219	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Gradually Decreasing'	506	0.61 (0.14-2.14)	0.68 (0.19-2.89)	>0.999
Class 3 – 'Puberty Onset – Minor Increase'	357	0.76 (0.76-3.21)	0.86 (0.32-2.62)	>0.999
Class 4 – 'Puberty Onset – Major Increase'	184	1.81 (0.25-6.43)	1.41 (0.28-5.43)	>0.999
Class 5 – 'Persistently High	67	1.21 (0.77-3.21)	1.09 (0.31-4.88)	>0.999
Psychotic Disorder at Age 24 (Male)				
Class 1 – 'Stable Average'	4,164	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Gradually Decreasing'	443	0.76 (0.54-2.01)	1.08 (0.23-5.01)	>0.999
Class 3 – 'Puberty Onset – Minor Increase'	311	1.02 (0.65-1.43)	0.92 (0.21-4.76)	>0.999
Class 4 – 'Puberty Onset – Major Increase'	105	2.12 (0.91-4.12)	1.62 (0.71-3.98)	>0.999
Class 5 – 'Persistently High	107	3.52 (0.44-15.09)	2.25 (0.62-10.12)	>0.999
Psychotic Disorder at Age 24 (Female)				
Class 1 – 'Stable Average'	4,219	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Gradually Decreasing'	506	0.40 (0.09-1.21)	0.60 (0.10-3.87)	>0.999
Class 3 – 'Puberty Onset – Minor Increase'	357	3.16 (1.29-5.12)	1.88 (0.70-5.06)	>0.999
Class 4 – 'Puberty Onset – Major Increase'	184	1.31 (0.65-3.21)	2.60 (0.66-8.21)	>0.999
Class 5 – 'Persistently High	67	1.21 (0.40-6.21)	2.74 (0.62-12.22)	>0.999
Depressive Episode at Age 24 (Males)				
Class 1 – 'Stable Average'	4,164	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Gradually Decreasing'	443	1.31 (0.71-2.44)	1.31 (0.67-2.55)	>0.999
Class 3 – 'Puberty Onset – Minor Increase'	311	1.62 (0.83-3.17)	1.60 (0.76-3.36)	>0.999
Class 4 – 'Puberty Onset – Major Increase'	105	3.21 (0.67-8.21)	2.23 (0.41-12.72)	>0.999
Class 5 – 'Persistently High	107	1.31 (0.30-5.67)	1.77 (0.65-4.39)	>0.999

Depressive Episode at Age 24 (Females)				
Class 1 – 'Stable Average'	4,219	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Gradually Decreasing'	506	1.20 (0.82-1.85)	1.35 (0.90-2.01)	>0.999
Class 3 – 'Puberty Onset – Minor Increase'	357	1.91 (1.15-2.91)	1.52 (1.08-2.29)	0.047
Class 4 – 'Puberty Onset – Major Increase'	184	5.21 (2.09-8.21)	6.28 (2.14-18.44)	0.006
Class 5 – 'Persistently High	67	1.73 (0.86-3.51)	1.94 (0.83-4.67)	>0.999

[a]*p*-values adjusted for multiple testing using Holm-Bonferroni method

2.5 Discussion

I delineated fasting insulin and BMI trajectories from early life using prospective repeated measurements in a large population-representative birth cohort and report distinct associations with adult schizophrenia-spectrum and depression outcomes. After adjusting for several relevant confounders, I found that persistently high fasting insulin from mid-childhood was associated with increased risk of schizophrenia-spectrum outcomes at 24 years, while BMI increases around the age of puberty onset were associated with depression at 24 years. Associations of fasting insulin and BMI trajectories with cardiometabolic risk factors such as social class, ethnicity, smoking and physical activity alongside adult metabolic syndrome provides face validity to the identified trajectories. Although the last data point for fasting insulin/BMI overlapped with outcome assessment, trajectories were differentiated by mid-childhood, suggesting a temporal relationship between exposure and outcome. Evidence for the association of 'puberty-onset' BMI increase and adult depression remained after adjusting for childhood emotional and behavioural problems, suggesting that reverse causality may not fully explain this finding. Whilst the same adjustment may be less capable of ruling out reverse causality for associations involving schizophrenia-spectrum outcomes; it is improbable that many participants had experienced psychosis before age 9 years since the prevalence of pre-pubertal psychosis is rare (McClellan and Werry, 1997). Therefore, reverse causality for associations with schizophrenia-spectrum outcomes is rare.

I found consistent evidence for an association between fasting insulin trajectories and schizophrenia-spectrum outcomes. Effect sizes were largest in the 'persistently high' trajectory, consistent with a dose-response relationship, and point estimates were larger in more clinically relevant outcomes. The findings complement meta-analyses reporting altered glucose-insulin homeostasis in FEP (Perry et al., 2016, Pillinger et al., 2017a). Moreover, I show that disruptions to glucose-insulin homeostasis detectable at FEP may begin much earlier in life. The point estimates partly attenuated after adjustment for confounders, so malleable lifestyle factors such as smoking, physical activity and diet must remain crucial targets for reducing the risk of incident cardiometabolic disorders in young people with psychosis. I also found that participants classified into the 'persistently high' fasting insulin trajectory, who had the highest risk of schizophrenia-spectrum outcomes, had mean BMI and FPG values within the reference range at age 24 years. Therefore, these individuals may be 'hiding in plain sight' in EIS since commonly measured physical indices may not identify them. Consequently, careful assessment and clinical considerations are needed to minimise the risk of cardiometabolic disorders in these individuals.

The findings suggest that altered glucose-insulin homeostasis could be a shared mechanism for schizophrenia and T2D, which could be genetic and/or environmental in origin. People with comorbid schizophrenia and T2D have a higher genetic predisposition to both disorders than controls (Hackinger et al., 2018), and genetic predisposition for schizophrenia is associated with insulin resistance in schizophrenia patients (Tomasik et al., 2019). Additionally, I found that the highest risks for schizophrenia-spectrum outcomes were found in the fasting insulin trajectory associated with lower birth weight and more perinatal stressful life events. I found similar patterns of association in BMI trajectories which were associated with depression. These findings are consistent with the developmental programming hypothesis (Barker et al., 1993), positing that disruption to early-life development can have wide-ranging and far-reaching impacts on adult health.

The findings regarding BMI trajectories with depression at age 24 years are in line with meta-analyses (Gariepy et al., 2010, Luppino et al., 2010) suggesting an association between BMI and risk of depression. Similar trajectories of BMI have previously been linked with adult T2D (Zhang et al., 2019), obesity (Rolland-Cachera and Peneau, 2013) and CVD (Barker et al., 2005). The character and composition of BMI trajectories I identified are consistent with previous studies, although the length of follow-up was longer than most previous studies (Mattsson et al., 2019).

The findings provide further insights into the link between BMI and depression (Luppino et al., 2010), showing that puberty-onset increases in BMI specifically are associated with risk of adult depression. This finding, together with the lack of evidence for an association between persistently high BMI and depression, indicates that BMI might be a risk indicator for depression rather than a risk factor. This is because individuals in the 'persistently-high' BMI trajectory would likely have been exposed to the "largest dose" of BMI. Therefore, if BMI were the risk factor, one would have expected the largest effect size for depression in that trajectory. Consequently, environmental and/or genetic factors influencing BMI during puberty are likely to be important risk factors for depression. For instance, social stressors such as bullying may predispose to altered eating behaviours and increased risk of depression in adolescents (Lee and Vaillancourt, 2018). Additionally, deviating childhood BMI trajectories have previously been associated with a greater risk of adolescent/adult eating disorders (Yilmaz et al., 2019), which are highly comorbid with depression (Welch et al., 2016).

Also, effects of the female sex hormone oestrogen may be relevant since the associations of puberty-onset BMI increases and depression appeared stronger in females. Changes in oestrogen levels are associated with depressive symptoms throughout the life course, including pregnancy (Schiller et al., 2015), menopause (Dalal and Agarwal, 2015) and puberty (Soares and Zitek, 2008). Oestrogen is associated with obesity (Li et al., 2017a) and may explain the genetic correlation of age at menarche with adult obesity (Bell et al., 2018) and depression (Lewis et al., 2018). Further research is needed

to identify factors influencing pubertal BMI increases, as they may represent critical preventative targets for depression.

I did not find consistent evidence for associations of fasting insulin trajectories with depression or of BMI trajectories with schizophrenia-spectrum outcomes. Previous research has reported mixed findings regarding the association between insulin resistance and depression in young adults (Timonen et al., 2006, Perry et al., 2020b). However, some estimates for the associations of BMI trajectories with schizophrenia-spectrum outcomes in the analyses had wide confidence intervals, possibly due to sample size. These particular findings require replication in larger samples of people with psychosis.

Strengths of the study include a longitudinal design with repeated measurements of fasting insulin and BMI between ages 1-24 years in a relatively large sample, enabling a detailed examination of dynamic cardiometabolic change from childhood to early adulthood. I included several relevant schizophrenia-spectrum and depression outcomes, which allowed me to examine for specificity and a biological gradient of evidence.

Limitations of the study include missing data. Whilst I used a robust method to handle missing data, FIML may be biased in instances where data are 'missing not at random' (Cham et al., 2017). However, the risk of bias in FIML is no greater than the bias associated with traditional complete-case methods (Little et al., 2014), and FIML permitted a larger sample size and therefore increased statistical power. Nevertheless, missing psychiatric outcome data may have affected the results. Furthermore, residual confounding could still be an issue. For example, I could not account for psychological stress since data on cortisol levels were available only at age 9 years, in a small sub-section of the cohort. In addition, the confidence intervals were relatively wide for the sex-stratified analysis, likely due to reduced statistical power. Therefore, replication of the work in larger samples is required. Finally, the ALSPAC dataset does not include an ICD diagnosis of schizophrenia as an outcome. However, the psychotic disorder outcome would likely meet the threshold for clinical intervention, and all the psychosis outcomes I included lie on the schizophrenia continuum.

In summary, I report that the cardiometabolic comorbidity of psychosis and depression may have distinct early-life origins. Disrupted glucose-insulin homeostasis from mid-childhood is associated with adult psychosis, and BMI increases starting around the time of puberty onset are associated with adult depression. Whilst residual confounding may still be an issue; the results suggest that these cardiometabolic markers could be among shared risk factors/indicators for adult cardiometabolic and psychiatric disorders and may represent novel targets for treatment/prevention of cardiometabolic disorders in people with psychosis and depression.

Section B: Summary of Main Findings and Conclusions

Three aspects of the findings from Section B enrich our understanding of the nature of association between cardiometabolic disorders and schizophrenia. First, the results suggest that disruptions to glucose-insulin homeostasis may predate the onset of psychosis. Reverse causality is unlikely to explain this finding since psychosis is rare in prepubertal children (McClellan and Werry, 1997). Therefore, these findings counter the traditionally attributed notion that the high prevalence of cardiometabolic disorders in schizophrenia can be explained fully by lifestyle and clinical factors such as diet, exercise, and antipsychotic medications (i.e., cardiometabolic disorders are a consequence of the illness).

Second, the associations persisted after adjusting for a detailed range of potential sociodemographic and lifestyle confounders, suggesting that disrupted glucose-insulin homeostasis could be a risk factor for psychosis. This argument is also strengthened given that the study was conducted in a relatively young sample who would have been less affected by chronic lifestyle factors than studies of older adults. In addition, given the relatively young age of the sample, it is unlikely that the results could be explained by antipsychotic medications, which are not recommended (National Institute for Health and Care Excellence., 2013) and rarely prescribed (Olfson, 2009) in childhood.

Third, while depression shows strong genetic (Anttila et al., 2018) and phenotypic (Buckley et al., 2009) overlap with schizophrenia, and has similar associations with cardiometabolic disorders (Firth et al., 2019), my results suggest the cardiometabolic origins of the two psychiatric disorders are distinct. The results indicate that primary glucose-insulin homeostasis may be specific to psychotic disorders such as schizophrenia, further strengthening the idea that disruption to glucose-insulin homeostasis may be a risk factor for psychosis.

SECTION C

TESTING POTENTIAL MECHANISMS OF ASSOCIATION BETWEEN CARDIOMETABOLIC DISORDERS AND SCHIZOPHRENIA

Section C Summary

Having established evidence for the direction and specificity of association between glucose-insulin homeostasis and psychosis in Section B, in Section C I have examined for potential shared mechanisms for cardiometabolic disorders and schizophrenia, focussing on a shared genetic basis and inflammation.

In Chapter 3, I used prospective ALSPAC data from up to 7,977 participants to examine whether genetic predisposition to T2D is associated with risk of schizophrenia-spectrum outcomes in early adulthood and *vice versa*. I also examined whether genetic predisposition to T2D or schizophrenia influences childhood inflammation, and whether this mediates the associations with risk of psychosis or T2D, respectively. Findings from this study have been published in *Schizophrenia Research* (Perry et al., 2020a). See Appendix C Manuscript 1 for the published manuscript.

In Chapter 4, I have used summary data from large-scale GWAS to rigorously examine, using multiple complementary genomic methods, for shared genetic overlap between schizophrenia, cardiometabolic and inflammatory traits. I also examined for a biologically plausible genetic common-causal basis for the physical and psychiatric traits.

In Chapter 5, I have conducted two-sample, uni- and multivariable MR analysis of summary data from large-scale GWAS to explore whether there is likely to be an unconfounded association between disruption to glucose-insulin homeostasis and schizophrenia; to further explore the direction of association between cardiometabolic traits and schizophrenia; and, to examine whether inflammation may be a common biological mechanism for comorbid cardiometabolic disorders and schizophrenia. Findings from this study have been published in *PLOS Medicine* (Perry et al., 2021a). See Appendix C Manuscript 2 for the published manuscript.

Together, these studies build a consistent body of evidence that indicates that a summation of genetic variation may influence biological pathways leading to changes in inflammatory pathways/immune function, which in turn simultaneously increases the risk of both disrupted glucose-insulin homeostasis and schizophrenia. The findings from these studies can help to explain why disruption to cardiometabolic indices can be detected from the onset of psychosis in young adults in the absence of chronic lifestyle or treatment factors, and in light of the results of Section B, may be detectable from childhood/adolescence, years before the onset of psychosis.

Chapter 3

Associations of Genetic Liability for Type 2 Diabetes and Schizophrenia with Schizophrenia-Spectrum Outcomes, Insulin Resistance, and Inflammation in the ALSPAC Birth Cohort

3.1 Introduction

Shared genetic liability may contribute to the comorbidity between cardiometabolic disorders and schizophrenia (Lin and Shuldiner, 2010). For example, the risks of insulin resistance (Chouinard et al., 2019) and impaired glucose tolerance (Ferentinos and Dikeos, 2012), two key precursors of T2D, are higher in unaffected relatives of patients with psychosis compared with controls. People with comorbid schizophrenia and T2D have a higher genetic predisposition for both disorders than controls (Hackinger et al., 2018), and an association between genetic predisposition for schizophrenia and insulin resistance has been reported in a clinical sample (Tomasik et al., 2019). However, each of these studies is limited by small sample sizes and thus limited statistical power.

Another limiting feature of existing studies is that they have included adult cases of established schizophrenia or T2D or have relied on blood measurements taken in adulthood. Therefore, confounding by cumulative effects of lifestyle and other factors is possible (Reinikainen et al., 2015). Population-based prospective studies have identified early markers of disease risk associated with T2D and schizophrenia. For instance, PEs in adolescence or young adulthood are associated with risk of schizophrenia in adulthood (Zammit et al., 2013, Poulton et al., 2000), and insulin resistance is a precursor of T2D (Martin et al., 1992). To the best of my knowledge, no studies have examined whether genetic predisposition for T2D or schizophrenia are associated with, respectively, PEs or insulin resistance in young adulthood. Demonstrating such associations with early markers of illness in young adults with lessened effects of cumulative lifestyle confounding would be consistent with the idea that shared genetic variation is a common mechanism for comorbid T2D and schizophrenia.

Although existing studies provide some evidence for a shared genetic basis for T2D and schizophrenia, underlying pathophysiologic mechanisms remain unclear. Low-grade inflammation may be one such mechanism, which is associated with insulin resistance (Bowker et al., 2020), T2D (Pradhan et al., 2001) and psychosis (Upthegrove et al., 2014). Population-based longitudinal studies have reported that higher levels of circulating inflammatory markers at baseline are associated with risks of psychosis and disrupted glucose-insulin homeostasis subsequently at follow-up (Khandaker et al., 2014, Perry et al., 2018). MR studies have reported associations of genetic variants regulating inflammatory biomarkers such as IL-6 and CRP with schizophrenia (Hartwig et al., 2017) and T2D (Bowker et al., 2020), suggesting that inflammation may be associated with schizophrenia and disrupted glucose-insulin homeostasis beyond any effects of confounding.

3.2 Aims and Objectives

I examined whether shared genetic variation and inflammation could be common mechanisms for T2D and psychosis using prospective, population-based data from the ALSPAC birth cohort. I tested whether: (i) genetic predisposition for T2D is associated with risk of schizophrenia-spectrum outcomes at age 18 years; (ii) genetic predisposition for schizophrenia is associated with insulin resistance at age 18 years; (iii) these associations may be mediated by childhood CRP or IL-6 levels measured at age 9 years.

3.3 Methods

3.3.1 Description of Cohort and Sample Selection

See Section 2.3.1 for a full description of the ALSPAC cohort. This study received ethics approval from the ALSPAC Ethics and Law Committee and local research ethics committees. All participants provided written or implied informed consent. In total, 7,977 participants had genotyping data, 3,768 participants had data on both genotyping and psychosis outcomes, and 2,344 participants had data on genotyping and insulin resistance as outcome. See Appendix C Figure 1 for a flow-chart of participants in the study. The analysis was conducted on participants without missing data for the covariates or outcomes of interest.

3.3.2 Assessment of Schizophrenia-Spectrum Outcomes at Age 18 Years

3.3.2.1 Psychotic Experiences

PEs were identified through the face-to-face, semi-structured PLIKSi conducted by trained psychology graduates. The PLIKSi comprised of an introductory set of questions on unusual experiences, and then 12 ‘ core’ questions eliciting key symptoms covering the three main domains of positive psychotic symptoms: hallucinations (visual and auditory); delusions (delusions of being spied on, persecution, thoughts being read, reference, control, grandiose ability and other unspecified delusions); and symptoms of thought interference (thought broadcasting, insertion and withdrawal). For these 12 core items, 7 stem questions were derived from the Diagnostic Interview Schedule for Children–IV (DISC–IV) and 5 stems from section 17-19 of the Schedules for Clinical Assessment in Neuropsychiatry version 2.0 (SCAN 2.0). After cross-questioning, interviewers rated PEs as not present, suspected, or definitely present. Interviewers rated down (i.e., suspected rather than definite, or none rather than suspected) if unsure. For suspected or definite PEs, interviewers also recorded the

frequency; effects on social/educational/ occupational function; help seeking; and attributions including fever, hypnopompic/ hypnogogic state, or illicit drugs. For interrater reliability, the interviewers recorded audio interviews at three time points, approximately 6 months apart, across the clinic duration (75 interviews in total). The average kappa value of PEs was 0.83, with no evidence of differences across time. Test-retest reliability was assessed using 162 individuals reinterviewed after approximately 47 days (kappa=0.76, SE=0.078), 46 of whom were reinterviewed by the same interviewer (kappa=0.86, SE=0.136). The primary outcome was presence of *definite* PEs, referring to at least one definite PE since age 12 years; the comparator group was suspected/no PEs. The outcome is therefore reflective of a six-year period prevalence of definite PEs. From the total number of participants with definite PEs at age 18 years (230, 4.9%), 80 participants (45.3%) had suffered definite PEs at least once in the month preceding assessment. From the total sample of participants reporting definite PEs, 146 participants (63.5%) reported auditory hallucinations, 63 participants (28.2%) reported any delusion, and 22 participants (9.9%) reported thought disturbance. See the main reporting study for further information (Zammit et al., 2013).

3.3.2.2 Psychotic Disorder

Psychotic disorder was defined (Zammit et al., 2013) as the presence of PEs when symptoms were not attributable to fever/sleep/drugs, had occurred at least once per month over the previous six months, and caused significant distress resulting in either help-seeking from a professional source (general practitioner, counsellor, mental health team), or significantly disrupted social/occupational function. From the total ALSPAC sample who underwent the PLIKSi, 46 participants (1.0%) met criteria for psychotic disorder. Psychotic disorder was included as a secondary outcome due to its lower prevalence in the study sample.

3.3.3 Assessment of Insulin Resistance at Age 18 Years

Insulin resistance was calculated as a binary variable from fasting plasma glucose and fasting insulin levels at age 18 years, using the validated HOMA-IR equation (Levy et al., 1998). There is no consensus-agreed cut-off for insulin resistance based on HOMA-IR in the literature since levels vary between populations (Wallace et al., 2004). Therefore, I used the 75th centile of the study population to define insulin resistance. The 75th centile cut-off has been used in previous research (Hedblad et al., 2000, Marques-Vidal et al., 2002, Geloneze et al., 2006, Cediel et al., 2016). In the ALSPAC sample, The 75th centile for HOMA-IR was 2.15.

3.3.4 Assessment of Polygenic Risk Scores for T2D and Schizophrenia

From the ALSPAC cohort, 8,812 participants were genotyped using the Illumina HumanHap550 quad genome-wide SNP genotyping platform by *23andMe* subcontracted to the Wellcome Trust Sanger Institute, Cambridge, UK and the Laboratory Corporation of America, Burlington, NC, USA. Individuals were excluded from further analysis by ALSPAC based on gender mismatches, minimal or excessive heterozygosity, disproportionate levels of individual missingness (>3%), evidence of cryptic relatedness (>10% of alleles identical by descent) and being of non-European ancestry (assessed by multidimensional scaling analysis including HapMap 2 individuals). Imputation of the target data was performed using Impute V2.2.2 against the 1000 genomes reference panel (Phase 1, Version 3; all polymorphic SNPs excluding singletons), using 2186 reference haplotypes (including non-Europeans), by ALSPAC. Following quality control assessment, imputation, and restricting to 1 young person per family, genetic data was available for 7,977 ALSPAC participants.

Polygenic risk scores (PRS) for schizophrenia and T2D were constructed for all 7,977 participants with genotype data, using training sets based on the second Psychiatric Genomics Consortium (PGC) Schizophrenia GWAS (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014) and a large T2D GWAS (Mahajan et al., 2014), respectively. Both GWAS analyses adjusted for principal components to reduce the impact of population stratification (Price et al., 2006). PRS were calculated using the PLINK (v1.9) (Chang et al., 2015, Purcell et al., 2007) 'score' command following the methodology described by the International Schizophrenia Consortium (Purcell et al., 2009). Prior to construction of scores, SNPs were removed from the analysis if they had a minor allele frequency (MAF) less than 0.01, an imputation quality less than 0.8 or if there was allelic mismatch between samples. Due to the presence of strand differences between ALSPAC and the T2D GWAS, and lack of allele frequency information in the T2D summary statistics, palindromic SNPs were also removed prior to construction of the T2D PRS. Because of the high linkage disequilibrium (LD) within the extended major histocompatibility complex (MHC; chromosome 6: 25-34Mb) only a single SNP was included to represent this region. SNPs were pruned for LD using the PLINK 'clump' command to remove SNPs in LD ($r^2 > 0.25$) with a more significant SNP in the training set. Windows of 500kb were used to assess inter-SNP LD for pruning.

For the primary analysis, PRS were constructed using a list of SNPs with the optimal *p*-value thresholds to capture phenotypic variance defined by both GWAS individually ($p \leq 10^{-5}$ for T2D (Mahajan et al., 2014) and $p \leq 0.05$ for schizophrenia (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014)). Scores were weighted by the logarithm of the odds ratio (OR) for schizophrenia or T2D reported by the GWAS training sets, for the schizophrenia and T2D PRS respectively. Ten principal components (PCs) were generated using unrelated individuals (IBS

< 0.05) and independent SNPs (with long range LD regions removed) using the `-- pca` command in PLINK v1.9. All PRS analyses were adjusted for the 10 PCs to reduce the risk of population stratification. Two PRS measures were calculated for T2D; the first including all SNPs associated with T2D, and the second after excluding a SNP located in the *FTO* gene region, which is understood to be associated with T2D only through its influence on BMI variation (Frayling et al., 2007); the latter was used in sensitivity analysis. Additionally, since the optimal *p*-value thresholds of both PRS scores differed, sensitivity analyses were conducted to examine PRS-outcome associations using a range of *p*-value thresholds from p=0.50 to genome-wide significance ($p<5x10^{-8}$).

3.3.5 Measurement of IL-6 and CRP at Age 9 Years

Data on IL-6 and CRP were available from 5,076 and 5,086 participants respectively. Blood samples were collected at non-fasting state, frozen at −80°C, and assayed in 2008 after a median of 7.5 years in storage. There was no evidence of freeze-thaw cycles during storage period. IL-6 was measured by ELISA (R&D systems, Abingdon, UK), and CRP was measured by automated particle-enhanced immunoturbidimetric assay (Roche UK, Welwyn Garden City, UK). All assay coefficients of variation were <5%. The minimum detection limit for IL-6 was 0.1pg/mL. This represents the lowest measurable analytic level that can be distinguished from zero. Those below this limit were assigned a value of zero (0.4% of the sample) and were also included in the analysis. The minimum detection limit for CRP was 0.03mg/L. Twenty-nine participants (0.6% of the sample) were below this limit and were assigned values of 0.01 (n=16) and 0.02 (n=13); they were also included in the analysis. 32 subjects had CRP levels >10mg/L and were excluded from analysis due to the risk of acute inflammatory state such as infection, which may have confounded results.

3.3.6 Assessment of Potential Confounders

I included sex (categorical), ethnicity (binary caucasian / non-caucasian due to the predominantly caucasian sample), social class (categorical) and BMI at age 18 years (continuous).

3.3.7 Statistical Analysis

I examined the distribution of PRS-T2D and PRS-schizophrenia using the Shapiro-Wilk test for normality, and from visual inspection of Q-Q plots. The distributions were p>0.05 and appeared normally distributed. PRS variables were standardized (Z-transformed).

3.3.7.1 Association between PRS and Outcomes at Age 18 Years

I conducted logistic regression analyses to examine the association between PRS-T2D and risks for schizophrenia-spectrum outcomes, and PRS-schizophrenia and insulin resistance at age 18 years. ORs indicate the increase in risk of outcome per SD increase in PRS. *P*-values for adjusted regression models in the primary analysis were corrected for multiple testing per the three outcomes included (definite PEs, psychotic disorder and insulin resistance) using the Holm-Bonferroni method (Holland, 1987). To test for linearity of associations, I included a quadratic term (PRS^2) in the logistic regression models.

3.3.7.2 Association between PRS scores and Childhood Inflammatory Markers at Age 9 Years

I used linear regression analyses to test associations of PRS for T2D or schizophrenia, separately, with IL-6 and CRP levels at age 9 years (Z-transformed values), before and after adjustments for potential confounders listed above.

3.3.7.3 Mediation by Childhood Inflammatory Markers at Age 9 Years

I performed mediation analyses to examine whether any evident associations may be mediated by any childhood inflammatory markers that also showed evidence for associations with PRS. I calculated direct and indirect effects between exposure (PRS-T2D or PRS-schizophrenia) and outcome (e.g., PEs or insulin resistance) taking into account the mediator variable (e.g., CRP). Evidence of an indirect effect is consistent with mediation. The indirect effect was bootstrapped using 5000 iterations to determine the 95% CIs.

3.3.8 Missing Data

I assessed the potential impact of missing data by comparing mean PRS score between the analytic sample and participants with missing data for psychosis and insulin resistance outcomes, using separate variance t-tests. I also performed logistic regression analysis to determine sociodemographic and other predictors (sex, ethnicity, BMI, and social class) of missing data.

3.4 Results

3.4.1 Baseline Characteristics of The Analytic Sample

Of the 3,768 participants with data on PRS-T2D and schizophrenia-spectrum outcomes, 283 met the criteria for suspected/definite PEs (7.5%), 183 for definite PEs (5.1%), 29 (0.7%) for psychotic disorder at age 18 (Table 15). Of the 2,344 participants with data on PRS-schizophrenia and insulin resistance, 173 met the criteria for insulin resistance at age 18 (7.3%).

Table 15: Baseline Characteristics of the Analytic Sample

Characteristic, *n* (%) unless otherwise stated	**All Sample**	**Definite PEs**	**Psychotic Disorder**	**No/Suspected PEs**
Male Sex	1846 (49)	71 (38)	7 (15)	1775 (49)
White British Ethnicity	3692 (98)	179 (98)	39 (95)	3513 (98)
Social Class				
I & II	1,582 (42)	62 (35)	5 (16)	1,456 (40)
III - non manual & manual	1,616 (43)	75 (43)	15 (48)	1,630 (44)
IV & V	565 (15)		11 (36)	
		38 (22)		583 (16)
BMI (kg/m^2) at 18 years, mean (SD)	22.71 (3.76)	23.37 (4.49)	22.73 (4.26)	22.60 (3.71)
HOMA at 18 years, mean (SD)	0.92 (0.73)	1.03 (0.75)	1.28 (1.00)	0.92 (0.73)
Insulin Resistance	251 (8)	25 (17)	7 (20)	209 (7)
Current Smoking	220 (7)	22 (15)	5 (18)	188 (7)
CRP (mg/L) at 9 years, mean (SD)	0.68 (2.52)	0.72 (2.61)	0.75 (1.33)	0.67 (2.49)

BMI=body mass index; HOMA=homeostatic model assessment for insulin resistance; CRP=C-reactive protein; PE=psychotic experiences

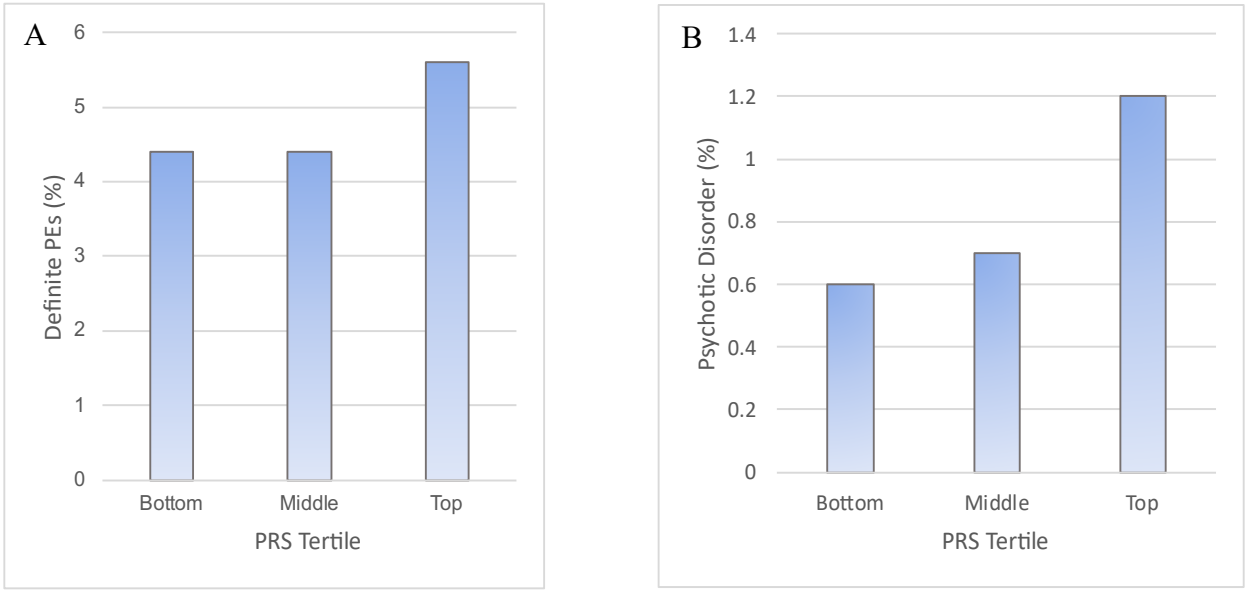
3.4.2 The Association of Genetic Predisposition for T2D with Schizophrenia-Spectrum Outcomes at Age 18 Years

The prevalence of schizophrenia-spectrum outcomes at age 18 years was higher for participants in the top third of PRS-T2D distribution compared with those in the bottom third (Figure 8). PRS-T2D was associated with definite PEs (adjusted OR=1.21; 95% CI, 1.01-1.45 per SD increase in PRS-T2D) and psychotic disorder (adjusted OR=1.51; 95% CI, 1.04-2.05 per SD increase in PRS-T2D) at age 18 years after controlling for sex, ethnicity, social class, and BMI (Table 16). Quadratic terms for PRS-T2D in these regression models were non-significant suggesting no evidence for departure from linearity (all $p>0.05$). The results for sensitivity analyses using PRS-T2D score excluding a SNP in the *FTO* gene region were similar (Table 17).

3.4.3 Association between Genetic Predisposition for Schizophrenia and Insulin Resistance at Age 18

There was weaker evidence for an association between PRS-schizophrenia and insulin resistance at age 18 (adjusted OR=1.10; 95% CI, 0.99-1.22 per SD increase in PRS-schizophrenia) after controlling for sex, ethnicity, social class, and BMI. The quadratic term for PRS-schizophrenia was non-significant suggesting no evidence for departure from linearity (p>0.05).

Figure 8: The Prevalence of Schizophrenia-Spectrum Outcomes at Age 18 Years Per Tertile of Genetic Risk for Type 2 Diabetes



A=Prevalence of psychotic experiences per tertile of genetic risk for type 2 diabetes; B=Prevalence of psychotic disorder per tertile of genetic risk for type 2 diabetes

Table 16: Odds Ratios (95% CIs) for Outcomes at Age 18 Years Per SD Increase in Genetic Risk for Type 2 Diabetes or Schizophrenia

Outcome / Risk Factor	Sample	OR (95% C.I.)		*p*-value	Corrected *p*-value[c]
		Unadjusted[a]	Adjusted for sex, ethnicity, social class and BMI[b]		
Definite PEs					
PRS-T2D	3,768	1.15 (0.99-1.34)	1.21 (1.01-1.45)	0.020	0.049
Psychotic Disorder					
PRS-T2D	3,768	1.42 (1.00-1.96)	1.51 (1.04-2.05)	0.016	0.042*
Insulin Resistance					
PRS-SCZ	2,344	1.16 (1.04-1.32)	1.10 (0.99-1.22)	0.089	0.089

PRS-T2D=polygenic risk for type 2 diabetes; PRS-SCZ=polygenic risk for schizophrenia; [a] The unadjusted analysis was adjusted for 10 principal components only; [b] Samples for adjusted analysis included 3,070 participants for psychosis outcomes and 1,970 participants for insulin resistance outcome; [c]*p*-value corrected from adjusted analysis using Holm-Bonferroni method; *evidence surpasses Holm-Bonferroni threshold

Table 17: Odds Ratios (95% CIs) for Schizophrenia-Spectrum Outcomes at Age 18 Years Per SD Increase in Genetic Risk for Type 2 Diabetes Excluding *FTO* Associated SNP

Risk Factor/outcome	Sample	OR (95% C.I.)		*p*-value	Corrected *p*-value[c]
		Unadjusted[a]	Adjusted for sex, ethnicity, social class and BMI[b]		
Definite PEs					
PRS-T2D without *FTO*	3,768	1.15 (0.99-1.34)	1.21 (1.02-1.46)	0.025	0.051
Psychotic Disorder					
PRS-T2D without *FTO*	3,768	1.42 (1.01-1.96)	1.50 (1.04-2.03)	0.016	0.048*

PRS-T2D=polygenic risk for type 2 diabetes; [a] The unadjusted analysis was adjusted for 10 principal components only [b] Samples for adjusted analysis included 3,070 participants; [c]*p*-value corrected from adjusted analysis using Holm-Bonferroni method; *evidence surpasses Holm-Bonferroni threshold

3.4.4 Associations of PRS Scores with Inflammatory Markers at Age 9 Years

Data on both PRS scores and serum IL-6 and CRP levels were available for 2,180 and 2,176 participants, respectively. After adjustments for sex, ethnicity, social class, and BMI, PRS-T2D was associated with CRP (β=0.03; 95% CI, 0.01-0.08, *p*=0.040), but not with IL-6 (β=0.01; 95% CI, -0.02 – 0.05, *p*=0.082). There was also trend level evidence for an association between PRS-schizophrenia and CRP (β=0.05; 95% CI, -0.01 – 0.10, *p*=0.061) but not with IL-6 (β=0.01; 95% CI, -0.04-0.09, *p*=0.670).

3.4.5 The Mediating Effect of Childhood CRP Levels on the Associations of PRS scores with Schizophrenia-Spectrum Outcomes or Insulin Resistance

Based on 1,955 participants with data on PRS-T2D, CRP levels at age 9 and PEs at age 18 years, CRP at age 9 years partially mediated the association between PRS-T2D and definite PEs at age 18 years. There was evidence of an indirect effect indicative of mediation; the coefficients were 0.28; 95% CI, 0.07-0.45, p=0.044 for direct effect; co-efficient=0.05; 95% CI 0.02-0.12, p=0.040 for indirect effect. Due to the low number of participants meeting the criteria for psychotic disorder I could not reliably test the mediation model with psychotic disorder as outcome. Since IL-6 levels at age 9 years were not associated with PRS-T2D, I did not perform mediation analysis using IL-6. There was no evidence for a mediating effect of CRP on the association between PRS-schizophrenia and insulin resistance at age 18; the coefficients were 0.14; 95% CI, -0.06-0.34, p=0.756 for direct effect; co-efficient=0.01; 95% CI, -0.01-0.03, p=0.180 for indirect effect.

3.4.6 Results for Sensitivity Analysis Using Different *P*-Value Thresholds for PRS

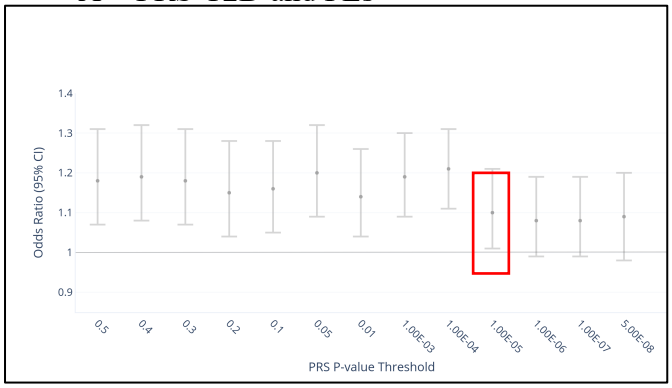
Figure 9 presents the associations between PRS-T2D and PEs alongside the associations between PRS-schizophrenia and insulin resistance, at different PRS p-value thresholds. The point estimates for the PRS-T2D-PEs associations were >1 for all p-value thresholds, though the strength of association weakened at more stringent p-value thresholds. A similar pattern was observed for the PRS-schizophrenia-insulin resistance association, where the evidence for a positive association attenuated at p-value thresholds more stringent than 1.00×10^{-4}.

3.4.7 Missing Data

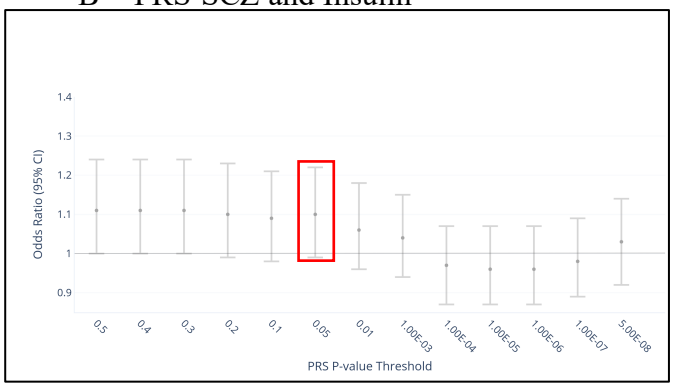
Fifty-three percent of participants with data on PRS-T2D had psychotic outcomes data missing, and 71% of participants with PRS-schizophrenia had insulin resistance outcome data missing (Appendix C Figure 1). Compared with the analytic sample, the missing sample had higher mean PRS-schizophrenia but lower PRS-T2D scores (Table 18). Male sex, lower social class and higher BMI predicted missing data for psychotic outcomes, and non-white ethnicity was associated with having missing data for insulin resistance (Table 19).

Figure 9: The Association of PRS Score and Outcome at Age 18 Years Across a Range of PRS *p*-value Thresholds

A = PRS-T2D and PEs



B = PRS-SCZ and Insulin



Point estimate highlighted in red box represents association at GWAS-defined optimum *p*-value threshold for the exposure. PRS-SCZ = Polygenic Risk Score for Schizophrenia; PRS-T2D = Polygenic Risk Score for Type 2 Diabetes; PEs = Psychotic Experiences

Table 18: Mean PRS-T2D and PRS-Schizophrenia in the Analytic and Missing Samples Compared Using Separate Variance T-Test

Outcome	***n***	**Mean Z-transformed PRS Score**	**Test Statistic, *p*-value**
PRS-T2D with Missing Psychosis-Risk Data			
Analytic sample	3,768	0.033	
Missing sample	4,209	-0.252	t=3.2, p=0.002
PRS-Schizophrenia with Missing Insulin Resistance Data			
Analytic sample	2,344	-0.083	
Missing sample	5,633	0.334	t=-4.7, p<0.001

Table 19: Predictors of Missing Outcome Data at Age 18 Years for Participants with Data on PRS

Predictor/Outcome	*n* (%) with data[1]	OR (95% C.I.)	*p*-value
Psychosis outcomes			
Sex[2]	7,870 (99)	1.69 (1.55-1.84)	<0.001
Ethnicity[3]	7,876 (89)	0.62 (0.11-3.40)	0.583
Social Class[4]	7,060 (89)	1.11 (1.08-1.13)	<0.001
BMI	5,062 (63)	1.05 (1.02-1.07)	0.001
Insulin Resistance			
Sex[2]	7,870 (99)	1.00 (0.92-1.03)	0.203
Ethnicity[3]	7,876 (89)	0.51 (0.42-0.60)	<0.001
Social Class[4]	7,060 (89)	1.01 (0.98-1.04)	0.827
BMI	5,062 (63)	1.06 (0.91-1.17)	0.547

[1]*n* with predictor from risk set of all participants with data on PRS (*n*=8,812); [2]Female sex is reference
[3]White-British is reference; [4]Social Class I is reference

3.5 Discussion

Using prospective birth cohort data, I found that genetic predisposition for T2D is associated with schizophrenia-spectrum outcomes at age 18 years in a linear fashion. The PRS-T2D findings were consistent using two genetic scores, one with and one without a SNP at the *FTO* locus, which is related to BMI (Frayling et al., 2007). Additionally, there was evidence for a dose-response pattern in the association between PRS-T2D and schizophrenia-spectrum outcomes; the effect size was largest for psychotic disorder, which is a more clinically relevant outcome than PEs. I also found some evidence, albeit slightly weaker, for an association between genetic predisposition for schizophrenia and insulin resistance at age 18 years. However, the sample of participants with missing data had higher mean PRS-schizophrenia scores than included participants, thus missing data may help explain the weaker evidence. Nonetheless, the findings provide evidence that the comorbidity between disrupted glucose-insulin homeostasis and schizophrenia arises partly due to shared genetic factors.

The point estimates across various *p*-value thresholds were similar in both combinations of genotype-phenotype analysis. However, in both cases, at more stringent *p*-value thresholds, the evidence of association weakened. This weakening effect is consistent with a previous study examining the association between PRS-schizophrenia and adolescent psychopathology (Jones et al., 2016), which also reported that PRS-schizophrenia was associated with attrition. Therefore, type II statistical error may contribute to the weaker associations between PRS-schizophrenia and insulin resistance at age 18 years.

The results of this study are in line with previous research, which found that people with comorbid schizophrenia and T2D have a higher genetic predisposition to both disorders compared to controls (Hackinger et al., 2018). The findings also align with a report of an association between PRS for schizophrenia and insulin resistance in a clinical sample of people with schizophrenia (Tomasik et al., 2019). Another study found evidence for a genetic overlap of schizophrenia with triglycerides and HDL (Andreassen et al., 2013), which are associated with insulin resistance (Laws and Reaven, 1992). However, one previous study found no evidence for an association between PRS-T2D and schizophrenia (Padmanabhan et al., 2016), though it featured a much smaller sample size than the present study and may have been underpowered to detect a difference.

Genetic liability for T2D or schizophrenia may increase the risk of both disorders via pleiotropic mechanisms. For example, genetic liability for schizophrenia may influence inflammatory pathways (Slopen et al., 2013), leading to disrupted glucose-insulin homeostasis. I found some evidence for the association of childhood CRP levels with both PRS-T2D and PRS-schizophrenia. However, I did not find an association with IL-6. This is perhaps unexpected since IL-6 stimulates the production of CRP (Calabro et al., 2003) and is associated with psychosis (Khandaker et al., 2014) and insulin resistance (Bowker et al., 2020).

However, it is also possible that genetic predisposition for T2D or schizophrenia influences CRP via mechanisms other than the IL-6 pathway. CRP plays an active role in hepatic insulin resistance, partly through impairment in insulin signalling independent of IL-6 (Xi et al., 2011). Interestingly, CRP has shown to be protective of schizophrenia in MR studies (Hartwig et al., 2017). However, the GWAS studies included in previous MR research measured phenotypic markers in adults. I used CRP measured in childhood, which may be reflective of a distinct biological environment.

I found evidence that genetic predisposition for T2D may influence risk of psychosis in early adulthood by increasing inflammation in childhood. Still, the magnitude of this mediating effect was small, suggesting that other mechanisms are likely to be involved. On the other hand, I found no evidence that childhood IL-6/CRP mediated the association between genetic predisposition for

schizophrenia and insulin resistance. The mediating effect of inflammation for the outcome of PEs is consistent with previous research reporting an association between genetic risk for schizophrenia and immune-related disorders (Stringer et al., 2014, Tylee et al., 2018).

Due to the relatively small number of participants with psychotic disorder in the sample and associated lack of power, I did not consider testing psychotic disorder in mediation analyses. Future longitudinal research conducted on larger samples may seek to perform a mediation analysis of CRP between PRS-T2D and more clinically relevant schizophrenia-spectrum outcomes. Other mediators for PRS-T2D and schizophrenia-spectrum outcomes may include non-immune mechanisms such as pleiotropic genes affecting distinct biological pathways. For example, a study examining the genetic overlap between T2D and schizophrenia highlighted, among others, *PROX1* as a potentially pleiotropic locus (Hackinger et al., 2018). *PROX1* acts both as a transcriptional activator and repressor and is implicated in human pancreatic beta-cell development and neurogenesis (Holzmann et al., 2015).

Strengths of this study include a larger sample compared with previous research in the field; considering different genetic scores for T2D to address the potential pleiotropic effect of PRS-T2D on BMI, and in using childhood inflammatory markers in a mediation model to test a hypothesis that inflammation may be a biological mechanism of association. Since the exposures in the study were genetic risk, the potential for confounding by environmental and lifestyle factors is limited. In addition, I was able to control for potential confounding effects of sex, BMI, social class, and inflammatory disease. Regarding ethnicity, participants of non-European genetic ancestry were removed at the stage of genotyping analysis. I also adjusted regression analyses for ethnicity since ethnicity is significantly associated with T2D-risk (Oldroyd et al., 2005). I adjusted for PCs (Price et al., 2006) in PRS analyses to further reduce the risk of population stratification bias.

An important limitation is missing data. Over half of the risk set with data on PRS had outcome data missing at follow-up. The missing sample had a higher mean score for PRS-schizophrenia but a lower mean score for PRS-T2D. Thus, the analyses may underestimate the true association between genetic predisposition for schizophrenia and insulin resistance. In contrast, the opposite might be the case for the association between PRS-T2D and schizophrenia-spectrum outcomes.

Furthermore, whilst PEs have been shown to reflect an increased risk for psychotic disorders (Zammit et al., 2013, Sullivan et al., 2020), and PEs lie on a continuum with clinical psychosis in the general population (van Os et al., 2009), the transition from PEs to clinical psychosis is low (Kaymaz et al., 2012) and PEs are also associated with other psychiatric phenotypes such as depression and anxiety disorders. Additionally, since the schizophrenia-spectrum outcomes were measured before the peak

age of psychosis onset (Eranti et al., 2013), some participants may not have yet developed psychotic symptoms. This point also applies to the sample of participants meeting the criteria for insulin resistance at age 18 years. Whilst I attempted to address these limitations by reversing the genotype and phenotype to more accurately capture schizophrenia/T2D liability, replication of the methods in a clinical sample is necessary. In addition, certain antipsychotic medications can have adverse effects on glycaemic indices (Leucht et al., 2013). At present, ALSPAC does not have treatment record linkage, so I could not adjust for antipsychotic treatment. This may have impacted the analyses examining PRS-schizophrenia and insulin resistance since a higher genetic predisposition for schizophrenia may be associated with antipsychotic use. Finally, one-off measurements of inflammatory markers in childhood may not reflect lifelong levels of inflammation. However, if non-differential, measurement error introduces a bias towards the null, so the results may underestimate the true association between PRS-T2D and IL-6 and CRP.

In conclusion, this study provides evidence that a summation of minor genetic variation, set at conception, representing lifetime risk for T2D or schizophrenia, may contribute a portion of the variance of the comorbidity of these disorders in adulthood. This genetic variation may influence inflammatory pathways to increase the risk of comorbidity. In future, similar research may seek to examine the associations between PRS for T2D and other mental disorders, including depression or bipolar disorder, both of which are known to have higher rates of cardiometabolic disorders than the general population (Martin et al., 2016). Such research may also help to test the specificity of the findings in this study.

Chapter 4

Evidence for Shared Genetic Aetiology between Schizophrenia, Cardiometabolic and Inflammatory Traits: Genetic Correlation and Colocalization Analyses

4.1 Introduction

The cardiometabolic comorbidity of schizophrenia is traditionally attributed to lifestyle factors, such as smoking and physical inactivity, and the adverse effects of antipsychotic medication (Leucht et al., 2013). However, cardiometabolic dysfunction is detectable in antipsychotic-naïve young adults with FEP, suggesting that lifestyle factors/medication may not be the full explanation (Perry et al., 2016).

For example, schizophrenia and cardiometabolic disorders share similar associations with elevated concentrations of circulating inflammatory markers such as CRP and IL-6, both cross-sectionally (Upthegrove et al., 2014, Wang et al., 2013) and longitudinally (Khandaker et al., 2014, Bowker et al., 2020). MR studies have similarly shown that genetically predicted levels of IL-6 and CRP could be causally linked with cardiometabolic disorders (Georgakis et al., 2020) and schizophrenia (Hartwig et al., 2017).

Therefore, schizophrenia, cardiometabolic and inflammatory traits could share pathophysiologic mechanisms, including a common genetic basis. An improved understanding of the mechanisms underlying the comorbidity between schizophrenia, cardiometabolic and inflammatory traits is pivotal to inform novel approaches to treatment and prevention.

Previous studies have predominantly used LDSC (Bulik-Sullivan et al., 2015a) to estimate the whole-genome correlation between schizophrenia and cardiometabolic traits, with one recent study reporting evidence of partial genetic similarity between schizophrenia and BMI (Bahrami et al., 2020). There is limited evidence for other cardiometabolic and inflammatory traits (Bulik-Sullivan et al., 2015a).

However, the LDSC approach may have limitations. First, LDSC could be susceptible to the 'missing heritability' problem, where subtle population stratification downwardly bias the effects of lower-frequency variants (Mathieson and McVean, 2012). Therefore, genetic correlation analysis which considers the relative frequency of variants is required. Second, LDSC estimates may be biased towards the null when opposing mechanisms exist (e.g., regions of positive and negative correlation nullifying each other when averaged (Shi et al., 2017)). This may be expected in a relatively heterogeneous condition like schizophrenia (Wolfers et al., 2018). Therefore, more fine-grained locus level genetic correlation analysis is required to identify genomic regions of interest. Third, while LDSC can provide evidence of overall genomic similarity between traits, it cannot provide information with which to consider biological plausibility or infer potential causality.

4.2 Aims and Objectives

I aimed to use a range of complementary genomic approaches in a stepwise manner to rigorously examine the potential for a common genomic basis for schizophrenia and a range of cardiometabolic and inflammatory traits previously reported to be associated with it (Vancampfort et al., 2015, Miller et al., 2014). I aimed to identify specific putative biological pathways underpinning the comorbidity and address limitations of previous approaches. First, in addition to LDSC to estimate the genome-wide correlation between traits, I used genetic covariance analyser (GNOVA (Lu et al., 2017)), a recent methodological extension of LDSC, to estimate genetic correlation after stratifying variants by MAF. Second, I used Heritability Estimation from Summary Statistics (ρ-HESS (Shi et al., 2017)) to identify positive or negative regions of locus-level genetic correlation that otherwise may be masked by LDSC. Finally, to estimate colocalization between clusters of traits and to identify putative common-causal variants amongst locally correlated genomic regions, I used hypothesis prioritization multi-trait colocalization (HyPrColoc (Foley et al., 2021)), a novel type of multi-trait colocalization analysis (Giambartolomei et al., 2018).

4.3 Methods

4.3.1 Summary Statistics for Schizophrenia, Cardiometabolic and Inflammatory Traits

For schizophrenia, I used publicly available summary data from the Psychiatric Genomics Consortium (PGC) (40,675 cases, 64,643 controls (Pardinas et al., 2018)). I used publicly available summary GWAS data for twelve cardiometabolic and inflammatory traits (fasting insulin, HOMA-IR, T2D, FPG, glucose tolerance, HbA1C, LDL, HDL, triglycerides, BMI, CAD, and CRP) from large-scale consortia (Table 20). All GWAS were conducted in mostly European samples and adjusted for population stratification, age, and sex. Ethical approval was obtained by the original GWAS authors as per each individual GWAS protocol.

Table 20: Summary GWAS Data Used For Cardiometabolic and Inflammatory Traits

Trait	Consortium (Author, Year)	Sample Size	Cases/Controls[a]	Participant Description	PMID
Schizophrenia	PGC (Pardinas et al., 2018)	105,318	40,675 / 64,643	European Adults	29483656
Fasting Insulin	MAGIC (Lagou, 2019)	140,595	-	European Adults	-
FPG	MAGIC (Scott et al., 2012)	133,010	-	European Adults	22885924
HOMA-IR	MAGIC (Dupuis et al., 2010)	46,186	-	European Adults	20081858
Glucose Tolerance	MAGIC (Scott et al., 2012)	42,854	-	European Adults	22885924
HbA1C	MAGIC (Wheeler et al., 2017)	123,665	-	European Adults	28898252
T2D	Mahajan et al, 2017 (DIAGRAM) (Mahajan et al., 2018)	898,130	74,124 / 824,006	European Adults	30297969
LDL	GLGC (Liu et al., 2017)	237,050	-	European Adults	29083408
HDL	GLGC (Liu et al., 2017)	237,050	-	European Adults	29083408
Triglycerides	GLGC (Liu et al., 2017)	237,050	-	European Adults	29083408
BMI	GIANT and UK Biobank (Pulit et al., 2019)	694,649	-	European Adults	30239722
CAD	CARDIoGRAM C4D (van der Harst et al, 2017) and UK Biobank (van der Harst and Verweij, 2018)	547,261	122,733 / 424,528	European Adults	29212778
CRP	CHARGE (Ligthart et al., 2018)	204,402	-	European Adults	30388399

FPG=fasting plasma glucose; HOMA-IR=homeostatic model assessment for insulin resistance; HbA1C=glycated haemoglobin; T2D=type 2 diabetes; LDL=low-density lipoprotein; HDL=high-density lipoprotein; BMI=body mass index; CAD=coronary artery disease; CRP=C-reactive protein; PGC=Psychiatric genomics consortium; MAGIC=Meta-analyses of glucose and insulin-related traits consortium; DIAGRAM=Diabetes genetics replication and meta-analyses; GLGC=Global lipids genetics consortium; GIANT=Genetic investigation of anthropometric traits; CARDIoGRAM=Coronary artery disease genome wide replication and meta-analysis; C4D=Coronary artery disease genetics consortia; CHARGE=Cohorts for heart and aging research in genomic epidemiology.
[a]Case/Control numbers supplied for binary traits

4.3.2 Statistical Analysis

4.3.2.1 LDSC for Genome-wide Correlations

Genome-wide SNP-heritability estimates (h_2), standard errors (SEs), and genome-wide genetic correlation estimates (r_g) between all trait-pairs were estimated using LDSC (Bulik-Sullivan et al., 2015b) and an LD reference panel from the 1000 Genomes Project's Phase 3 European (1kG CEU) sample. Quality control (QC) steps on each GWAS dataset prior to analysis were: 1) filtering SNPs that were not included within the HapMap3 reference panel or had MAF <5% within the 1kG CEU reference sample; 2) filtering SNPs within the major histocompatibility complex (MHC) due to the complex LD structure within the region (Miretti et al., 2005). I used a Bonferroni-adjusted threshold of $p<0.004$ to define strong evidence of genome-wide genetic correlation.

4.3.2.2 MAF-Stratified Genetic Correlation

MAF-stratified genetic correlations between schizophrenia and other traits were estimated using GNOVA (Lu et al., 2017). GNOVA is an extension to classical LDSC, allowing estimates of genetic correlation across continuous annotations (e.g., MAF). QC methods were the same as above. MAF quartiles were defined by the GNOVA authors (Lu et al., 2017) and calculated using genotyping data from the 1kG CEU reference sample. MAF cut-offs for each quartile were as follows: Q1=0.05-0.11; Q2=0.11-0.22; Q3=0.22-0.35; Q4=0.35-0.50. I used a Bonferroni-adjusted threshold of $p<0.004$ to define significant evidence of MAF-stratified genetic correlation.

4.3.2.3 Locus-Level Genetic Correlation

Next, I explored locus level correlation between schizophrenia and traits with at least nominal evidence of either whole-genome or MAF-stratified genetic correlation. I accepted a less-stringent significance threshold to select traits for locus level correlation analysis to allow for an examination of opposing mechanisms (Shi et al., 2017), which may have biased 'averaged' correlation estimates (e.g., from LDSC or GNOVA) toward the null. I used ρ-HESS (Shi et al., 2017) to estimate partitioned heritability and genetic correlations within pre-defined genomic LD-blocks based on European participants (Berisa and Pickrell, 2016), allowing for greater resolution of the correlation within each LD block. No sample overlap between data from different consortia was assumed, as recommended, due to the heterogeneity of analysed trait-pairs (i.e., a psychiatric trait with a cardiometabolic/inflammatory trait) (Shi et al., 2017). Where no SNPs were available for analysis within a particular LD block, that LD block was removed from analysis of that trait pair. I used a Bonferroni-adjusted threshold dependent on the number of LD blocks tested between pairs of traits to determine significant evidence of locus-level correlation (between $p<3.14\text{x}10^{-5}$ and $p<2.7\text{x}10^{-5}$).

4.3.2.4 Multi-trait Colocalization

To provide greater resolution and allow for a consideration of biological plausibility in genomic regions with evidence of local correlation, I used HyPrColoc (Foley et al., 2021). HyPrColoc estimates the posterior probability of colocalization across multiple traits at a single causal variant by enumerating putative causal configurations. In doing so, HyPrColoc can identify distinct clusters of traits which colocalize at independent putative causal variants within the genomic region of interest. To conduct this stage of analysis, I identified the lead SNP for schizophrenia within each LD-block showing Bonferroni-significant evidence of locus-level correlation with cardiometabolic and

inflammatory traits. For each trait, I included all SNPs located 500kb either side of the schizophrenia lead SNP. I did not consider regions within the MHC. The primary analysis used the recommended variant-specific prior configuration (prior 1=$1x10^{-4}$; prior 2=0.02) and regional and alignment threshold settings (0.5 for both). Assistance with the codes for the colocalization analysis was provided by Dr Nick Bowker (University of Cambridge).

4.3.2.5 Colocalization Sensitivity Analysis

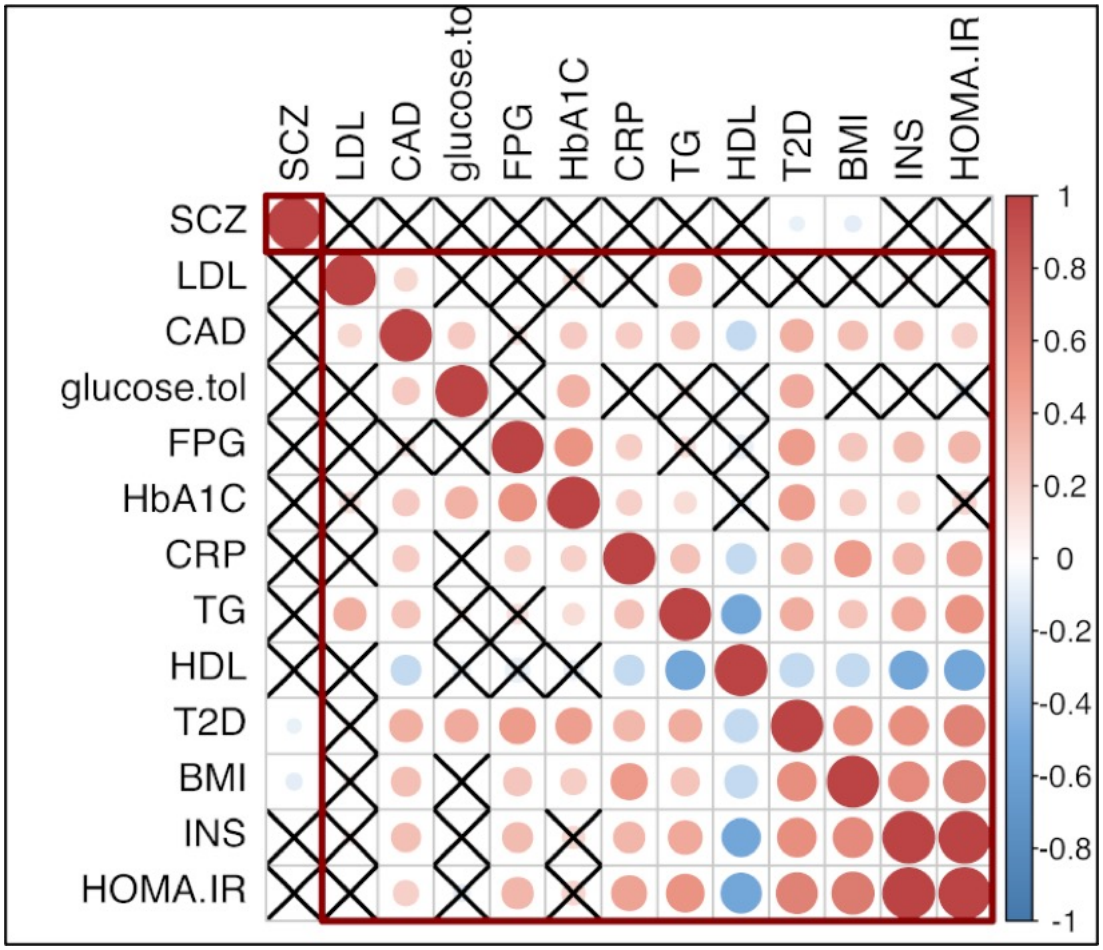
To test the strength of evidence for colocalization and also cluster stability, I repeated colocalization analysis over: a) increasingly stringent prior settings, (0.02, 0.01, 0.001); and b) increasingly stringent regional and alignment threshold settings (0.5, 0.6, 0.7, 0.8, 0.9). To visualise cluster stability across the permutations, heatmaps were drawn based on a similarity matrix between clusters. Where there was evidence for potential colocalization, stacked regional association plots were drawn to visually inspect putative candidate SNPs, their strength of association within each putative colocalized trait, and the LD structure in the genomic region.

4.4 Results

4.4.1 Genome-wide Correlation between Schizophrenia and Cardiometabolic/Inflammatory Traits

Using LDSC, I found Bonferroni-significant evidence of correlation of schizophrenia with BMI (r_g=-0.09; 95% C.I., -0.06, -0.12; p=$1.83x10^{-5}$; h_2=0.21; SE=0.007) and T2D (r_g=-0.07; 95% C.I., -0.03- -0.12; p=0.002; h_2=0.04; SE=0.002). In hierarchical clustering, two clusters were formed: schizophrenia in the first, and all other included traits in the second (Figure 10). See Appendix C Table 1 for complete LDSC results for all trait pairs.

Figure 10: Whole Genome Correlations between Schizophrenia, Cardiometabolic and Inflammatory Traits



X indicates correlations that did not meet Bonferroni-corrected evidential threshold (p=0.004). Hierarchical clusters are indicated with red boxes. SCZ=schizophrenia; LDL=low-density lipoprotein; CAD=coronary artery disease; glucose.tol=two-hour glucose; FPG=fasting plasma glucose; HbA1C=glycated haemoglobin; CRP=C-reactive protein; TG=triglycerides; HDL=high-density lipoprotein; T2D=type 2 diabetes mellitus; BMI=body mass index; INS=fasting insulin; HOMA.IR=homeostatic model assessment for insulin resistance.

4.4.2 MAF-Stratified Genetic Correlation between Schizophrenia and Cardiometabolic / Inflammatory Traits

I found a trend of nominal evidence for correlation in the lowest MAF-quartile between schizophrenia and a range of cardiometabolic and inflammatory traits (fasting insulin (r_g=0.22; p=0.029); triglycerides (r_g=0.14; p=0.020); CAD (r_g=0.24; p=0.025); HDL (r_g=-0.11; p=0.053); T2D (r_g=0.06; p=0.076); CRP (r_g=0.18; p=0.088)); in the second-lowest MAF-quartile between schizophrenia and LDL (r_g=0.06; p=0.037); and in the highest MAF-quartile between schizophrenia and both BMI (r_g=-0.13; p=0.006) and T2D (r_g=-0.12; p=0.012) (Table 21).

Table 21: MAF Stratified Genetic Correlations between Schizophrenia, Cardiometabolic and Inflammatory Traits

Trait	MAF[1]							
	Q1		Q2		Q3		Q4	
	r_g	*p*	r_g	*p*	r_g	*p*	r_g	*p*
T2D†	0.062	0.076	0.045	0.585	-0.056	0.433	-0.120	0.012
FPG	0.078	0.296	0.052	0.971	0.143	0.345	-0.127	0.173
Fasting Insulin†	0.223	0.029	0.110	0.266	0.095	0.174	-0.050	0.790
HOMA-IR	-0.122	0.507	-0.147	0.284	0.056	0.652	-0.067	0.570
Glucose Tolerance	0.030	0.220	-0.112	0.670	-0.063	0.691	0.023	0.842
HbA1C	0.013	0.459	0.032	0.844	-0.097	0.350	0.044	0.414
HDL†	-0.114	0.053	-0.070	0.280	0.043	0.105	0.051	0.572
LDL†	0.053	0.371	0.063	0.037	0.036	0.291	0.039	0.231
Triglycerides†	0.136	0.020	0.025	0.106	-0.003	0.188	0.031	0.201
BMI†	-0.109	0.219	-0.086	0.187	-0.078	0.143	-0.127	0.006
CAD †	0.235	0.025	-0.025	0.761	0.054	0.484	-0.041	0.446
CRP†	0.181	0.088	0.018	0.859	0.091	0.196	-0.049	0.358

HDL=high-density lipoprotein; LDL=low-density lipoprotein; FPG=fasting plasma glucose; BMI=body mass index; T2D=type 2 diabetes mellitus; HOMA-IR=homeostatic model assessment of insulin resistance; HbA1C=glycated haemoglobin; CAD=coronary artery disease; CRP=C-reactive protein; r_g=genetic correlation estimate.; [1]MAF split into quartiles; Q1=lowest to Q4=highest; †indicates traits taken to next stage of analysis based upon nominal evidence of whole or stratified genetic correlation

4.4.3 Locus-Level Genetic Correlation Between Schizophrenia and Cardiometabolic / Inflammatory Traits

All included cardiometabolic and inflammatory traits showed Bonferroni-significant evidence of at least one region of local genetic correlation with schizophrenia. BMI exhibited 78 regions of Bonferroni-significant local genetic correlation with schizophrenia, the most of any trait. All traits showed evidence of opposing mechanisms with schizophrenia (Table 22). See Figure 11 for Manhattan Plots of locus-level correlation between schizophrenia, cardiometabolic and inflammatory traits. See Appendix C Table 2 for the full numerical results from locus-level correlation analysis for all trait pairs.

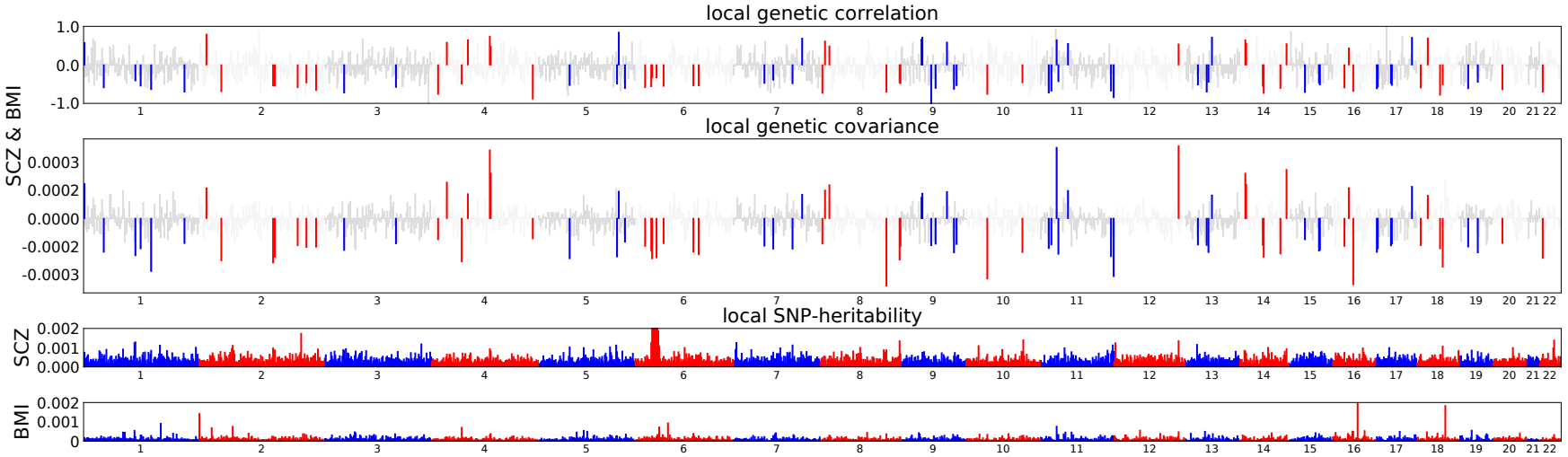
Table 22: Summary of Local Genetic Correlation Analyses between Schizophrenia, Cardiometabolic and Inflammatory Traits

Trait	LD Blocks, No.	Bonferroni *p*-value Threshold	Regions of local correlation[a] with schizophrenia, No.
BMI	1,684	2.70×10^{-5}	78
Fasting Insulin	1,676	2.98×10^{-5}	30
T2D	1,591	3.14×10^{-5}	8
CRP	1,684	2.70×10^{-5}	5
Triglycerides	1,684	2.70×10^{-5}	5
HDL	1,684	2.70×10^{-5}	4
Coronary Artery Disease	1,676	2.98×10^{-5}	4
LDL	1,684	2.70×10^{-5}	2

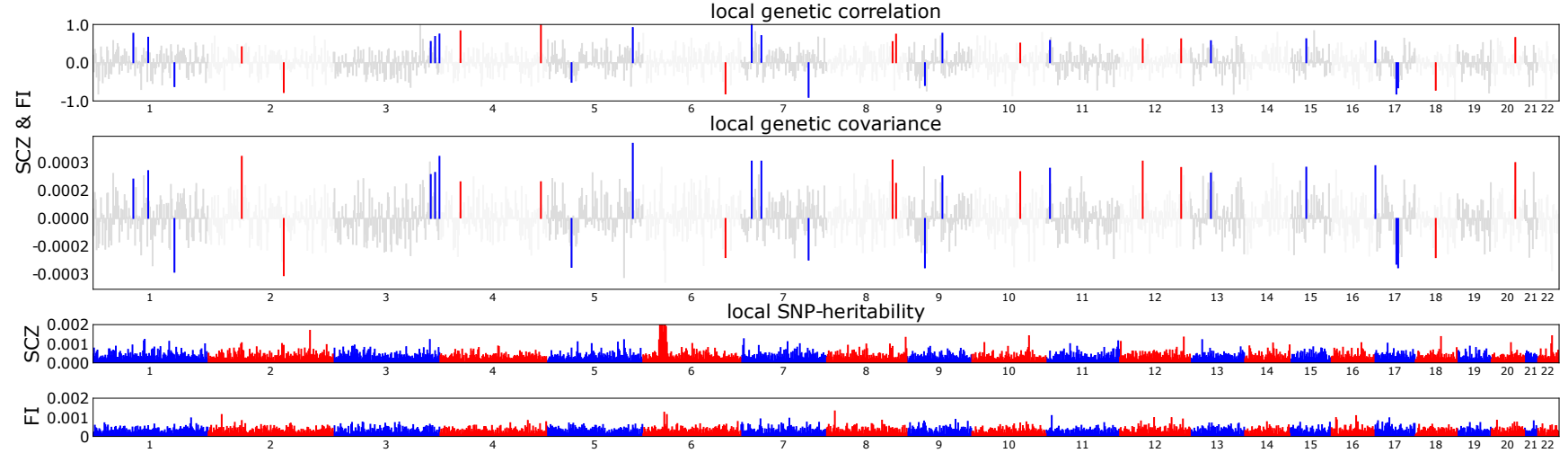
BMI = body mass index; T2D = type 2 diabetes mellitus; CRP = C-reactive protein; HDL = high-density lipoprotein; LDL = low-density lipoprotein; [a]regions with evidence of local correlation surpassing Bonferroni significance threshold

Figure 11: Manhattan Plots Showing Regions of Local Genetic Correlation between Schizophrenia and Cardiometabolic/Inflammatory Traits

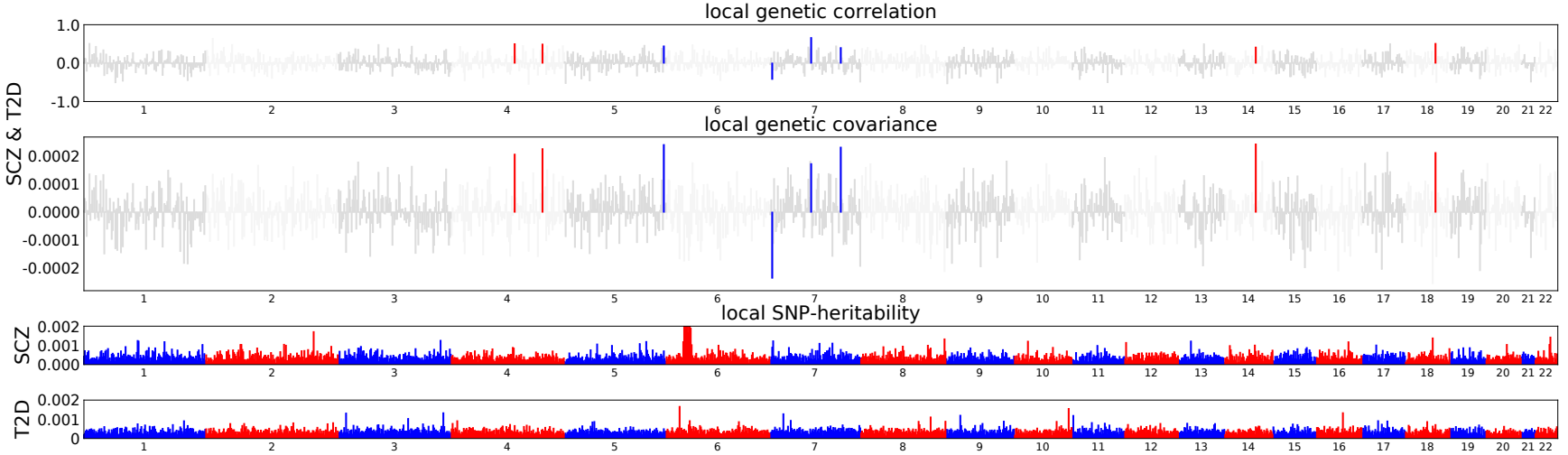
A. Body Mass Index



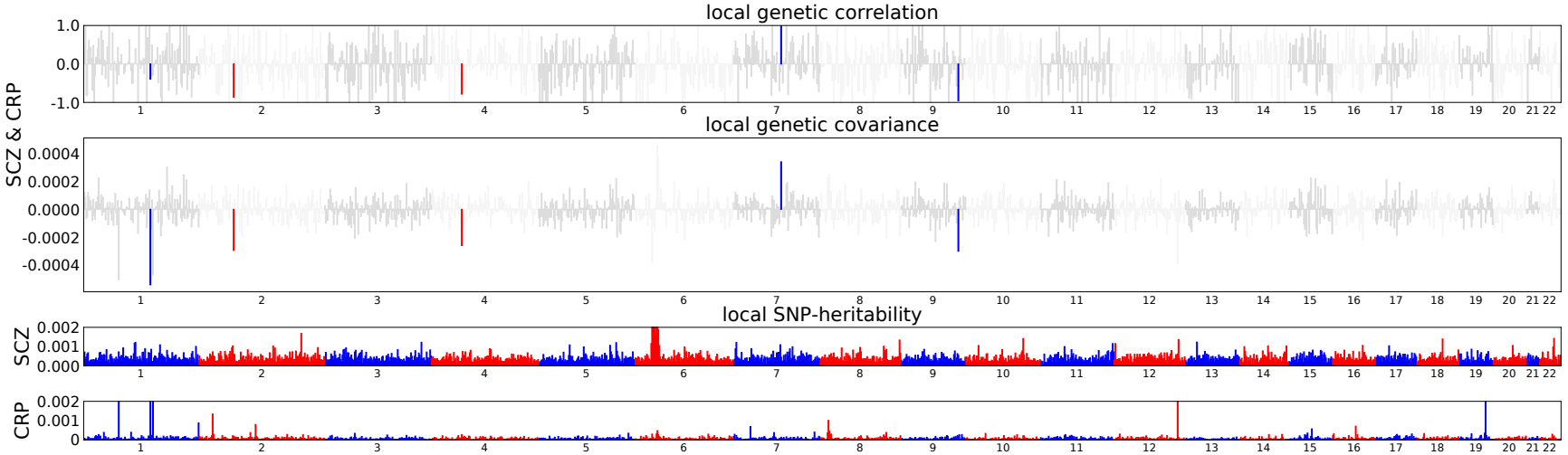
B. Fasting Insulin



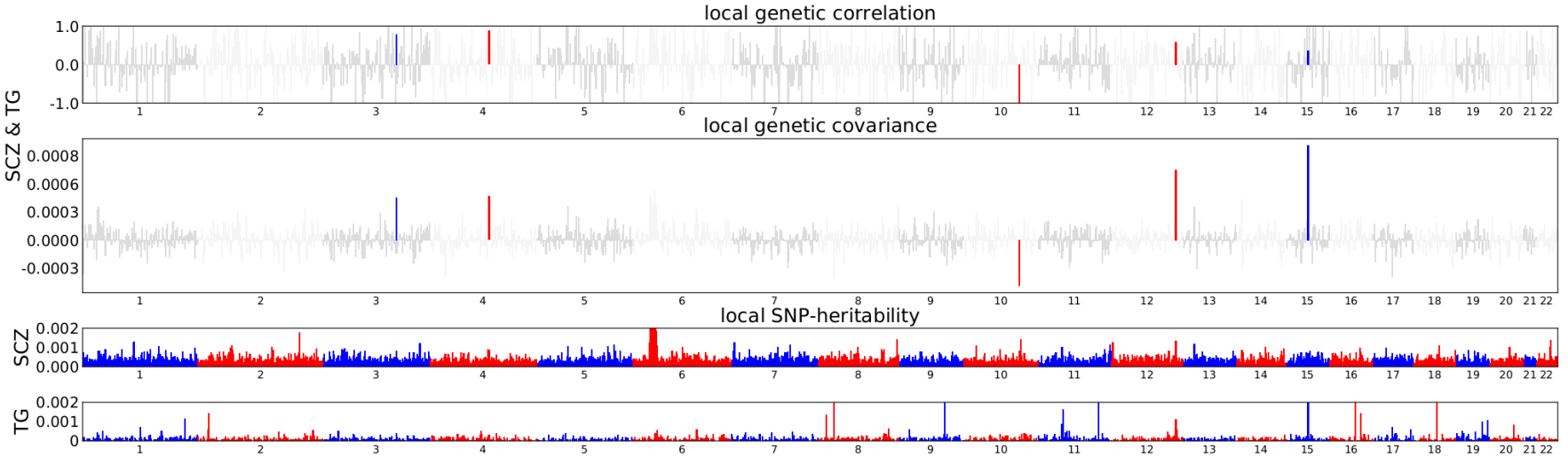
C. Type 2 Diabetes



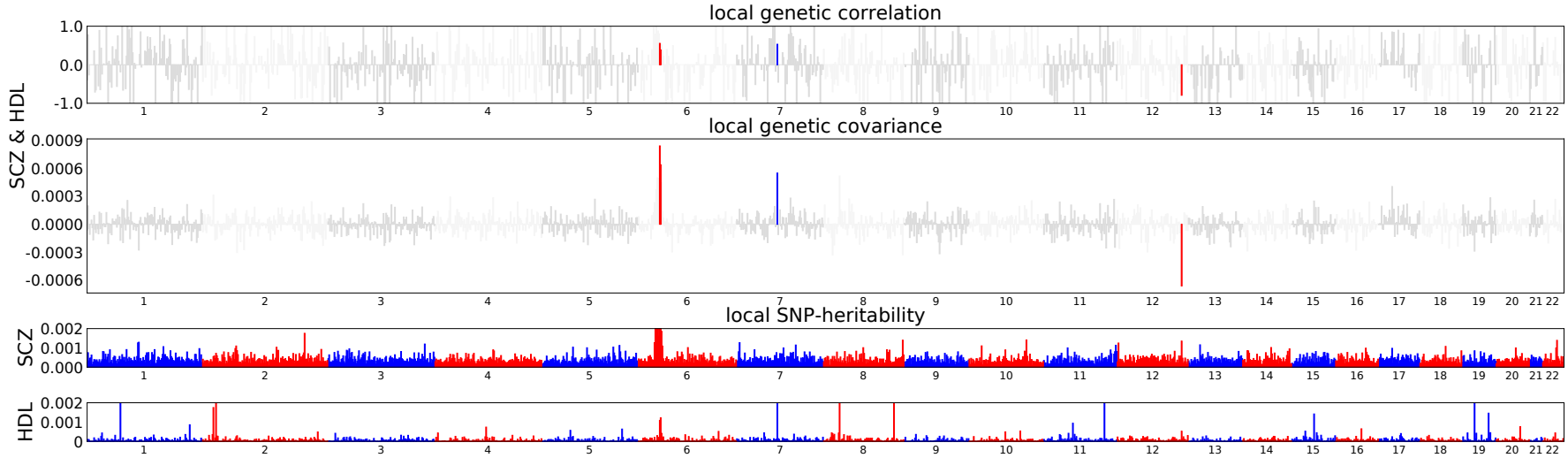
D. C-Reactive Protein



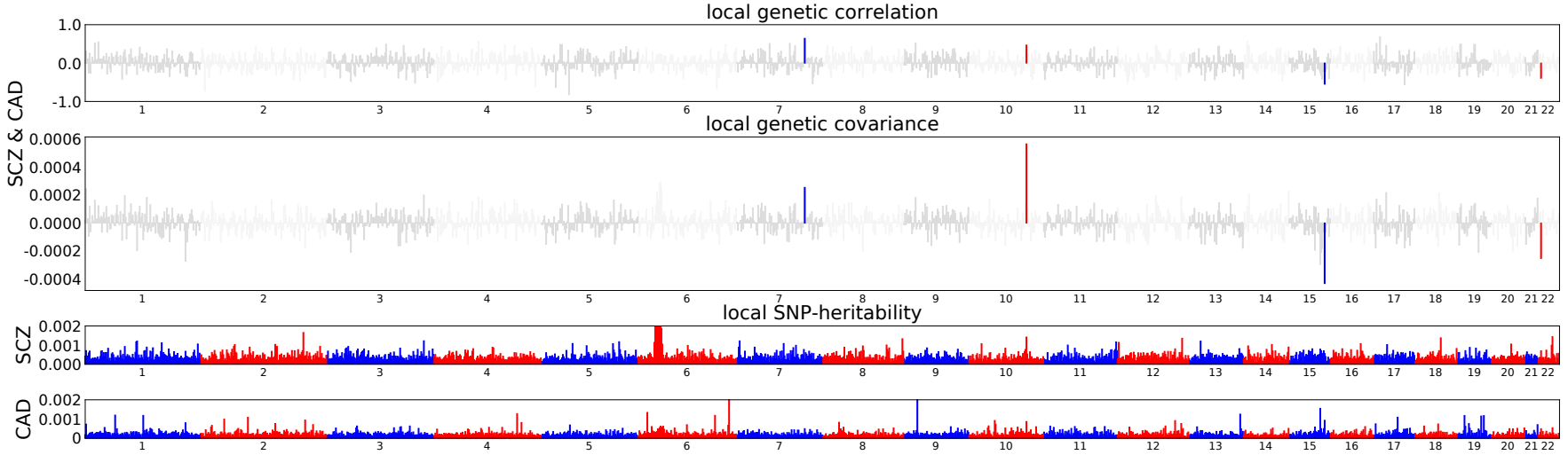
E. Triglycerides



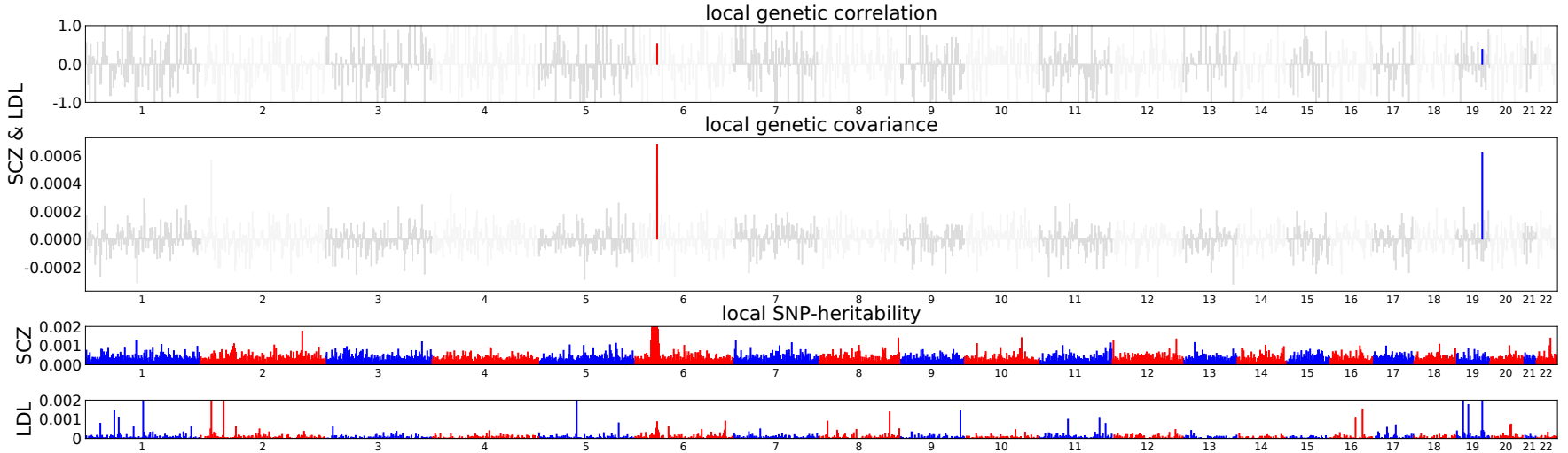
F. High-Density Lipoprotein



G. Coronary Artery Disease



H. Low-Density Lipoprotein



Manhattan plots showing local genetic correlation estimates (top panel); local covariance estimates (second panel); and local SNP heritability estimates (bottom two panels), at LD blocks across chromosomes 1-22. Areas coloured red/blue in top two panels correspond to LD-blocks surpassing Bonferroni significance threshold.

4.4.4 Multi-Trait Colocalization between Schizophrenia and Cardiometabolic / Inflammatory Traits

I found the strongest evidence for colocalization (posterior probability for colocalization (PP_{coloc}) >0.80) between schizophrenia, cardiometabolic and inflammatory traits at seven loci, which included missense (rs13107325; rs6265), intronic (rs17514846; rs8192675; rs3800229) and synonymous (rs3814883) variants, and one intergenic variant (rs12782894) (Table 23). See Figure 12 for stacked regional association plots for four colocalized variants with strong evidence for colocalization (rs8192675/*SLC2A2*, rs13107325/*SLC39A8*, rs6265/*BDNF*, rs17514846/*FURIN*). I found additional evidence for colocalization (PP_{coloc}=0.54-0.79) at five loci, including intronic (rs11191514; rs2108349; rs6031855; rs340874) and synonymous (rs2239647) variants (Table 23). See Appendix C Figure 2 for stacked regional association plots of the remaining variants with evidence for colocalization.

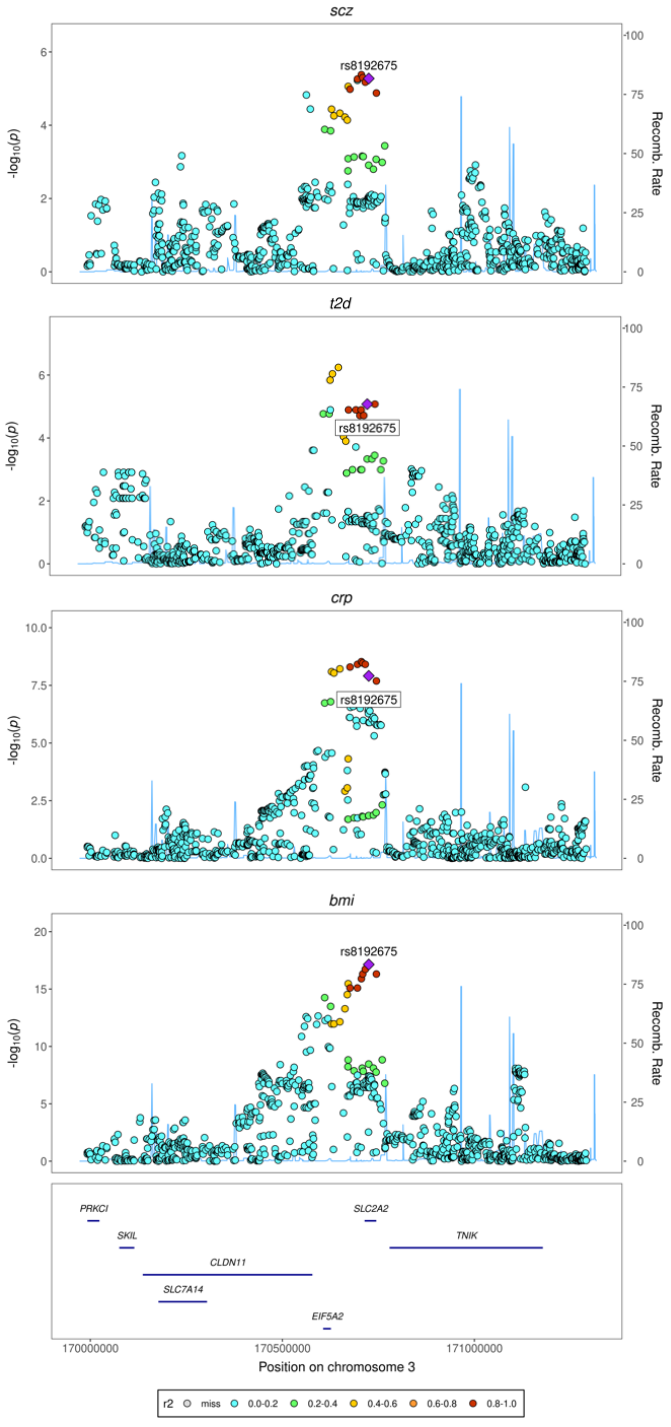
Table 23: Results from Colocalization Analysis between Schizophrenia, Cardiometabolic and Inflammatory Traits

Candidate SNP	Gene Implicated	Variant Type	Colocalized Traits	PP_{coloc}[a]	$PP_{explained}$[b]	N SNPs[c]
rs17514846	*FURIN*	Intron	SCZ, CAD	1.00	1.00	1071
rs3814883	*TAOK2*	Synonymous	SCZ, BMI	0.99	0.99	193
rs8192675	*SLC2A2*	Intron	SCZ, BMI, CRP, T2D	0.93	0.50	919
rs3800229	*FOXO3*	Intron	SCZ, BMI	0.89	0.96	872
rs12782894	*	*	SCZ, BMI	0.88	0.68	1255
rs13107325	*SLC39A8*	Missense	SCZ, HDL, TG, BMI, T2D	0.86	1.00	936
rs6265	*BDNF*	Missense	SCZ, BMI, CRP, CAD	0.86	0.75	925
rs2239647	*AKAP6*	Synonymous	SCZ, BMI, T2D	0.79	0.66	1584
rs11191514	*CNNM2*	Intron	SCZ, BMI, CAD	0.77	0.30	710
rs2108349	*GRB10*	Intron	SCZ, FI	0.60	0.88	1272
rs6031855	*YWHAB*	Intron	SCZ, BMI	0.59	0.28	990
rs340874	*PROX1*	Intron	SCZ, T2D	0.54	0.66	1324

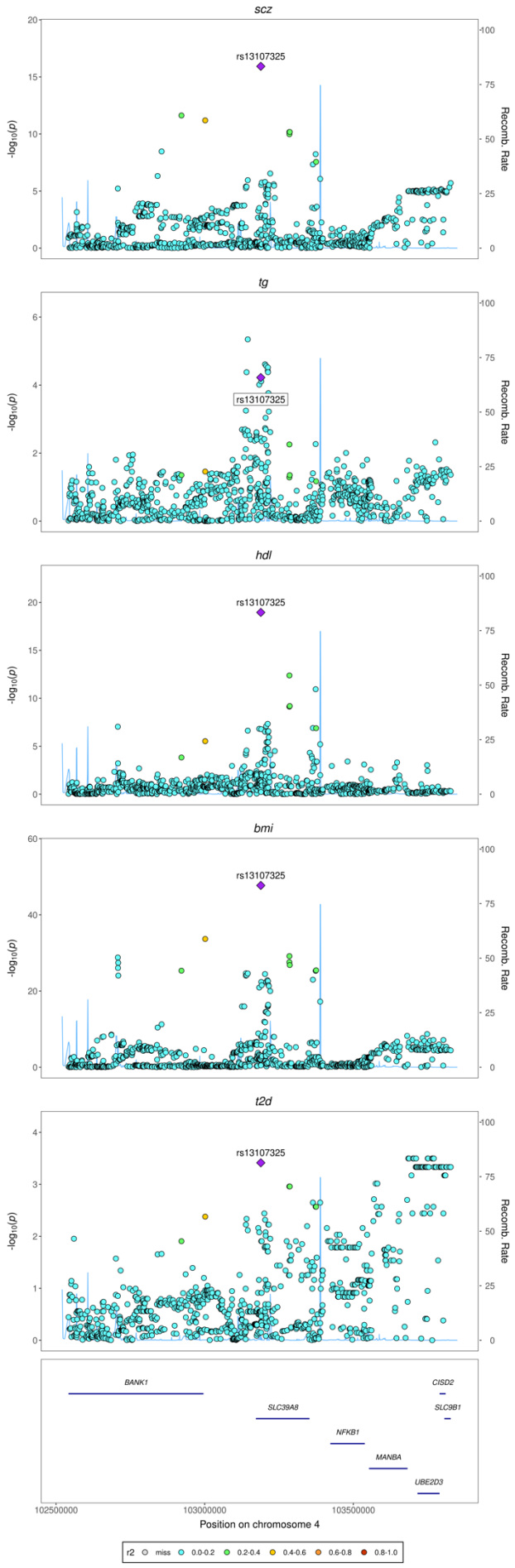
SCZ=schizophrenia; BMI=body mass index; CAD=coronary artery disease; HDL=high-density lipoprotein; TG=triglycerides; LDL=low-density lipoprotein; T2D=type 2 diabetes; CRP=C-reactive protein; FI=fasting insulin
[a]PP_{coloc} indicates posterior probability of single shared causal SNP at default prior and threshold settings
[b]$PP_{explained}$ indicates the amount of shared trait variance explained by the candidate SNP
[c]Corresponds to the number of SNPs present in all datasets; *Intergenic

Figure 12: Examples of Regional Genetic Association Plots for Four Loci Returning Strong Evidence for Colocalization between Schizophrenia, Cardiometabolic and Inflammatory Traits

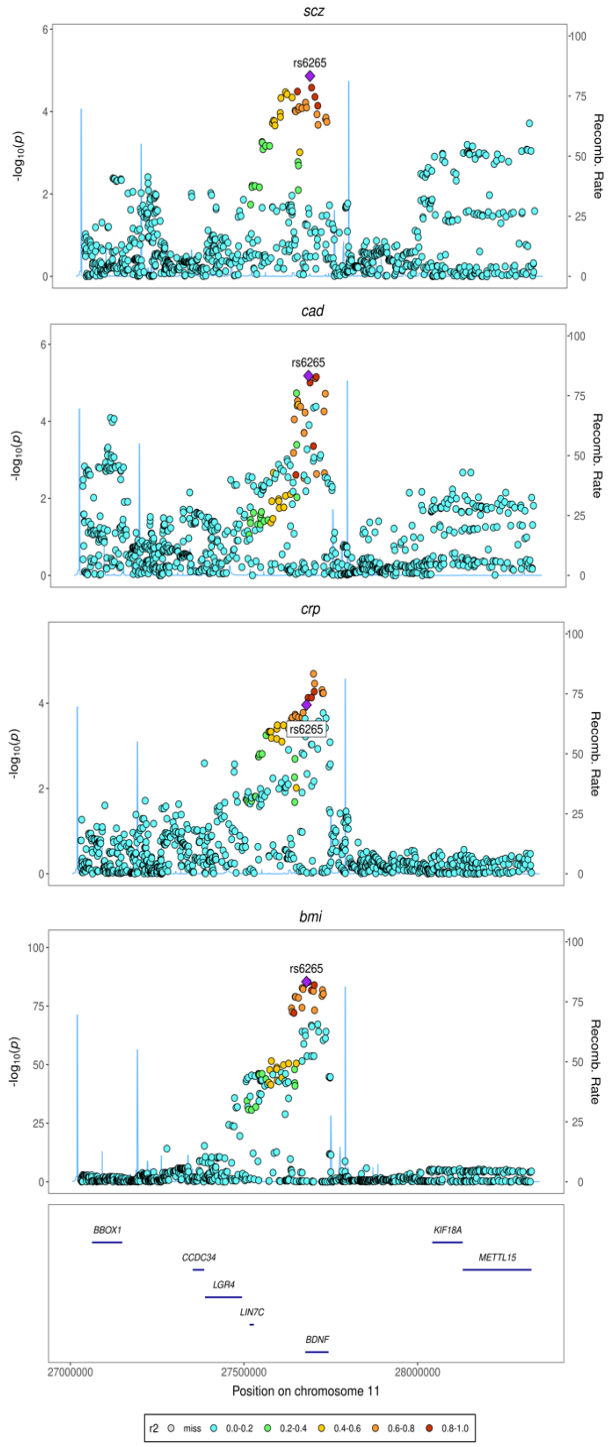
A. rs8192675 – *SLC2A2*



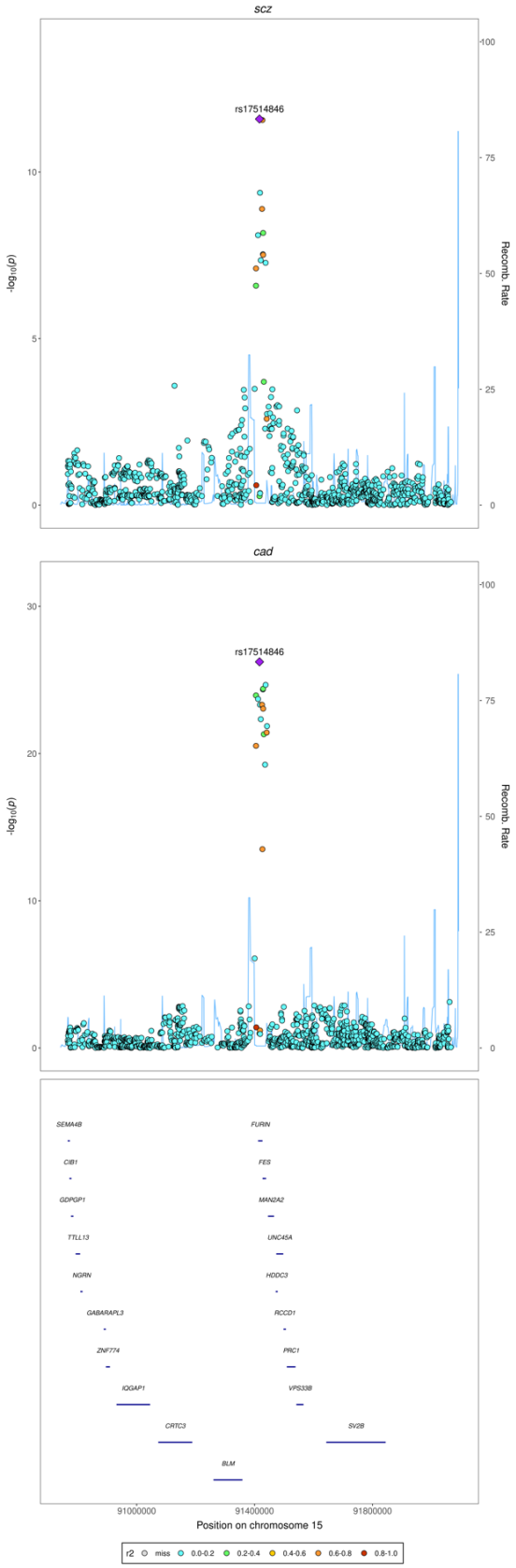
B. rs13107325 – *SLC39A8*



C. rs6265 – *BDNF*



D. rs17514846 - *FURIN*



Regional association plots denote chromosomal location (x axis) and strength of association with listed trait ($-\log_{10(p)}$) (y axis). SNP r^2 was estimated from the EPIC-Norfolk cohort. See Appendix C Figure 2 for regional association plots of the remaining colocalized variants described in Table 23. scz=schizophrenia; bmi=body mass index; tg=triglycerides; hdl=high-density lipoprotein; t2ds=type 2 diabetes; crp=c-reactive protein

4.4.5 Colocalization Sensitivity Analysis

Trait clusters for all loci were stable in sensitivity analysis, returning in all instances the same candidate colocalized variant over increasingly stringent prior and threshold configurations. See Appendix C Table 3 and Appendix C Figure 3 for full sensitivity analysis results and heatmap sensitivity analysis plots. To summarise the sensitivity analysis results, clusters at rs17514846 and rs3814883 were stable across all permutations of priors. Clusters at two loci (rs12782894; rs3800229) were stable till prior settings surpassed the most stringent level of 0.99. Clusters at three loci (rs8192675; rs13107325; rs2239647) were stable till regional/alignment thresholds surpassed a stringent level of 0.8, and then T2D was dropped from the clusters and the PP_{coloc} increased for the remaining traits. Clusters at rs6265 were stable till regional/alignment thresholds surpassed 0.7, then CRP was dropped and the PP_{coloc} increased for the remaining traits. Clusters at rs2108439 were stable till regional/alignment thresholds surpassed 0.6. Clusters at the remaining three variants (rs340874; rs11191514; rs6031855) were stable only at the recommended prior settings and regional/alignment thresholds.

4.5 Discussion

Using a complementary set of approaches leveraging GWAS summary data, I tested whether schizophrenia, cardiometabolic and inflammatory traits may share common genetic aetiology. First, I report evidence for partial genome-wide genetic correlation of schizophrenia with T2D and BMI. Second, I report that a 'cardiometabolic risk increasing' pattern of partial genetic correlation between schizophrenia, cardiometabolic and inflammatory traits may be confined to relatively lower-frequency genetic variants. Yet, a 'cardiometabolic risk lowering' pattern of partial genetic correlation may be present amongst the highest-frequency common genetic variants. I identified numerous regions of Bonferroni-significant locus-level genetic correlation between schizophrenia, cardiometabolic and inflammatory traits, which I interrogated using colocalization analysis. In doing so, I found robust and biologically plausible evidence for 12 colocalized SNPs that may at least in part contribute toward the comorbidity between schizophrenia, inflammation and cardiometabolic disorders. Together, the results suggest that the comorbidity between schizophrenia, inflammation and cardiometabolic disorders could be partly attributable to shared genes rather than being fully explained by lifestyle factors and medication side-effects.

Findings from the LDSC analysis are in line with previous research. For example, a similar negative correlation between schizophrenia and BMI was recently reported (Bahrami et al., 2020), and a large

genotyping meta-analysis has shown an inverse association between polygenic risk for schizophrenia and obesity (Zheutlin et al., 2019). Additionally, observational evidence indicates that low birthweight (Wahlbeck et al., 2001, Abel et al., 2010, Nielsen et al., 2013) and thinness in childhood (Zammit et al., 2007, Weiser et al., 2004, Sorensen et al., 2006) are associated with a higher risk for developing schizophrenia in adulthood. However, the LDSC SNP-heritability estimate suggests that only a modest fraction of phenotypic variance could be explained by the additive effects of shared genetic variants. This can be interpreted in one of two ways; either shared genetic architecture is only likely to explain a small fraction of the variance of phenotypic comorbidity, or LDSC estimates have been downwardly biased due to the limitations described in Section 4.2.

I also found a weak Bonferroni-significant overall negative genetic correlation between schizophrenia and T2D, which was not found in a previous LDSC study that used data from a smaller T2D GWAS (Bulik-Sullivan et al., 2015a). This finding is inconsistent with observational studies suggesting increased T2D risk in psychosis (Ward and Druss, 2015, Perry et al., 2016). The observed partial negative correlation between schizophrenia and BMI could explain this since T2D and BMI are highly genetically correlated (Zhang et al., 2017). This finding could also highlight the importance of environmental influences upon cardiometabolic risk in schizophrenia, given the small amount of phenotypic explained variance from LDSC. Alternatively, the finding could be explained by both the missing heritability phenomenon and presence of opposing mechanisms, and findings from the MAF-stratified and locus-level correlation analyses support the relevance of the latter interpretation.

Stratifying LDSC by MAF helps address the limitation of missing heritability and suggests that similar risk-increasing genetic architecture between schizophrenia and cardiometabolic disorders is likely to be confined to relatively lower-frequency GWAS-detectable variants. I identified a consistent pattern at nominal significance of 'cardiometabolic risk-increasing' partial correlation in the lowest MAF-quartile of schizophrenia with fasting insulin, triglycerides, CAD, HDL, T2D and CRP, and in the second-lowest MAF quartile of schizophrenia with LDL, which aligns with observational findings (Vancampfort et al., 2015, Miller et al., 2014). These findings also align with previous GWAS research, which leveraged pleiotropy with cardiovascular traits to improve detection of schizophrenia risk variants to reduce the impact of missing heritability (Andreassen et al., 2013). However, I also identified a pattern of 'cardioprotective' partial correlation in the highest MAF-quartile of schizophrenia with BMI and T2D, in line with whole-genome correlation estimates. These results suggest the presence of opposing mechanisms, which may be related to the heterogeneity of schizophrenia. Nevertheless, the Bonferroni significance threshold was not met for most traits in MAF stratified analysis. So, future replication of my work with better-powered GWAS is necessary to confirm these findings.

I found numerous regions of Bonferroni-significant locus-level correlation between included trait-pairs. Across all trait pairs, I found evidence of multiple regions of positive and negative correlation with schizophrenia, indicating opposing mechanisms. This may help explain the weaker evidence found in LDSC and GNOVA and the results of previous LDSC research (Bulik-Sullivan et al., 2015a), which found limited evidence for genetic correlation between schizophrenia and cardiometabolic traits. This is because the combination of regions of positive and negative correlation may have biased estimates toward the null (Shi et al., 2017).

I found twelve loci indicating evidence of colocalization between traits at the default prior configuration. Many of these were stable over increasingly stringent settings in sensitivity analysis, suggesting robust evidence for colocalisation. Several loci exhibited stronger evidence for colocalization after one weaker trait was dropped at more stringent thresholds. Of the seven loci returning the strongest evidence of colocalization (PP_{coloc}>0.80), four (rs6265; rs8192675; rs3800229; rs17514846) relate to pathways involving BDNF. BDNF is an important member of the neurotrophin family and is associated with a range of clinical features of schizophrenia (Notaras et al., 2015); is involved in the regulation of cardiometabolic function (Tasci et al., 2012); and is associated with cardiometabolic function in schizophrenia (Nurjono et al., 2014).

First, rs6265 (Val66Met) is a missense SNP in the *BDNF* gene. Val66Met reduces intracellular trafficking and activity-dependent secretion of BDNF (Egan et al., 2003). Interestingly, meta-analytic evidence suggests lower BDNF levels in people with schizophrenia (Cui et al., 2012), which may contribute to disease-specific changes of neuronal synaptic plasticity and the immune system (Zakharyan and Boyajyan, 2014). The Val66Met polymorphism may additionally influence food intake and body weight (Hong et al., 2012) in humans.

Second, rs8192675 is located in an intronic region of *SLC2A2*, which encodes the facilitated glucose transporter GLUT2. GLUT2 regulates the entry of glucose into the pancreatic β-cell, thus initiating the cascade of events leading to insulin secretion. GLUT2 is also highly expressed in both the liver, where it regulates both glucose uptake and output and the hypothalamus, where it regulates synaptic activity and neurotransmitter release (Jurcovicova, 2014). Variants in *SLC2A2* impair GLUT2 expression and are strongly associated with T2D (Sansbury et al., 2012). Rs8192675 is associated with increased diabetic symptomatology but may also be related to favourable T2D treatment response (Rathmann et al., 2019). Impaired GLUT2 expression is associated with lower levels of BDNF (Maekawa et al., 2013), and conversely, higher levels of BDNF are associated with a protective effect on GLUT2 in pancreatic β-cells, reducing T2D risk (Bathina and Das, 2019).

Third, rs3800229 lies in an intron of *FOXO3,* which regulates diverse cellular processes, for example, adult stem cell homeostasis (Eijkelenboom and Burgering, 2013) and immuno-metabolic processes (Lundell et al., 2019). *FOXO3* is associated with brain development and intracranial volume (Renault et al., 2009) and is associated with poor cognition in schizophrenia (Smeland et al., 2017). Interestingly, *FOXO3* is implicated as a potential therapeutic target for obesity (Deng et al., 2018) and mediates the inhibitory actions of insulin in diverse pathways, including cell metabolism and survival (Lee and Dong, 2017). *FOXO3* signalling can be disrupted by BDNF, mediated by the phosphatidylinositol 3-kinase (PI3K)/Akt pathway (Zhu et al., 2004). The PI3K/AKT pathway has roles in insulin sensitivity, neuronal development, dopamine regulation, and the immune system (Hers et al., 2011) and is implicated as a putative mechanism linking schizophrenia and T2D (Liu et al., 2013).

Fourth, the rs17514846 variant lies in an intron of *FURIN*, which encodes a protease that processes latent precursor proteins into their biologically active products. *FURIN* is expressed in neuroendocrine, liver, gut, and brain tissues. A recent GWAS found a significant association between rs17514846 and CAD (Webb et al., 2017), and rs17514846 regulates *FURIN* expression in monocytes, which modulates their migration and proliferation in atherosclerotic plaques (Turpeinen et al., 2011). Furthermore, rs17514846 is in high-LD with rs4702, a genome-wide significant variant for schizophrenia (Schizophrenia Working Group of the Psychiatric Genomics, 2014) which lies in the 3' untranslated region of *FURIN,* leading to reduced gene expression and impaired BDNF secretion (Hou et al., 2018).

Outside of BDNF-related pathways, there is biological plausibility for additional colocalized variants. One of these is rs13107325, a missense SNP in *SLC39A8,* which encodes a protein responsible for metal ion transport and homeostasis. Rs13107325 has been associated with weight gain (Pulit et al., 2019), lipid dysfunction (Willer et al., 2013), changes in brain volume (Luo et al., 2019) and brain metal homeostasis, the latter of which may influence schizophrenia risk (Carrera et al., 2012). Rs340874 is a genome-wide significant variant for T2D (Mahajan et al., 2014) and lies in an intron of *PROX1*. *PROX1* has been proposed as a possible genetic mechanism for comorbid schizophrenia and T2D (Hackinger et al., 2018) and is associated with pancreatic beta-cell development and neurogenesis (Holzmann et al., 2015). Rs2108349 lies in an intron of *GRB10*, which encodes an inhibitor of insulin receptor signalling (Morrione, 2000). The variant is in high-LD with rs2237457, which is associated with schizophrenia treatment resistance (Li and Meltzer, 2014). Finally, two variants, rs3814883 in *TAOK2* and rs11191514 in *CNNM2,* are each associated with schizophrenia (Guan et al., 2016, Li et al., 2017b), and both are associated with increased risks of cardiometabolic and cardiovascular disorders (Lv et al., 2017, Zhang et al., 2017).

The main strengths of this study include the use of several complementary genomic analysis methods that refine genetic correlation estimates to putative common-causal SNPs. This can inform future basic research and highlight potential pathways that might be investigated for therapeutic potential for schizophrenia and its associated cardiometabolic comorbidity. The findings represent a consistent pattern of evidence across complementary methods, which can address the limitations of previous research.

The main limitations of this study are as follows: GWAS power might have affected the results. For example, some correlation estimates for MAF-stratified analysis did not reach the Bonferroni-corrected evidential threshold. Therefore, the results should be interpreted with caution and require replication when better powered GWAS are available. I considered traits for further analysis based upon a nominal threshold since correlation estimates, which are either (i) averaged across the whole genome (LDSC); or (ii) averaged across MAF quartiles (GNOVA), may have been biased toward the null where opposing mechanisms exist, and results from the locus-level correlation analyses suggested this was the case for all analysed trait pairs. Differences in GWAS statistical power between traits may also be partly responsible for the different numbers of regions of local correlation identified in the HESS analyses. In future, better-powered GWAS may identify more regions of locus-level correlation between trait pairs. For the MAF-stratified analysis, due to limitations in current GWAS power, I could only include SNPs with MAF >5% in the lowest MAF-quartile, a limitation common to genetic correlation methods. Such variants are therefore best described as a lower-frequency tranche of common genetic variation. As GWAS methods and sample sizes improve, sufficient power may be achieved to consider variants on the rare end of the MAF spectrum.

Secondly, HyPrColoc assumes the presence of at most one causal SNP in the region, a limitation common to colocalisation methods. Yet, HyPrColoc estimates may only become unreliable when the secondary causal variants explain a similar amount of trait variation as the primary shared variant (Foley, 2019).

Thirdly, I could only include one inflammatory marker, CRP, since large-scale GWAS of other inflammatory biomarkers are scarce. Despite CRP being a generalized marker of inflammation, future replication of the work with a more extensive set of upstream inflammatory markers may help test specific inflammatory pathways. Future research may also consider other mental disorders, for example, depression, which is genetically correlated with schizophrenia (Anttila et al., 2018) and is also observationally associated with cardiometabolic disorders (Lamers et al., 2018). Finally, some level of similarity in genetic architecture might be expected between any set of complex disease traits; however, the results of this study show a consistent pattern across a number of independent analytic methods, suggesting that chance associations are unlikely to fully explain the results.

In conclusion, I present evidence indicating a shared genetic basis for schizophrenia, cardiometabolic and inflammatory traits. The results suggest that the commonly observed comorbidity between these conditions may be at least partly heritable. The results indicate that the shared genetic aetiology may be confined to relatively lower-frequency common genetic variants. The majority of loci showing evidence for colocalization are biologically plausible, with several implicating pathways involved in regulating BDNF and glucose transport. Together, the results highlight putative pathophysiological mechanisms that could underly the comorbidity, which may form the basis for future basic and therapeutics research, both for schizophrenia and its associated cardiometabolic comorbidity.

Chapter 5

The Potential Shared Role of Inflammation in Insulin Resistance and Schizophrenia: A Bi-Directional Two-Sample Mendelian Randomization Study

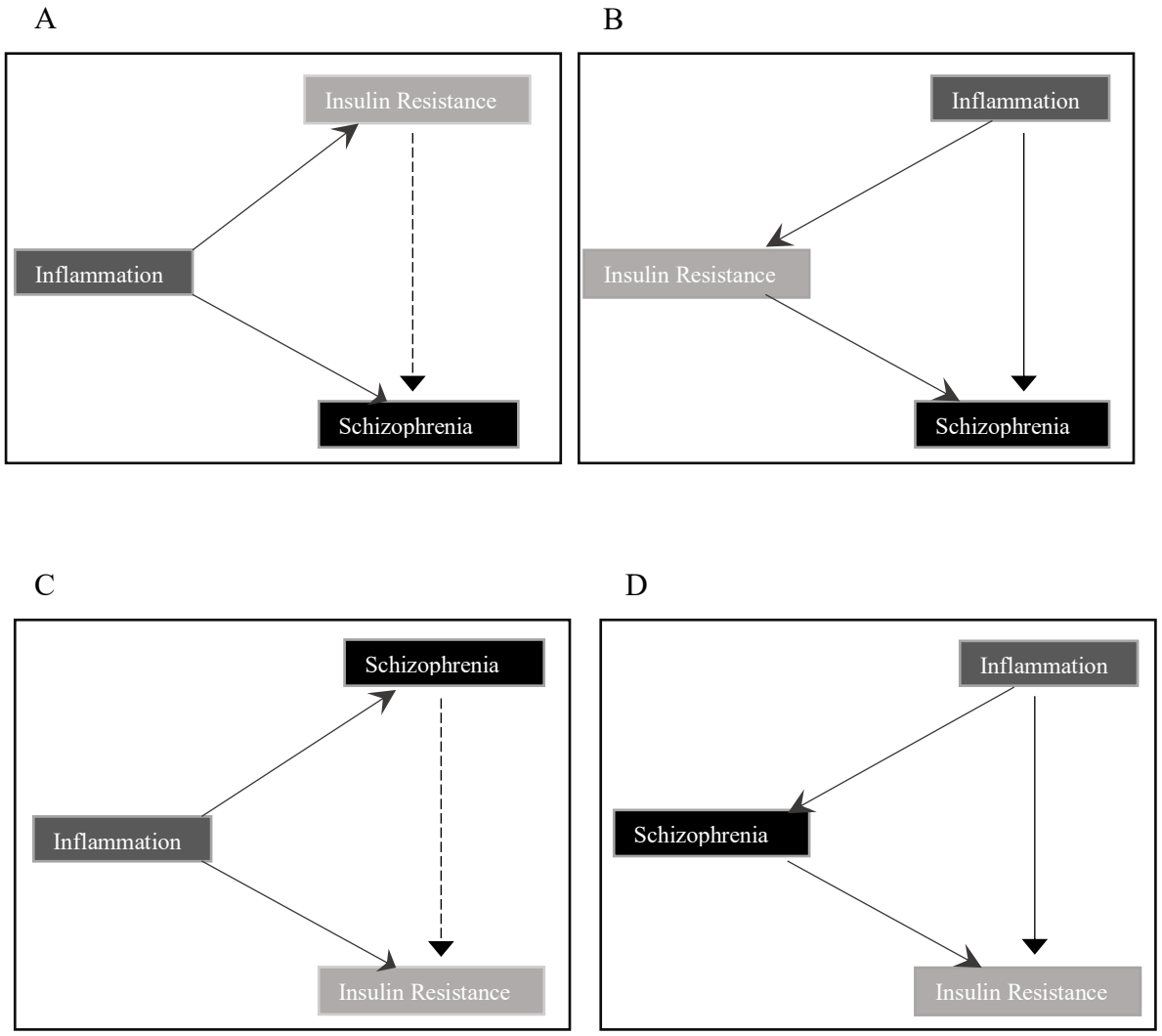
5.1 Introduction

Most existing research examining the cardiometabolic comorbidity of schizophrenia is cross-sectional. Therefore, existing studies cannot confirm whether cardiometabolic disorders are a cause or consequence of illness (i.e., reverse causality). Additionally, whilst previous studies have adjusted for potential confounders, residual confounding, which is a limitation of both cross-sectional and longitudinal research, could still be relevant. MR analysis can address these limitations by using genetic variants inherited randomly at conception as unconfounded proxies of a modifiable exposure to examine whether the exposure may have a causal effect on a disease outcome (Smith, 2010). MR studies of cardiometabolic traits and schizophrenia are scarce, have focused on a limited set of cardiometabolic exposures, and have reported mixed findings (Li et al., 2018, Polimanti et al., 2017). To the best of my knowledge, MR studies examining associations between a wide range of cardiometabolic traits and schizophrenia are lacking. Such studies may help identify common potentially causal risk factors and pathophysiologic mechanisms for these physical and psychiatric illnesses.

Inflammation could be pathophysiologically related to cardiometabolic disorders and schizophrenia. Higher levels of circulating inflammatory markers have been associated with both psychosis and cardiometabolic disorders, both cross-sectionally and longitudinally (Dandona et al., 2004, Khandaker et al., 2014, Upthegrove et al., 2014). MR studies have reported potential causal associations between inflammation, particularly CRP and IL-6, and schizophrenia (Hartwig et al., 2017, Khandaker et al., 2017). CRP and IL-6 are also implicated in the pathogenesis of insulin resistance (Kim et al., 2009) and may exaggerate the effects of insulin resistance on psychosis-risk in young adults (Perry et al., 2018). However, to the best of my knowledge, no MR studies have examined whether inflammation could be pathophysiologically related to insulin resistance and schizophrenia, for example, via mediating or common-causal mechanisms.

Therefore, I have conducted a study to examine evidence in support of four scenarios regarding the potential relationships between inflammation, insulin resistance and schizophrenia: a) inflammation is a common cause (confounder) between insulin resistance and schizophrenia; b) insulin resistance mediates an association between inflammation and schizophrenia; c) inflammation is a common cause (confounder) between schizophrenia and insulin resistance; d) schizophrenia mediates an association between inflammation and insulin resistance. See Figure 13 for directed acyclic graphs (DAGs) illustrating the proposed mechanisms.

Figure 13: Directed Acyclic Graphs Outlining Potential Mechanisms of Association between Inflammation, Insulin Resistance and Schizophrenia



5.2 Aims and Objectives

First, I carried out MR analyses to test whether ten cardiometabolic traits related to insulin resistance (fasting insulin; triglycerides, HDL; LDL; FPG; BMI; glucose tolerance; leptin, glycated HbA1C; T2D) could be causally associated with schizophrenia. To test the direction of association, I used genetically predicted levels of cardiometabolic traits as exposures and schizophrenia as the outcome, and *vice versa*. Next, I examined whether inflammation could be a shared mechanism linking insulin resistance and schizophrenia by including genetic variants for each cardiometabolic trait that were also associated with a marker of inflammation. Finally, I used multi-variable MR (MVMR) analysis to control for genetic associations of cardiometabolic traits with CRP, an archetypal general inflammatory marker that I used as a general measure for systemic inflammation.

5.3 Methods

5.3.1 Selection of Genetic Variants Related to Cardiometabolic Traits and Schizophrenia

For fasting insulin, triglycerides, and HDL, I used a set of 53 SNPs reported to be associated with insulin resistance from a recent meta GWAS of 188,577 European adults which adjusted for BMI (Lotta et al., 2017a). In this study, I included SNPs reaching genome-wide significance for the corresponding trait. Summary statistics for genome-wide significant SNPs were also obtained for six related continuous (FPG, HbA1C, LDL, BMI, leptin, glucose tolerance) and one binary (T2D) cardiometabolic traits from recent large GWAS (Table 24). See Appendix C Tables 4-11 for the SNPs included for each exposure. I obtained summary statistics for schizophrenia from a recent GWAS from the PGC (Pardinas et al., 2018) based on 40,675 cases and 64,643 European controls. The degree of sample overlap between exposure and outcome samples was likely to be low since the data were obtained from different consortia (Shi et al., 2017). The study was a secondary analysis of the above publicly-available data. Informed consent was sought for all participants per the original GWAS protocols, and all ethical approvals for the GWAS were obtained by original GWAS authors.

Table 24: GWAS used for SNP Selection in MR Analysis

Cardiometabolic Exposure	Consortia	Ethnic Origin	Sample	Setting[a]	GWAS-significant SNPs, No.	Inflammation-related SNPs, No.[b]
Fasting Insulin (Lotta et al., 2017)	MAGIC	European	108,557 (101,393 adults and 7,164 adolescents)	Meta-GWAS of 19 European Cohort Studies, participants with diabetes excluded	53	5
Triglycerides (Lotta et al., 2017)	EPIC-InterAct, FPLD1	European	188,577 adults	Meta-GWAS of 45 population-based cohort studies and case-control studies.	53	5
HDL (Lotta et al., 2017)	EPIC-InterAct, FPLD1	European	188,577 adults	Meta-GWAS of 45 population-based cohort studies and case-control studies.	53	4
LDL (Willer et al., 2013)	GLGC	European, East Asian, South Asian, African	173,082 adults	Meta-GWAS of 45 population-based cohort studies and case-control studies.	79	13
Fasting Plasma Glucose (Manning et al., 2012)	MAGIC	European	58,074 adults	Meta-GWAS of 29 European population-based cohort studies. participants with diabetes excluded.	22	2
T2D (Mahajan et al., 2018)	DIAGRAM	European, East Asian, South Asian, Mexican, Mexican American	435,387 adults; (81,412 with T2DM and 370,832 controls)	Meta-GWAS of two large prospective European cohort studies	152	7
BMI (Locke et al., 2015)	GIANT	European, African, Asian	339,224 adults	Meta-GWAS of 125 European cohort studies, adjusted for age	97	6
HbA1C (Wheeler et al., 2017)	MAGIC	European, African American, East Asian, South Asian	159,940 adults	Meta-GWAS of 82 population-based cohort studies. Participants with diabetes excluded.	60	7
Glucose Tolerance (Saxena et al., 2010)	MAGIC	European	15,234 adults	Meta-GWAS of 9 population-based cohort studies. Participants with diabetes excluded.	7	0
Leptin (Kilpeläinen et al., 2016)	-	European	82,315 adults	Meta-GWAS of 32 population-based cohort studies adjusted for age and sex.	5	0

SNP=Single Nucleotide Polymorphism; GWAS=Genome-Wide Association Study; HDL=High-Density Lipoprotein; LDL=Low-Density Lipoprotein; T2D=Type 2 Diabetes Mellitus; BMI=Body Mass Index; HbA1C=Glycated Haemoglobin; MAGIC=Meta-Analysis of Glucose and Insulin Related traits Consortium; GLGC=Global Lipids Genetics Consortium; DIAGRAM=Diabetes Genetics Replication and Meta-Analysis; GIANT=Genetic Investigation of Anthropometric Traits

[a]See original GWAS publication for detailed demographic and setting information for studies included in meta-GWAS.

[b]Number of SNPs with pleiotropy for inflammation at genome-wide significance

5.3.2 Statistical Analysis

I obtained summary-level data (SNP rs number; β-coefficient or log OR; standard errors or 95% confidence intervals; effect allele; other allele; *p*-value; effect allele frequency; sample size; number of cases/controls) from each GWAS. Where a specific instrument SNP was not available in the outcome dataset, I located proxy SNPs using LD tagging (r^2>0.8) via *LDlink* (Machiela and Chanock, 2015). Alleles were harmonised based on matching alleles and the resulting instruments were clumped for LD to ensure independence (10,000kb pairs apart, r^2<0.001). In the event of palindromic SNPs, the forward strand was inferred where possible using allele frequency information. I performed bidirectional analysis (i.e., with schizophrenia as exposure and cardiometabolic traits as outcomes) to examine direction of association. Statistical analysis was conducted using the *TwoSampleMR* package (v0.5.4) (Hemani et al., 2018) for R (R Core Team, 2017).

The primary MR analysis method was inverse variance weighted (IVW) regression when at least two exposure SNPs were available for analysis. IVW consists of a weighted linear regression of SNP-exposure SNP-outcome effect estimates. The IVW estimate is the inverse variance weighted mean of ratio estimates from two or more instruments (Burgess et al., 2013), and assumes that all SNPs are valid instruments or that the sum of directional bias is zero. Since the intercept is an estimate of average pleiotropic effects across instrumental variables, in an IVW approach the intercept is fixed to 0. When one exposure SNP was available for analysis, I used the Wald ratio method.

I also conducted weighted median and MR-Egger regression analysis. The weighted median is the median of the weighted empirical distribution function of individual SNP ratio estimates. This method provides a consistent effect estimate if more than 50% of the information comes from valid SNPs (Bowden et al., 2016a). MR-Egger regression consists of a weighted linear regression similar to IVW, with the assumption that horizontal pleiotropic effects and SNP-exposure associations are uncorrelated (Bowden et al., 2015), therefore the intercept is not fixed. MR Egger regression provides a valid effect estimate even if all SNPs are invalid instruments but assumes that uncertainty in the SNP-exposure association estimates is negligible (Bowden et al., 2017).

For the binary outcome of schizophrenia, the estimates for continuous exposures (FI, HDL, triglycerides, LDL; FPG; BMI; HbA1C; glucose tolerance, leptin) represent log-odds ratios converted into ORs representing the increase in risk of schizophrenia per SD of exposure, and 95% CIs. For binary exposures (T2D), the estimates represent the OR for schizophrenia per unit increase in the log-odds of T2D. For continuous cardiometabolic outcomes, β-coefficients represent the SD increase in exposure per unit increase in the log-odds of schizophrenia, with SEs.

I performed several sensitivity analyses to check the validity of the results. Heterogeneity among SNPs included in each analysis was examined using the Cochran Q test. I checked for horizontal pleiotropy using the MR Egger regression intercept alongside a more recent and robust method to detect horizontal pleiotropy and outliers, 'MR pleiotropy residual sum and outlier' (MR-PRESSO) (Verbanck et al., 2018). MR-PRESSO relies on a regression framework where the variants' effects on the outcome are regressed on the same variants' effects on exposure, with the slope of the regression line providing an estimate of the causal effect of the exposure on the outcome (Verbanck et al., 2018). The MR-PRESSO global test evaluates overall horizontal pleiotropy amongst all instrumental variables in a single MR test by comparing the observed distance of all the variants to the regression line (residual sum of squares) to the expected distance under the null hypothesis of no horizontal pleiotropy (Verbanck et al., 2018). The MR-PRESSO outlier test evaluates the presence of specific horizontal pleiotropic outlier variants by using the observed and expected distributions of the tested variant. Finally, the MR-PRESSO distortion test evaluates the significance of the distortion between the causal estimate before and after removal of the horizontal pleiotropic outlier variants (detected from the outlier test of MR-PRESSO). Using MR-PRESSO, I used the global test to examine for horizontal pleiotropy, and where evident, used the method to correct the IVW-estimate via outlier removal.

I examined for measurement error in SNP-exposure associations using the I^2_{GX} statistic (Bowden et al., 2016b).

5.3.2.1 Analysis using Inflammation-Related SNPs

Next, I repeated MR analysis using only inflammation-related SNPs for each cardiometabolic risk factor as an instrumental variable for the outcome of schizophrenia. I did this to test the hypothesis that these SNPs may represent a biological mechanism involving inflammation. This could be via, for example, a common causal basis (Panel A in Figure 13) or via vertical (mediating) pleiotropy (Hemani et al., 2018) (Panel B in Figure 13). I used *Phenoscanner v2* (Staley et al., 2016) to examine each SNP associated with each cardiometabolic risk factor, to identify SNPs that were also associated with a measure of inflammation, defined as blood concentration/count of cytokines (such as chemokines, interferons, interleukins, lymphokines, or tumour necrosis factors), acute phase or inflammatory proteins (e.g., CRP), or immune cells (e.g., neutrophils, lymphocytes). Primarily, I considered inflammation-related SNPs at genome-wide significance ($p<5\times10^{-8}$) to maximise specificity. However, I also performed a sensitivity analysis by including inflammation-related SNPs at a less-stringent nominal significance threshold ($p<1\text{x}10^{-4}$) used previously to increase sensitivity

toward inflammation-related SNPs (Ligthart et al., 2015). See Appendix C Tables 12-19 for inflammation-related SNPs and associated inflammation-related pleiotropy. Using the same method, I identified genome-wide inflammation-related schizophrenia SNPs (Appendix C Table 20) and used them as instrumental variables in MR analysis examining cardiometabolic traits as outcomes.

5.3.2.2 Adjustment for Inflammation

As a sensitivity analysis to estimate whether any associations evident above may be explained by inflammation, I conducted MVMR analysis (Burgess and Thompson, 2015, Sanderson et al., 2018) using the genome-wide significant SNPs for fasting insulin, triglycerides and HDL, representative of an insulin resistance phenotype as exposures, with schizophrenia as the outcome, after conditioning on the associations of those SNPs with CRP. I chose CRP because it is a widely used downstream measure of systemic inflammation, and publicly available data from large-scale GWAS for CRP are available. Summary statistics for CRP were obtained from a recent large GWAS based on 204,402 participants (Ligthart et al., 2018). For CRP as an exposure in MVMR, I used independent SNPs reported to be conditionally associated with CRP and located within the *CRP* gene coding region. See Appendix C Table 21.

5.3.2.3 Correction for Multiple Testing

Statistical significance was estimated using the Holm-Bonferroni correction method (Holm, 1979), correcting for the number of exposures tested at each stage of analysis.

5.4 Results

5.4.1 MR Analyses using All Genetic Variants Associated with IR and Other Cardiometabolic Traits

There was no evidence for associations between genetically-predicted levels of cardiometabolic traits and schizophrenia, using the primary IVW analysis method. Evidence using the weighted median method for associations between genetically-predicted levels of triglycerides (weighted median OR=1.26; 95% C.I., 1.06-1.50; corrected p=0.090) and HDL (weighted median OR=0.79; 95% C.I., 0.65-0.95; corrected p=0.126) with schizophrenia did not survive correction for multiple testing (Table 25).

Table 25: MR Analyses of Cardiometabolic Traits and Schizophrenia using All SNPs

Risk Factor	SNPs, N[a]	Method	Odds Ratio (95% C.I.)	*p*-value	Corrected *p*-value[b]
Fasting Insulin	9	IVW	1.13 (0.76-1.70)	0.548	1.000
		Weighted Median	0.98 (0.68-1.41)	0.920	1.000
		MR Egger	9.24 (1.82-46.97)	0.028	0.280
Triglycerides	9	IVW	1.16 (0.86-1.56)	0.334	1.000
		Weighted Median	1.26 (1.06-1.50)	0.009	0.090
		MR Egger	1.31 (0.84-2.03)	0.308	1.000
HDL	14	IVW	0.94 (0.71-1.23)	0.649	1.000
		Weighted Median	0.79 (0.65-0.95)	0.010	0.126
		MR Egger	0.67 (0.45-0.99)	0.067	0.670
Fasting Plasma Glucose	18	IVW	1.07 (0.87-1.31)	0.522	1.000
		Weighted Median	1.01 (0.84-1.23)	0.887	1.000
		MR Egger	1.13 (0.74-1.74)	0.584	1.000
Type 2 Diabetes	27	IVW	0.93 (0.78-1.12)	0.470	1.000
		Weighted Median	0.93 (0.80-1.09)	0.375	1.000
		MR Egger	1.03 (0.66-1.62)	0.895	1.000
Body Mass Index	81	IVW	1.05 (0.89-1.24)	0.554	1.000
		Weighted Median	1.07 (0.92-1.24)	0.383	1.000
		MR Egger	1.43 (0.97-2.10)	0.103	1.000
HbA1C	36	IVW	1.01 (0.76-1.32)	0.956	1.000
		Weighted Median	1.12 (0.82-1.51)	0.483	1.000
		MR Egger	1.33 (0.79-2.23)	0.295	1.000
Glucose Tolerance	7	IVW	0.98 (0.85-1.14)	0.800	1.000
		Weighted Median	1.10 (0.87-1.15)	0.993	1.000
		MR Egger	1.85 (0.95-3.32)	0.094	0.940
LDL	74	IVW	0.99 (0.93-1.05)	0.679	1.000
		Weighted Median	0.97 (0.90-1.03)	0.322	1.000
		MR Egger	0.98 (0.90-1.07)	0.692	1.000
Leptin	4	IVW	1.97 (0.90-4.31)	0.091	0.910
		Weighted Median	1.18 (0.66-2.11)	0.579	1.000
		MR Egger	3.29 (0.56-17.22)	0.358	1.000

HDL=high-density lipoprotein; HbA1C=glycated haemoglobin; LDL=low-density lipoprotein; IVW=inverse variance weighted regression; SNPs=single nucleotide polymorphisms; [a]Number of SNPs remaining after clumping for independence; [b] Each analysis method (IVW, Weighted Median and MR Egger) corrected using the Holm-Bonferroni method for 10 cardiometabolic markers
Estimates represent ORs for schizophrenia per SD increase in exposure (per unit-increase in log-odds of exposure for T2DM)

5.4.2 MR Analyses using Inflammation-Related Genetic Variants for IR and Other Cardiometabolic Traits

After testing only genome-wide significant inflammation-related variants for cardiometabolic traits, I found evidence for associations of inflammation-related genetically-predicted fasting insulin (Wald Ratio OR=2.95; 95% C.I., 1.38-6.34; corrected p=0.035) and HDL (Wald Ratio OR=0.55; 95% CI, 0.36-0.84; corrected p=0.035) with schizophrenia. I could not include any genome-wide significant inflammation-related variants for triglycerides, leptin, or glucose tolerance. In the sensitivity analysis featuring inflammatory-related cardiometabolic variants at a less stringent significance threshold,

evidence persisted for associations of inflammation-related genetically-predicted fasting insulin (IVW OR=1.74; 95% C.I., 1.08-2.98; corrected p=0.030) and HDL (IVW OR=0.78; 95% C.I., 0.62-0.92; corrected p=0.036) with schizophrenia. In addition, there was evidence for an association of genetically-predicted inflammation-related triglycerides (IVW OR=1.24; 95% C.I., 1.07-1.55; corrected p=0.036) with schizophrenia (Table 26; Figures 14-15).

5.4.3 Adjustment for Inflammation

MVMR analysis for inflammation-related SNPs of fasting insulin, triglycerides and HDL with schizophrenia showed that the univariable associations fully attenuated after controlling for the genetic associations of these variants with CRP, in analyses involving both inflammation-related SNPs at genome-wide and nominal significance levels. Controlling for CRP had negligible effect on MR estimates based on all genetic variants (Tables 27-28; Figure 16).

5.4.4 Test for Bidirectionality using Schizophrenia as Exposure

I did not find statistically significant MR associations between schizophrenia and any cardiometabolic trait after correction for multiple testing (Table 29). Similarly, I did not find statistically significant MR associations of inflammation-related schizophrenia variants with cardiometabolic traits after correction for multiple testing (Table 30).

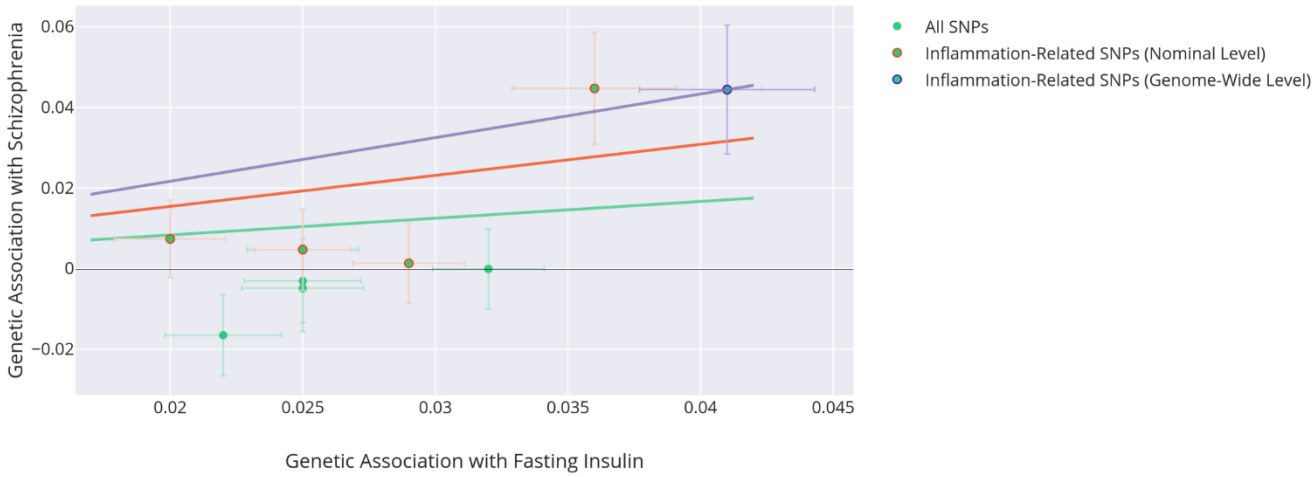
Table 26: MR Analyses of Inflammatory-Related Cardiometabolic SNPs and Schizophrenia

Risk Factor	Method	Genome-Wide Significant Inflammatory-Related SNPs				Nominally Significant Inflammatory-Related SNPs			
		SNPs, No.	Odds Ratio (95% C.I.)	*p*-value	Corrected *p*-value[a]	SNPs, No.	Odds Ratio (95% C.I.)	*p*-value	Corrected *p*-value[a]
Fasting Insulin	IVW / Wald Ratio	1	2.95 (1.38-6.34)	0.005	0.035	5	1.74 (1.08-2.98)	0.003	0.030
	Weighted Median						1.40 (0.83-2.34)	0.203	1.000
	MR Egger						7.20 (1.03-50.54)	0.141	0.987
Triglycerides	IVW / Wald Ratio	0	*	*	*	4	1.24 (1.07-1.55)	0.004	0.036
	Weighted Median						1.26 (1.06-1.50)	0.009	0.063
	MR Egger						1.29 (1.02-1.63)	0.167	0.987
HDL	IVW / Wald Ratio	1	0.55 (0.36-0.84)	0.005	0.035	7	0.78 (0.62-0.92)	0.004	0.036
	Weighted Median						0.77 (0.64-0.94)	0.008	0.056
	MR Egger						0.68 (0.51-0.91)	0.047	0.288
Fasting Plasma Glucose	IVW	2	1.53 (0.39-5.97)	0.537	1.000	4	1.04 (0.36-2.98)	0.945	1.000
	Weighted Median						1.08 (0.63-1.86)	0.776	1.000
	MR Egger						8.44 (0.65-120.54)	0.409	1.000
Type 2 Diabetes	IVW	7	0.94 (0.59-1.48)	0.776	1.000	10	0.97 (0.71-1.33)	0.850	1.000
	Weighted Median		1.05 (0.26-4.32)	0.941	1.000		1.05 (0.74-1.48)	0.781	1.000
	MR Egger		1.40 (0.32-6.08)	0.668	1.000		1.42 (0.59-3.38)	0.458	1.000
HbA1C	IVW	7	1.20 (0.67-2.13)	0.546	1.000	10	1.02 (0.64-1.61)	0.942	1.000
	Weighted Median		0.93 (0.46-1.85)	0.832	1.000		0.95 (0.54-1.69)	0.865	1.000
	MR Egger		1.68 (0.39-7.21)	0.508	1.000		1.18 (0.41-3.37)	0.767	1.000
Body Mass Index	IVW	4	1.23 (0.88-1.71)	0.229	1.000	12	1.48 (0.76-2.87)	0.249	1.000
	Weighted Median		1.15 (0.80-1.65)	0.451	1.000		1.16 (0.85-1.58)	0.350	1.000
	MR Egger		0.77 (0.33-1.79)	0.650	1.000		3.36 (0.61-18.45)	0.399	1.000
LDL	IVW	13	0.96 (0.79-1.17)	0.687	1.000	23	0.93 (0.79-1.10)	0.420	1.000
	Weighted Median		0.91 (0.80-1.04)	0.181	1.000		0.91 (0.80-1.04)	0.129	0.987
	MR Egger		0.81 (0.58-1.14)	0.254	1.000		0.82 (0.62-1.11)	0.220	0.987
Leptin	IVW	0	*	*	*	2	1.56 (0.77-3.17)	0.221	0.987
Glucose Tolerance	IVW	0	*	*	*	2	1.06 (0.82-1.56)	0.882	1.000

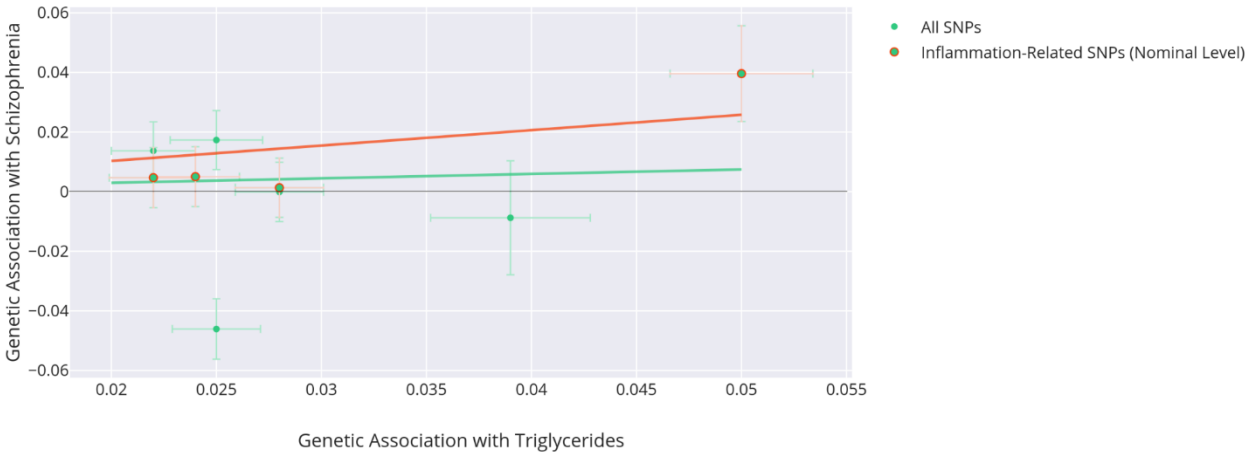
HDL=high-density lipoprotein; HbA1C=glycated haemoglobin; LDL=low-density lipoprotein; IVW=inverse variance weighted regression; SNPs=single nucleotide polymorphisms
[a]Each analysis method (IVW, Weighted Median and MR Egger) corrected using the Holm-Bonferroni method; *no identified inflammatory-related SNPs
Estimates represent ORs for schizophrenia per SD increase in exposure (or per unit-increase in log-odds of binary exposures e.g., T2D

Figure 14: MR Analyses Testing Associations of the Insulin Resistance Phenotype With Schizophrenia and Highlighting Inflammation-Related SNPs.

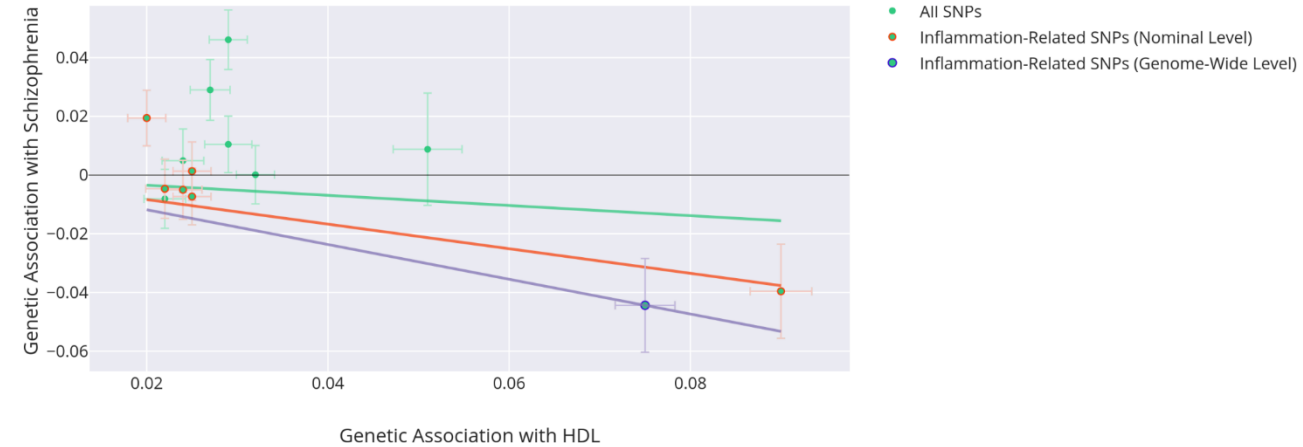
A. Fasting Insulin



B. Triglycerides

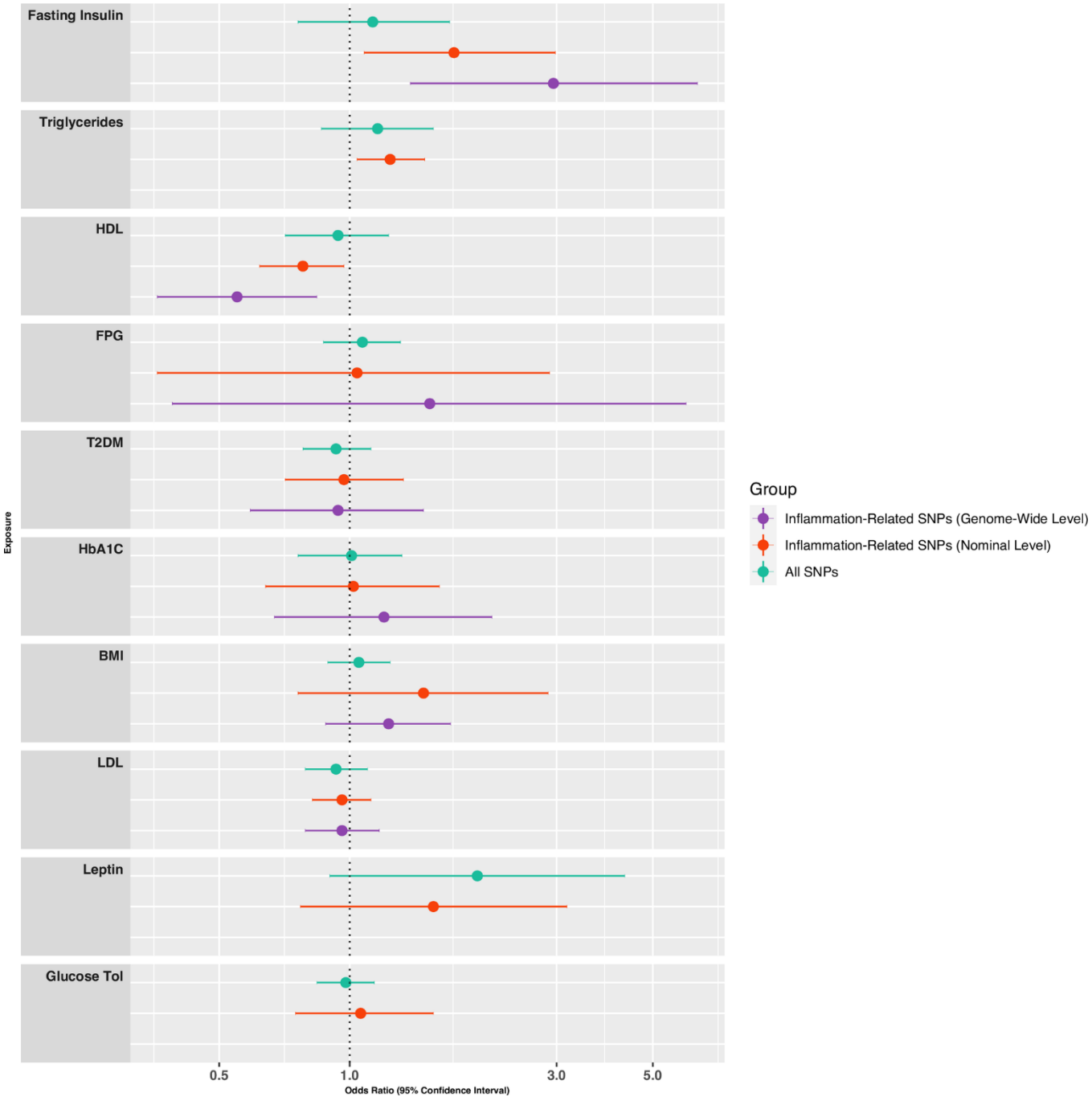


C. HDL



Points in plots represent the association of the genome-wide significant insulin-resistance single nucleotide polymorphisms (SNPs) and their association with schizophrenia (Y axis) and the exposure (X axis). SNPs are denoted by green points in the plot. Inflammation-related SNPs at genome-wide significance are denoted by a purple border. Inflammation-related SNPs at nominal significance are denoted by a red border. Whiskers represent SNP standard errors. Lines on the plot represent inverse-variance weighted (>1 SNP) or linear regression (1 SNP) of all-SNPs (green line), inflammation-related SNPs at genome-wide significance (purple line) and inflammation-related SNPs at nominal significance (red line).

Figure 15: MR Analyses Testing Associations between Cardiometabolic Traits and Schizophrenia



Forest plot presents ORs and 95% CIs for associations between cardiometabolic traits and schizophrenia using IVW / Wald Ratio MR analyses based on all single nucleotide polymorphisms (SNPs) associated with each risk factor (green), inflammation-related SNPs at genome-wide significance (purple), and inflammation-related SNPs at nominal significance (red). See Table 26 for the number of SNPs used in each analysis. HDL=High Density Lipoprotein; T2DM=Type 2 Diabetes Mellitus; BMI=Body Mass Index; FPG=Fasting Plasma Glucose; LDL=Low-Density Lipoprotein; HbA1C=Glycated Haemoglobin; Glucose Tol= Glucose Tolerance.

Table 27: Multivariable MR (MVMR) Results for Insulin Resistance-Phenotype Exposures (All-SNP analysis) with Addition of CRP as Exposure

Risk Factor	no. SNPs	Odds Ratio (95% C.I.)[a]	P-value
Fasting Insulin	10	0.96 (0.66-1.38)	0.813
CRP	*2*	*0.88 (0.62-1.23)*[b]	*0.456*
Triglycerides	10	0.98 (0.88-1.10)	0.756
CRP	*2*	*1.00 (0.0.65-1.56)*[b]	*0.987*
HDL	15	1.00 (0.86-1.18)	0.937
CRP	*2*	*0.92 (0.71-1.76)*[b]	*0.489*

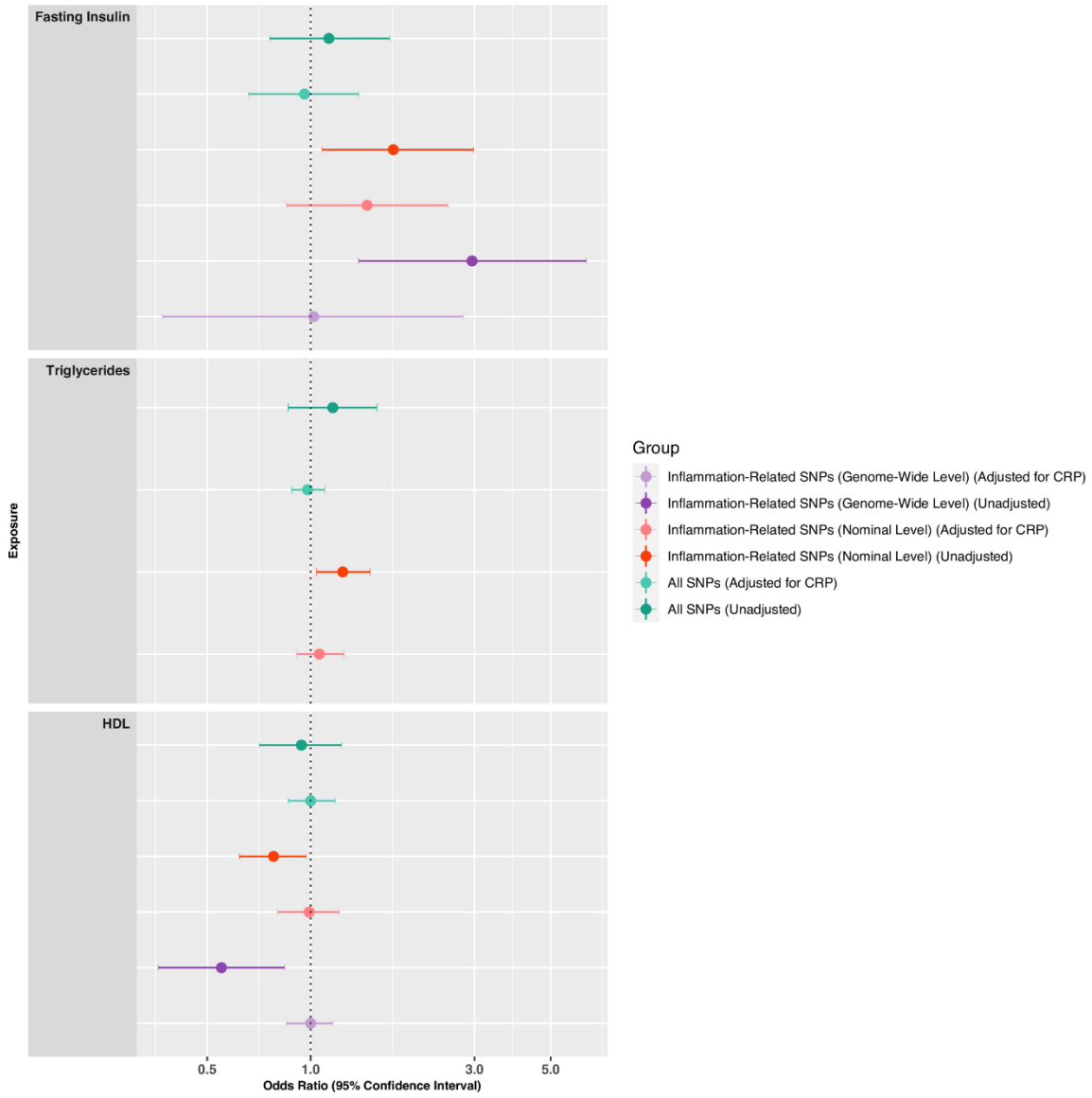
CRP=C-reactive protein; HDL=high-density lipoprotein; SNPs=single nucleotide polymorphisms
[a]Results for IVW MVMR analysis; [b]I did not perform univariable MR analysis for CRP since this was not a goal of the study. Univariable MR has been conducted and replicated for CRP and estimates are published elsewhere (Lin et al., 2019, Hartwig et al., 2017)

Table 28: Multivariable MR (MVMR) Results for Insulin Resistance-Phenotype Exposures (Inflammation-Related-SNP analysis) with Addition of CRP as Exposure

Risk Factor	Genome-Wide Significant Inflammation-Related SNPs			Nominally Significant Inflammation-Related SNPs		
	no. SNPs	Odds Ratio (95% C.I.)[a]	*P*-value	no. SNPs	Odds Ratio (95% CI)[a]	*P*-value
Fasting Insulin	1	1.02 (0.37-2.78)	0.975	5	1.46 (0.85-2.51)	0.307
CRP	*2*	0.94 (0.40-2.18)[b]	0.881	2	1.27 (0.80-2.02)	0.308
Triglycerides	-	-	-	4	1.06 (0.91-1.25)	0.447
CRP	-	-	-	2	0.70 (0.45-1.45)	0.343
HDL	1	1.00 (0.85-1.16)	0.849	7	0.99 (0.81-1.21)	0.731
CRP	*2*	0.90 (0.72-1.12)[b]	0.367	2	0.90 (0.76-1.08)	0.251

CRP=C-reactive protein; HDL=high-density lipoprotein; SNPs=single nucleotide polymorphisms
[a]Results for IVW MVMR analysis; [b]I did not perform univariable MR analysis for CRP since this was not a goal of the study. Univariable MR has been conducted and replicated for CRP and estimates are published elsewhere (Lin et al., 2019, Hartwig et al., 2017)

Figure 16: Multivariable MR Analysis Testing Associations between Insulin Resistance Phenotypes and Schizophrenia After Controlling for Genetic Associations with CRP



Forest plot presents ORs and 95% CIs for associations between insulin resistance phenotypes and schizophrenia using IVW / Wald Ratio MR and MVMR analyses based on all single nucleotide polymorphisms (SNPs) associated with each risk factor and unadjusted for CRP (dark green), all SNPs associated with each risk factor and adjusted for CRP (light green), inflammation-related SNPs at genome-wide significance and unadjusted for CRP (dark purple), inflammation-related SNPs at genome-wide significance and adjusted for CRP (light purple), inflammation-related SNPs at nominal significance and unadjusted for CRP (dark red), and inflammation-related SNPs at nominal significance and adjusted for CRP (light red). See Tables 27 and 28 for the number of SNPs used in each analysis. HDL=High Density Lipoprotein.

Table 29: Bidirectional MR Analyses Using All SNPs for Schizophrenia With Cardiometabolic Outcomes

Outcome	SNPs, No.	Method	β (S.E)	P-value	Corrected *p*-value[a]
Fasting Insulin	101	IVW	0.01 (0.02)	0.496	1.000
		Weighted Median	0.02 (0.02)	0.268	1.000
		MR Egger	-0.05 (0.08)	0.542	1.000
Triglycerides	101	IVW	0.00 (0.02)	0.970	1.000
		Weighted Median	0.00 (0.03)	0.987	1.000
		MR Egger	0.05 (0.11)	0.642	1.000
HDL	101	IVW	-0.02 (0.03)	0.521	1.000
		Weighted Median	-0.01 (0.03)	0.901	1.000
		MR Egger	-0.04 (0.05)	0.051	0.510
Fasting Plasma Glucose	105	IVW	0.01 (0.01)	0.339	1.000
		Weighted Median	0.01 (0.01)	0.454	1.000
		MR Egger	0.00 (0.06)	0.994	1.000
Type 2 Diabetes	109	IVW	-0.01 (0.06)	0.845	1.000
		Weighted Median	0.00 (0.08)	1.000	1.000
		MR Egger	0.14 (0.30)	0.645	1.000
Body Mass Index	101	IVW	-0.03 (0.02)	0.220	1.000
		Weighted Median	-0.03 (0.02)	0.146	1.000
		MR Egger	0.18 (0.10)	0.081	0.729
HbA1C	104	IVW	0.01 (0.01)	0.911	1.000
		Weighted Median	0.01 (0.02)	0.730	1.000
		MR Egger	0.01 (0.07)	0.948	1.000
Glucose Tolerance	101	IVW	0.08 (0.07)	0.278	1.000
		Weighted Median	0.12 (0.10)	0.233	1.000
		MR Egger	0.24 (0.35)	0.496	1.000
LDL	101	IVW	-0.06 (0.03)	0.079	0.790
		Weighted Median	-0.06 (0.05)	0.080	0.800
		MR Egger	-0.22 (0.14)	0.113	0.904
Leptin	101	IVW	0.02 (0.02)	0.239	1.000
		Weighted Median	0.01 (0.03)	0.677	1.000
		MR Egger	-0.02 (0.09)	0.810	1.000

HDL=high-density lipoprotein; HbA1C=glycated haemoglobin; LDL=low-density lipoprotein; SNPs=single nucleotide polymorphisms; IVW=inverse variance weighted regression; β=beta coefficient; S.E=standard error. [a]Adjusted using the Holm-Bonferroni method for multiple testing.

Table 30: Bidirectional MR Analyses Using Inflammation-Related SNPs for Schizophrenia With Cardiometabolic Outcomes

Outcome	SNPs, No.	Method	β (S.E)	P-value	Corrected P-value[a]
Fasting Insulin	3	IVW	0.04 (0.05)	0.409	1.000
		Weighted Median	0.03 (0.06)	0.666	1.000
		MR Egger	0.00 (0.09)	0.976	1.000
Triglycerides	3	IVW	0.20 (0.09)	0.034	0.340
		Weighted Median	0.20 (0.08)	0.009	0.090
		MR Egger	0.28 (0.30)	0.306	1.000
HDL	1	Wald Ratio	-0.26 (0.21)	0.202	1.000
LDL	3	IVW	0.11 (0.07)	0.953	1.000
		Weighted Median	0.06 (0.07)	0.341	1.000
		MR Egger	-0.01 (0.10)	0.895	1.000
Body Mass Index	4	IVW	-0.02 (0.09)	0.750	1.000
		Weighted Median	0.00 (0.05)	0.832	1.000
		MR Egger	0.05 (0.02)	0.705	1.000
Type 2 Diabetes	3	IVW	-0.18 (0.34)	0.598	1.000
		Weighted Median	0.10 (0.30)	0.729	1.000
		MR Egger	0.24 (0.80)	0.789	1.000
Fasting Plasma Glucose	4	IVW	-0.02 (0.07)	0.780	1.000
		Weighted Median	-0.05 (0.04)	0.317	1.000
		MR Egger	-0.16 (0.10)	0.258	1.000
HbA1C	3	IVW	-0.07 (0.07)	0.269	1.000
		Weighted Median	0.06 (0.05)	0.137	1.000
		MR Egger	-0.18 (0.12)	0.292	1.000
Glucose Tolerance	4	IVW	0.12 (0.26)	0.648	1.000
		Weighted Median	0.10 (0.31)	0.732	1.000
		MR Egger	-0.09 (0.49)	0.872	1.000
Leptin	4	IVW	-0.03 (0.07)	0.646	1.000
		Weighted Median	-0.04 (0.08)	0.619	1.000
		MR Egger	-0.10 (0.12)	0.526	1.000

HDL=high-density lipoprotein; HbA1C=glycated haemoglobin; LDL=low-density lipoprotein; SNPs=single nucleotide polymorphisms; IVW=inverse variance weighted regression; β=beta coefficient; S.E=standard error. [a]Adjusted using the Holm-Bonferroni method for multiple testing.

5.4.5 Test for Horizontal Pleiotropy

Using the MR-Egger regression intercept test, I found evidence of potential horizontal pleiotropy for BMI and HDL in the all-SNP analysis, but no evidence for horizontal pleiotropy for any cardiometabolic exposure in the inflammation-related SNP analysis. Using MR-PRESSO however, I found evidence that horizontal pleiotropy was likely to have affected estimates for all cardiometabolic exposures in the all-SNP analysis (*p* value for global test all ≤0.020), and both LDL and T2D in the inflammation-related SNP analysis. Following MR-PRESSO outlier correction, evidence strengthened for the association of triglycerides with schizophrenia in the all-SNP analysis (MR-PRESSO IVW β=0.23, S.E. 0.06, *p*=0.008), but outlier-corrected IVW estimates for other exposures

were not significantly altered. In bidirectional analyses, both MR-PRESSO and the MR-Egger regression intercept suggested horizontal pleiotropy affecting the outcomes of HDL, BMI, and LDL (all $p<0.05$). There was evidence for a weak protective effect of schizophrenia on BMI following outlier correction (β=-0.04, S.E. 0.02, $p=0.014$). MR-PRESSO additionally revealed possible horizontal pleiotropy affecting the outcomes of fasting insulin, triglycerides and T2DM (p for MR-PRESSO global test all <0.05), but outlier-corrected IVW estimates were not significantly altered. See Appendix C Tables 22-29 for full horizontal pleiotropy sensitivity analysis results.

5.4.6 Test for Heterogeneity of Instruments

In the analyses based on all-SNPs, the majority of cardiometabolic traits demonstrated evidence of heterogeneity, which was reduced in the inflammation-related SNP analysis. See Appendix C Tables 22-29 for full heterogeneity of instruments sensitivity analysis results.

5.4.7 Test for Measurement Error

Results for the I^2_{GX} tests for SNP-exposure associations revealed some evidence for potential measurement error which may have biased MR Egger analyses in the analyses with leptin, glucose tolerance, T2DM and schizophrenia as exposures. See Appendix C Table 30.

5.5 Discussion

I conducted bidirectional uni- and multi-variable two-sample MR analyses using large publicly available genomic datasets to first examine for associations that support a causal relationship between insulin resistance/related cardiometabolic traits and schizophrenia, and second, to examine whether there is evidence in support of the hypothesis that inflammation may be a common causal mechanism for insulin resistance and schizophrenia. Using the primary IVW analysis method, I did not find evidence supporting a causal association between genetically predicted cardiometabolic traits and schizophrenia. However, I found weak evidence using the weighted median method to support a causal association of genetically predicted levels of triglycerides and HDL with schizophrenia, but these associations did not survive correction for multiple testing, and the estimates may have been affected by horizontal pleiotropy.

I found more consistent evidence for an association of an insulin resistance phenotype of fasting insulin, triglycerides, and HDL (Lotta et al., 2017a) with schizophrenia when I examined only genetic variants also associated with inflammation. Using two *p*-value cut-offs for inflammation-related SNPs, I found that the strength of association with schizophrenia increased as the specificity toward inflammation-related SNPs increased. In MVMR analyses adjusting for CRP, those estimates attenuated to the null. I found no evidence in bidirectional analyses supporting a causal relationship of schizophrenia with insulin resistance (Panels C&D in Figure 13). Together, the results are therefore most consistent with inflammation as a common cause for insulin resistance and schizophrenia (Panel A in Figure 13).

Three aspects of the results point toward inflammation as a common cause for insulin resistance and schizophrenia. First, I did not find convincing evidence for a causal relationship between insulin resistance and schizophrenia (likely ruling out Panel B in Figure 13). Second, in the analyses of inflammation-related variants for the cardiometabolic traits, I found more consistent evidence supporting a potential causal relationship of fasting insulin, HDL and triglycerides with schizophrenia, and the strength of association with schizophrenia increased as the specificity toward inflammation-related SNPs increased. Third, I used MVMR to evidence that after controlling for CRP, an archetypal generalized marker of inflammation, the associations between inflammation-related genetic variants for insulin resistance and schizophrenia wholly attenuated. This result suggests that the observed associations for the inflammation-related variants are at least in part explained by inflammation. Together, the results are consistent with the idea that inflammation may be a common causal mechanism for insulin resistance and schizophrenia.

Evidence for a common-causal mechanism between insulin resistance and schizophrenia may help to explain why schizophrenia is associated with higher rates of insulin resistance even in the early stages of illness when the cumulative effects of medication and lifestyle factors are relatively small (Perry et al., 2016, Pillinger et al., 2017a). Anti-inflammatory agents, of which several have shown promise in treating the symptoms of schizophrenia (Cakici et al., 2019), should therefore be considered a putative therapeutic target for the prevention or treatment of cardiometabolic disorders in schizophrenia.

I used CRP, an archetypal downstream inflammatory marker, as a means of gauging the effect of systemic inflammation in MVMR analysis, rather than hypothesizing a specific role for CRP in the relationship between insulin resistance and schizophrenia. Nevertheless, CRP has observationally shown in both cross-sectional (Fernandes et al., 2016) and longitudinal (Metcalf et al., 2017) research to be associated with schizophrenia. However, such findings are limited by the potential for residual confounding and reverse causality. Interestingly, MR findings have reported that genetically predicted CRP may have a protective effect on schizophrenia (Hartwig et al., 2017), with authors positing that a genetically attenuated ability to produce CRP may predispose to more insidious and chronic infections. In MVMR analysis, attenuation of insulin resistance-schizophrenia associations after controlling for CRP is consistent with inflammation being associated with both exposure and outcome, albeit 'negatively' with the latter. Further research is needed to explore potential mechanisms of association between CRP and schizophrenia.

Many of the SNPs included in the inflammation-related analysis were associated with neutrophils and lymphocytes. A raised neutrophil to lymphocyte ratio (NLR) is a marker of systemic inflammation and is known to be associated with schizophrenia (Karageorgiou et al., 2018) and insulin resistance (Lou et al., 2015). However, I could not identify large GWAS studies conducted in European populations for NLR or for other inflammatory markers, which I might have used in MVMR analyses in place of CRP.

Based on the current results, one cannot completely rule out the possibility that insulin resistance may mediate an inflammation-schizophrenia association (Panel B in Figure 13). There was weak evidence that did not survive correction for multiple testing for an association of triglycerides and HDL with schizophrenia using the weighted median method. In the MR-PRESSO sensitivity analysis, evidence from the outlier-corrected IVW analysis suggested a possible association between triglycerides and schizophrenia. These findings are broadly similar to one previous MR study (Polimanti et al., 2017), which reported only weak evidence of an association between HOMA-IR and schizophrenia. Another MR study (Li et al., 2018) reported a genetic association between fasting insulin and schizophrenia, although the evidence attenuated after adjustment for BMI. To account for BMI, I obtained summary

statistics for genetic variants related to insulin resistance after controlling for BMI (Lotta et al., 2017b). The previous MR study included an ethnically heterogeneous sample, increasing the potential for population stratification bias. I used genetic data from a more ethnically homogenous GWAS of schizophrenia (Pardinas et al., 2018). Nevertheless, while the results in the all-SNP analysis suggested weak evidence for triglycerides and HDL, which may reflect an insulin resistance phenotype, the evidence did not survive correction for multiple testing and requires replication in future when larger GWAS samples are available.

Regarding additional findings, after outlier correction, I found that schizophrenia had a weak protective effect on BMI. This finding complements estimates from previous research, which has reported a negative genetic correlation between schizophrenia and BMI (Bahrami et al., 2020). This finding suggests that weight gain associated with schizophrenia is unlikely to be a feature of the illness itself but could be attributed to iatrogenic or lifestyle effects. Moreover, the 'lean insulin resistance' phenotype may be associated with higher levels of inflammation (Ding et al., 2016). In addition, the 'lean' nature of the phenotype may mean that critical cardiometabolic investigations may be overlooked, particularly in younger patients.

Strengths of this study include the use of a large set of cardiometabolic traits and large GWAS datasets, through which I could test specific biological mechanisms. I selected SNPs reaching genome-wide significance from large GWAS and meta-GWAS for insulin resistance and related cardiometabolic traits. I performed a comprehensive set of sensitivity analyses to check the validity and robustness of the findings. Furthermore, whilst weak-instrument bias may be a factor in MR analysis, in two-sample MR this bias tends toward the null (Davies et al., 2015) so would not explain the positive associations described in this study. I corrected for multiple testing to minimise potential type I error.

This study has some limitations. I did not select SNPs in known coding regions for the exposures, for example, the *IRS-1* gene for insulin resistance (Carvalho et al., 1999). I took this step on the assumption that many mechanisms at play may not yet be fully understood. For example, whilst the heritability of cardiometabolic traits such as obesity is as high as 70%, the variance currently explained by known genetic variants is a small fraction of this (Herrera et al., 2011). In addition, selecting SNPs from many different GWAS studies featuring large sample sizes may increase the risk of sample-overlap between exposure and outcome samples and can bias the results in either direction, depending on the proportion of overlap (Hemani et al., 2018). Also, for the primary inflammation-related SNP analysis, I chose a stringent *p*-value threshold to define inflammation-related SNPs. In doing so, I may have overlooked some SNPs with genuine inflammatory associations. As a result, only one genome-wide significant inflammation-related genetic variant was included in the analysis

of fasting insulin and HDL, and none could be included for triglycerides. Therefore, these results should be considered with caution. However, I attempted to address this limitation by relaxing the *p*-value threshold for inflammation-related SNPs, thereby allowing a larger number of SNPs to be included, and the results for fasting insulin, HDL and triglycerides were consistent. Yet, the inclusion of inflammation-related genetic variants at a relaxed significance threshold may have increased the risk of weak instrument bias for those analyses. In the future, better-powered GWAS may identify more SNPs for analysis and at greater resolution, potentially unearthing a larger number of inflammation-related SNPs and at greater strength of association, which would be helpful to confirm the findings.

Additionally, the full range of gene products from the genetic variants I used as proxies for the cardiometabolic traits is unknown. So, I cannot comment on potential biological mechanisms of association other than inflammation, which may also be relevant. Finally, the analyses were based on primarily European participants, so it is unclear whether the results of this study apply to other populations. Large-scale GWAS and replication of these analyses in different populations are required to answer this question.

In conclusion, it is well established that certain antipsychotic drugs and lifestyle factors such as smoking, lack of exercise and poor diet are important contributors to cardiometabolic comorbidity in people with schizophrenia. In addition, the findings from this study suggest that inflammation may be a common cause for schizophrenia and insulin resistance, which may at least partly explain why they so commonly co-occur in clinical practice. Lifestyle modification and careful prescription of certain antipsychotic medications remain crucial malleable targets to reduce the significant impact of comorbid cardiometabolic disorders on the quality and length of life in people with schizophrenia. However, findings from this study suggest that targeting inflammation could be an important therapeutic target for the treatment and prevention of cardiometabolic disorders in people with schizophrenia.

Section C: Summary of Main Findings and Conclusions

In Chapter 3, using prospective ALSPAC data, I report that genetic predisposition for disrupted glucose-insulin homeostasis was associated with an increased risk of schizophrenia-spectrum outcomes at age 18 years and *vice versa*. These findings provide evidence for shared genetic liability for comorbid schizophrenia and disrupted glucose-insulin homeostasis. I also report evidence for a mediating effect of childhood inflammation on the association between genetic predisposition for disrupted glucose-insulin homeostasis and psychosis risk in adulthood. These results suggest that genetic variation may influence biological pathways leading to inflammatory changes, which in turn increases the risk of both disrupted glucose-insulin homeostasis and schizophrenia in adulthood.

In Chapter 4, I found further evidence from large samples that cardiometabolic and inflammatory traits share genetic overlap with schizophrenia. I also found a set of biologically plausible common-causal variants that could influence biological pathways, particularly involving BDNF and glucose transport, which could influence inflammation, glucose-insulin homeostasis, and risk of schizophrenia.

In Chapter 5, using MR, I found evidence supporting that inflammation-related insulin resistance may be causally related to schizophrenia. These findings suggest that inflammation may be a common cause of schizophrenia and comorbid cardiometabolic disorders.

Together, the results from Chapters 3-5 comprising Section C suggest that a summation of genetic variation may influence biological pathways leading to changes in inflammation/immune function, which in turn increases the risk of both disrupted glucose-insulin homeostasis and schizophrenia.

SECTION D

IMPROVING THE PREDICTION OF CARDIOMETABOLIC RISK IN SCHIZOPHRENIA

Section D Summary

In Section D, I aimed to examine and improve the clinical prediction of cardiometabolic risk in schizophrenia. In this section, I focussed my attention on young people at the onset of psychotic illness since primary prevention is the best means to reduce the risk of adverse cardiometabolic outcomes (Weintraub et al., 2011).

Therefore, in Chapter 6, I performed a systematic review of cardiometabolic risk prediction algorithms developed either for the general or psychiatric populations and examined whether they may be suitable for young people with psychosis. In this detailed review of over 100 studies, I found that all existing algorithms were developed in relatively older adults; most were at high risk of bias; most were not externally validated; and few considered relevant predictors such as antipsychotic medication. Further, I performed a validation analysis in ALSPAC of three cardiometabolic risk prediction algorithms commonly used in clinical practice, testing their predictive ability in a sample of young adults who had/were at risk of developing psychosis. I found that the algorithms substantially underpredicted cardiometabolic risk in the younger psychosis-risk population. Therefore, I concluded no existing cardiometabolic risk prediction algorithms can be recommended for use in young people with psychosis. Findings from this study have been published in *Acta Psychiatrica Scandinavica* (Perry et al., 2020c). See Appendix D for the published manuscript.

Given the lack of an appropriate algorithm for young people with psychosis, in Chapter 7, I used patient data from three EIS to develop and externally validate the first cardiometabolic risk prediction algorithm tailored specifically for young people with psychosis, the Psychosis Metabolic Risk Calculator (PsyMetRiC). I developed two versions of PsyMetRiC, one with and one without biochemical results for clinical practicality. I developed PsyMetRiC in consultation with a young person's advisory group to maximise patient acceptability. I performed a detailed set of analyses to examine the predictive performance and potential clinical usefulness of PsyMetRiC, and developed an online data visualisation app. The findings from this study have been accepted for publication in *The Lancet Psychiatry*. The manuscript is currently at the proofing stage with the journal.

Chapter 6

Cardiometabolic Risk Prediction Algorithms for Young People with Psychosis: A Systematic Review and Validation Analysis

6.1 Introduction

Physical comorbidity is a leading cause of significantly higher mortality rates and reduced life expectancy for people with schizophrenia compared with the general population (Laursen et al., 2019, Plana-Ripoll et al., 2019, Plana-Ripoll et al., 2020). Therefore, there is a clear and crucial need for clinical tools to identify cardiometabolic risk in this group in order to optimise care and improve long-term outcomes. Yet, a recent report of a small sample of people with chronic schizophrenia suggests that some commonly used cardiometabolic risk prediction algorithms return differing risk prediction scores when tested on the same participants. This calls into question the reliability and suitability of such algorithms for relatively older people with chronic schizophrenia, let alone young people with psychosis (Berry et al., 2018).

Recent evidence suggests that the physical comorbidity associated with schizophrenia starts early. Markers of developing cardiometabolic disorders are a feature that distinguish cases of first-episode psychosis from matched general population controls (Perry et al., 2016, Pillinger et al., 2017a) and are associated with young adults at risk of developing psychosis (Perry et al., 2018). The field of early intervention in psychosis rests on the premise that intervening early could improve longer-term outcomes, and this premise applies equally to the treatment of cardiometabolic disorders. Therefore, cardiometabolic risk prediction algorithms may be a valuable tool for healthcare professionals to help tailor treatment plans for young people with psychosis that could help to reduce both long-term physical and psychiatric morbidity. However, such a tool could only be clinically useful if the predictions it makes are accurate. It is unclear as to whether this may or may not be the case.

6.2 Aims and Objectives

I conducted a systematic review to identify and compare existing cardiometabolic risk prediction algorithms developed for the general or psychiatric populations and consider their suitability for young people with psychosis. Next, I performed an exploratory validation analysis using data from ALSPAC to examine the predictive performance of any algorithms highlighted as potentially suitable by the review in a sample of young adults with or at risk of developing psychosis. To explore the impact of age on predictive performance, I reassessed model performance after artificially increasing the age of participants to the mean age of the original algorithm development study, leaving all other predictors unchanged.

6.3 Methods

6.3.1 Systematic Review

6.3.1.1 Literature Search

I conducted a systematic literature search of EMBASE (1947-present), Ovid MEDLINE (1946-present), PsychINFO (1806-present), Web of Science (from inception), and the first twenty pages of Google Scholar (Haddaway et al., 2015) to 1st December 2019. I also searched the references of included studies. The search strategy is presented below. MeSH headings (denoted with *) and text terms were used:

Group 1: metabolism* (OR) metabolic* (OR) diabetes mellitus* (OR) cardiovascular diseases* (OR) obesity* (OR) cardiometabolic

(AND)

Group 2: risk assessment* (OR) risk* (OR) outcome assessment* (OR) patient outcome assessment* (OR) prognosis*

(AND)

Group 3: calculator (OR) computers* (OR) algorithms* (OR) software* (OR) tool.

I applied the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-analyses) guidelines (Moher et al., 2009). The systematic review was registered on PROSPERO (CRD42019150377).

6.3.1.2 Study Selection

The inclusion criteria were as follows;

(1) Studies reporting the development and/or validation of cardiometabolic risk algorithms designed for either the general or psychiatric populations;

(2) studies which: reported in combination the development and validation (internal or external) of an original algorithm; reported the development but not validation of an algorithm; reported the first validation of a previously developed but not validated algorithm; or reported a new recalibration of a previously developed algorithm;

(3) Cardiometabolic risk was defined as CVD (stroke, myocardial infarction, hypertension, unstable angina) and its pre-determinants including T2D, prediabetes, obesity, or dyslipidaemia;

(4) Studies reported in any language;

(5) Published and unpublished research, conference proceedings and academic theses.

The exclusion criteria were as follows:

(1) algorithms designed specifically for other defined health groups (e.g., post-operative patients or patients with any physical health diagnoses at baseline);

(2) studies reporting validation without recalibration of previously validated algorithms.

Titles and abstracts were screened independently by three researchers (Dr Benjamin Perry; Mr Owen Crawford; Miss Soomin Jang) prior to full-text screening. Any discrepancies were resolved in consultation with a senior researcher (Professor Golam Khandaker). Data were extracted from studies that met the inclusion criteria. Searches were re-run immediately prior to the final analyses, and further studies retrieved for inclusion using the processes outlined above.

6.3.1.3 Data Extraction and Synthesis

I extracted data on general characteristics (e.g., population, location, study type, type of risk predicted), the characteristics of included participants (e.g., age, sex, ethnicity), and characteristics of the developed/validated algorithms (e.g., included predictors, algorithm performance statistics). Risk of bias was assessed using the 'Prediction model Risk Of Bias Assessment Tool' (PROBAST) (Wolff et al., 2019), which aims to identify shortcomings in study design, conduct, or analysis that could lead to systematically distorted estimates of model predictive performance. PROBAST includes four domains for potential sources of bias in prediction model studies (participants, predictors, outcome, and analysis) which are then summarised by an overall judgement; either low-risk, high-risk or unclear-risk of bias (Wolff et al., 2019). I plotted the range and frequency of predictors included in studies. I illustrated the relative weighting of different predictors in the single included study that featured psychiatric predictors. Algorithm performance was compared using statistics relating to model discrimination (how well an algorithm discriminates people at higher-risk from people at lower-risk, e.g. Harrell's C Statistic, where a score of 1.0 indicates perfect discrimination, and a score of 0.5 indicates the model is no better than chance) and model calibration (the accuracy of absolute-

risk estimates, e.g. calibration plots) (Alba et al., 2017). I also examined the events-per-variable ratio (EPV) (the ratio of outcome events: predictors considered in algorithm development) of each study to assess the potential risk for model overfit (Peduzzi et al., 1996). An EPV of 10 or more had previously been considered satisfactory (Pavlou et al., 2015), though more recently, higher EPV ratios are often advised (Ogundimu et al., 2016). Where an EPV ratio was not reported, I calculated it where possible from the information available in the study. Finally, I considered the likely suitability of included algorithms for young people with psychosis. I summarized and compared studies with a narrative synthesis (Rodgers, 2009).

6.3.2 Exploratory Validation Analysis

6.3.2.1 Data Source

See Section 2.3.1 for a full description of the ALSPAC cohort. Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and Local Research Ethics Committees. All participants provided informed consent.

6.3.2.2 Study Sample

I included participants who at either age 18 or 24 years were identified as experiencing PEs or psychotic disorder. See Section 2.3.3 and Section 3.3.2 for detail on assessment of PEs and psychotic disorder at age 18 and 24 years. I excluded participants who already met the outcome criteria at age 18 years, and participants who had missing data on all included variables. Additionally, I conducted a *post-hoc* sensitivity analysis to examine the potential impact of sample size; I reperformed the analysis including all participants from the total ALSPAC sample at age 18 years who did not meet the criteria for the outcome at age 18 years, and who did not have missing data on all included variables. In total, after exclusions, I included 505 participants. See Appendix D Figures 1-2 for flow-charts of included participants.

6.3.2.3 Outcome

I used the harmonized definition (Alberti et al., 2009) of the metabolic syndrome measured at age 24y as the outcome, which it is an established precursor of T2D (Shin et al., 2013) and CVD (Wilson et al., 2005), and is an appropriate cardiometabolic outcome for young adults. See Table 2 for the diagnostic criteria. For blood-based predictors (FPG, HDL and triglycerides), fasting samples were

taken at 0900 after a 10-hour fast (water only) at age 18 years. Samples were immediately spun, frozen and stored at –80ºC and measurements were assayed within 3 to 9 months of the samples being taken with no previous freeze-thaw cycles. FPG was measured by an ultrasensitive ELISA (Mercodia, Uppsala, Sweden) automated microparticle enzyme immunoassay. Its sensitivity was 0.07 mU/L, and inter- and intra-assay coefficients of variation were <6%. Plasma lipid concentrations were measured by modification of the standard Lipid Research Clinics Protocol by using enzymatic reagents for lipid determination.

6.3.2.4 Predictors

I included all available predictors from QRISK3 (Hippisley-Cox et al., 2017), QDiabetes (Hippisley-Cox and Coupland, 2017) and PRIMROSE (Osborn et al., 2015), which were the three algorithms highlighted as being potentially the most suitable for young people with psychosis. These included age, Townsend deprivation score, body mass index (BMI), ethnicity, smoking, antipsychotic medication use, antidepressant use, corticosteroid use, psychosis, depression, family history of cardiovascular disease or type 2 diabetes, hypertension, FPG, cholesterol:HDL ratio, systolic blood pressure, total cholesterol, HDL, alcohol intake, and year of assessment. See Appendix D Methods & Appendix D Table 1 for a detailed description of the coding and assessment of predictors.

6.3.2.5 Missing Data

To address the impact of missing data, I used multiple imputation using chained equations (MICE) (Buuren, 2011) for variables which: 1) had <40% missing data (Lee, 2011) from the sample of participants with data on the outcome; 2) had suitable auxiliary variables available to use as 'indicators of missingness', to reduce the impact of bias attributed by the risk of data being 'missing not at random' (Dong and Peng, 2013). Auxiliary variables were selected based upon contributing to reducing the fraction of missing information (Madley-Dowd et al., 2019). Multiple imputation of 10 datasets was used to replace missing continuous predictor data, using the *MICE* package (Buuren, 2011) in *R* (R Core Team, 2017). Equivalent biochemical and questionnaire data taken at age 15 years were included as auxiliary predictor variables in *MICE*. Box-and-Whisker and Density plots were used to check similarities of observed and imputed data. Rubin's rules were used to pool analyses.

6.3.2.6 Statistical Analysis

Estimated six-year risk estimates for metabolic syndrome were calculated for QDiabetes (Hippisley-Cox and Coupland, 2017), QRISK3 (Hippisley-Cox et al., 2017) and PRIMROSE (Osborn et al., 2015), by applying the published fully-specified algorithms to the sample. QDiabetes and PRIMROSE comprise different models depending on the availability of blood test results. Therefore, I used the model which performed best in the original model development studies (Hippisley-Cox and Coupland, 2017, Osborn et al., 2015). For QDiabetes, the best performing model included FPG; for PRIMROSE, the best performing model included lipids. QDiabetes and QRISK3 estimate risk separately for males and females. Algorithm performance was assessed using measures of discrimination (Harrell's C-statistic and R^2); and a measure of calibration (calibration plots). Calibration plots included grouped observations, which were split at each 0.2 of predicted risk. First, I calculated model performance using actual participant age (18y). To assess the impact of age on model performance, I artificially substituted every participants' age in ALSPAC to the mean age from the original algorithm development study (QDiabetes=44.9y; QRISK3=42.9y; PRIMROSE=49.5y), leaving all other predictors unchanged. I re-ran each algorithm and compared the model performance statistics described above. Statistical analysis was carried out in *R* version 3.6.0 (R Core Team, 2017).

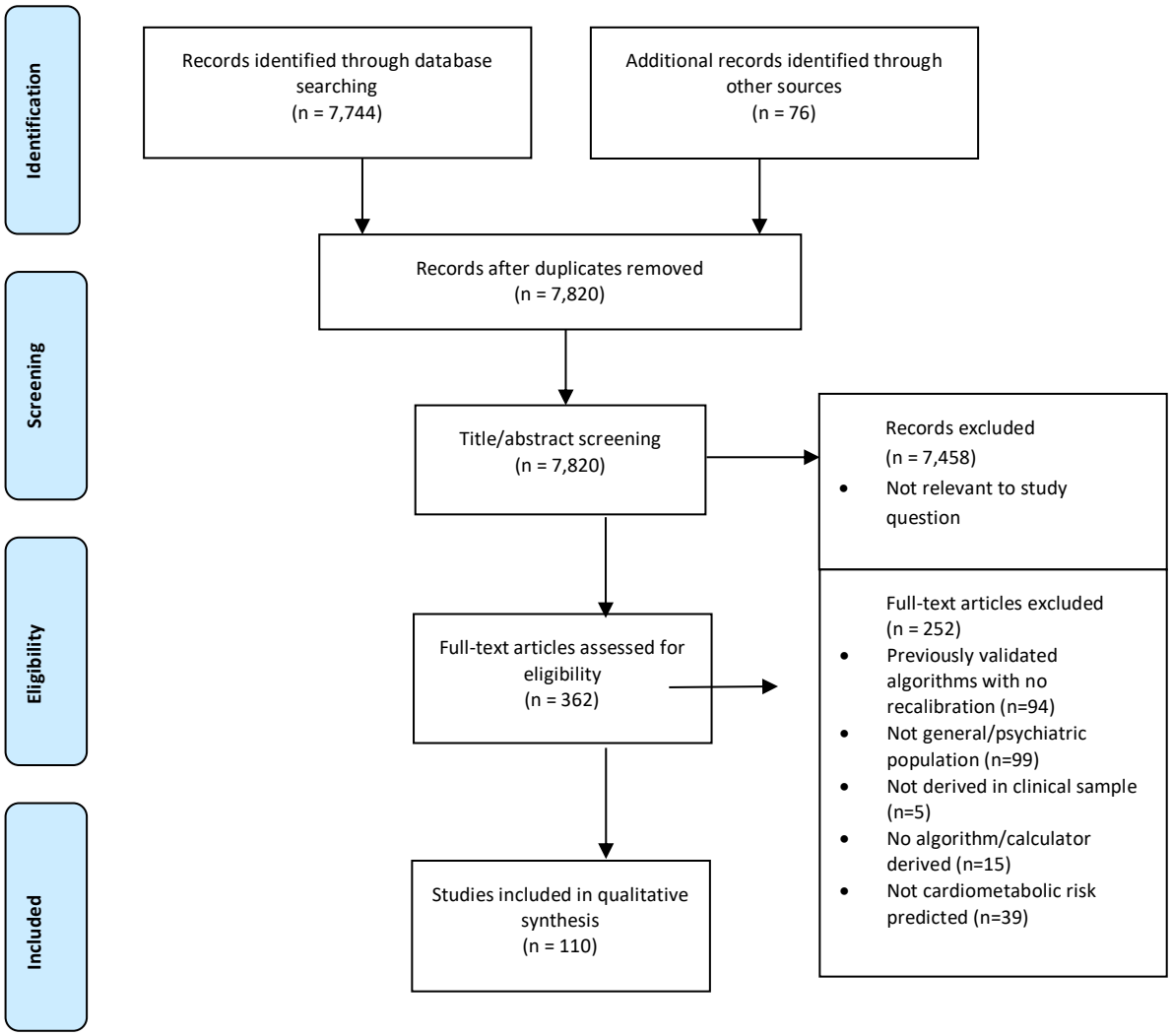
6.4 Results

6.4.1 Systematic Review

6.4.1.1 Study Selection and Quality Assessment

The literature search returned 7,744 results after removing duplicates. I reviewed 362 full texts, of which 110 studies met inclusion criteria. See Appendix D Results for a full list of studies included in the systematic review. See Figure 17 for the PRISMA diagram. Three studies were not published in peer-reviewed journals but were published either as conference proceedings (Hossain, 2018), a thesis (Boucher, 2019) or a pre-print (Gupta, 2019). Reporting quality was relatively poor across the majority of studies, with 108 studies (98%) either at unclear or high-risk of bias following assessment with the PROBAST tool (Wolff et al., 2019). See Appendix D Table 2 for full PROBAST results.

Figure 17: Systematic Review PRISMA Diagram



6.4.1.2 Study Characteristics

Appendix D Table 3 reports in detail the characteristics of included studies. To summarise, all studies were conducted on general population samples of healthy adults, except one which was conducted on patients with severe mental illness, defined as either schizophrenia, other psychotic disorder, or bipolar disorder (Osborn et al., 2015). The majority of included studies were conducted in high-income or upper-middle-income countries, with the UK, USA and China best represented. Eleven studies were conducted in lower- or middle-income countries. Sample sizes were highly variable in both development (from *n*=100 participants (Park et al., 2009) to *n*=8,136,705 participants (Hippisley-Cox and Coupland, 2017)) and validation cohorts (from *n*=90 participants (Friedland et al., 2009) to *n*=2,671,298 participants (Hippisley-Cox et al., 2017)). Sixty-one studies (55%) assessed the risk of fatal or non-fatal CVD; 31 studies (28%) assessed the risk of T2D; five studies (5%)

assessed the risk of either prediabetes or T2D; three studies (3%) assessed the risk of metabolic syndrome or obesity; and three studies (3%) assessed the risk of stroke or transient ischaemic attack.

Lengths of predicted risks ranged from one (Noda et al., 2010) to 30 (Wickramasinghe et al., 2014, Pencina et al., 2009) years. The most common risk prediction timeframes were either ten-year risk (38 studies, 35%) or five-year risk (14 studies, 13%). Thirty-nine studies (35%) performed external validation of an original algorithm. Fourty studies (36%) performed internal validation by sub-setting the initial cohort or bootstrap methods. All algorithms were designed using either Cox Proportional Hazards or derivations of logistic regression analysis. Most studies selected variables for inclusion from previous research or clinical importance (50 studies, 45%) or using statistical methods, i.e., forward, or backward selection (31 studies, 28%). Seventeen studies (15%) used simple univariable analysis of each considered predictor, which is the least preferable since it cannot assess interactions between two or more variables. Eleven studies (10%) used machine learning techniques for variable selection.

6.4.1.3 Participant Characteristics

All studies were conducted in adults. The mean age of participants based on the 76 studies that reported mean age was 50.50 (SD 9.31) years. No studies included a mean age of participants below 35 years. Eighty-nine studies (81%) reported the sex distribution of the derivation cohort (mean 55.29% male (SD 17.27)), and 42 studies (38%) reported for the validation cohort (mean 52.25% male (SD 14.44)). The majority of studies included roughly equal sex distribution, apart from nine studies which included only (Paynter et al., 2009, Ridker et al., 2007) or mostly females (Yatsuya et al., 2016, Yatsuya et al., 2013, Abd El-Wahab et al., 2019, Choe et al., 2018, Park et al., 2009, Paynter et al., 2011, Ayala Solares et al., 2019), and 12 studies which included only (Ridker et al., 2008, Assmann et al., 2002, Brand, 1976, Dunder et al., 2004, Ferrario et al., 2005, L'Italien et al., 2000, Noda et al., 2010, Voss et al., 2002, Zhang et al., 2005) or mostly males (Wong et al., 2016, Nanri et al., 2015, Wickramasinghe et al., 2014). Thirty-three studies (30%) reported the ethnic makeup of their sample, where samples ranged from being ethnically completely homogenous in 18 studies (16%) to relatively heterogeneous, with less than 66% of participants falling into the most common ethnic group (Anderson et al., 2015, Robinson et al., 2011, Ha et al., 2018, Pylypchuk et al., 2018). See Appendix D Table 3 for detailed results on participant characteristics of included studies.

6.4.1.4 Algorithm Characteristics

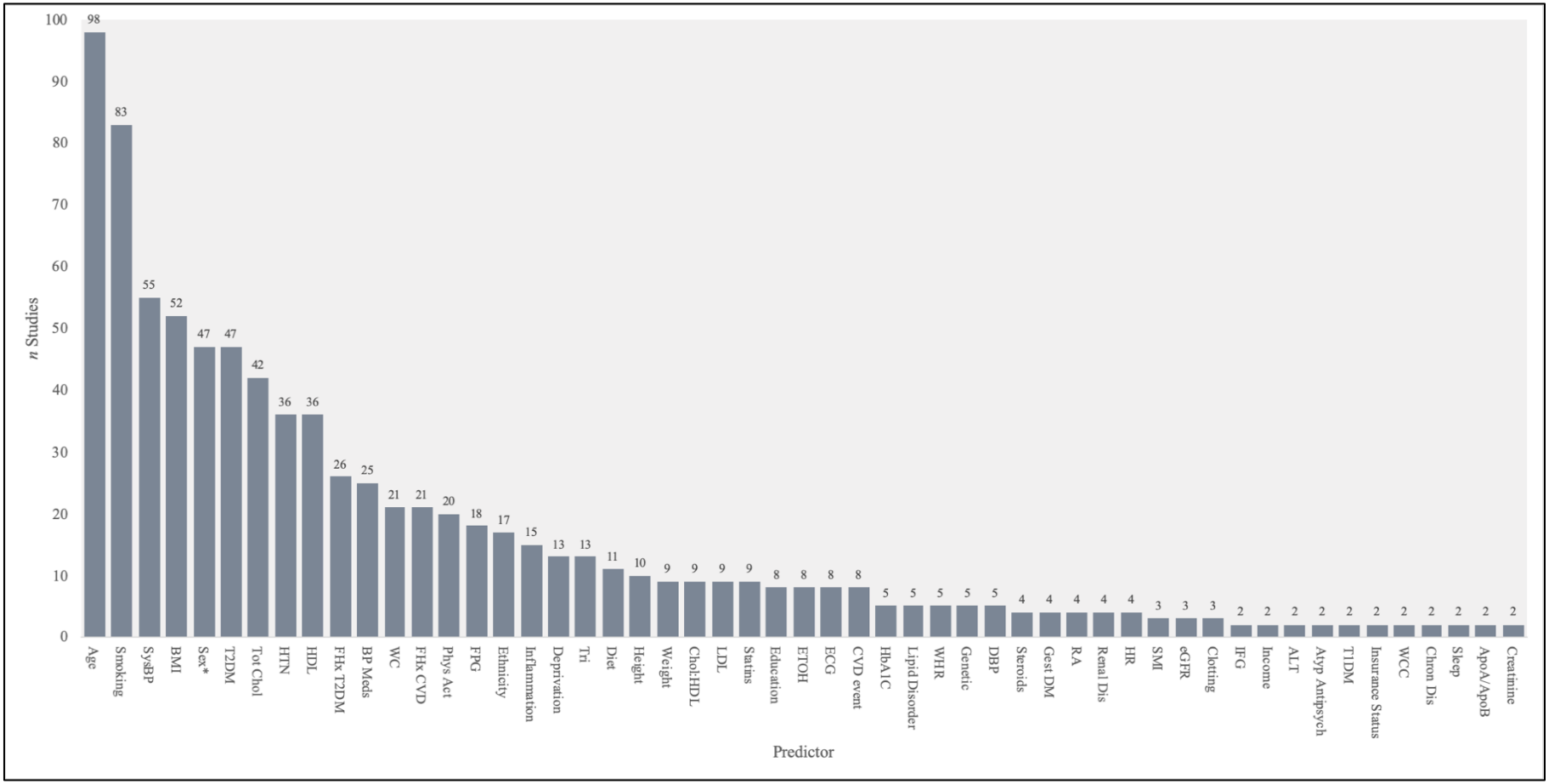
6.4.1.4.1 Predictors Included in Existing Algorithms

Figure 18 shows the frequency of different predictors included in studies. The most common predictors were age (98 studies, 89%), smoking (83 studies, 75%) and systolic blood pressure (55 studies, 50%). The number of predictors considered for each algorithm varied between four (Gao et al., 2010, Gao et al., 2009, Chen et al., 2009, Wen et al., 2017) to 473 predictors (Alaa et al., 2019). EPV varied between 2.1 (Griffin et al., 2000) and 5,075.4 (Hippisley-Cox et al., 2017). Twenty studies featured EPV ratios that were likely <10. See Appendix D Table 4 for a detailed description of algorithm characteristics of included studies.

6.4.1.4.2 Performance of Existing Algorithms

Discrimination statistics were presented in 93 studies (85%), and calibration statistics were presented in 62 studies (56%). From the 80 studies that included both model development and validation analysis, 35 (44%) reported performance statistics from both development and validation cohorts, 27 (34%) reported only validation cohort statistics, and ten (13%) reported development only statistics. Most commonly overall, studies reported both discrimination and calibration statistics (35 studies, 32%). Next most commonly, studies reported measures for discrimination, calibration, and sensitivity/specificity (23 studies, 21%). Eleven studies (10%) reported no model performance statistics. Discrimination was primarily assessed with the area under the curve (AUC / C-statistic). Reported C statistics ranged between 0.61 (Davies et al., 2010) to 0.97 (Park et al., 2009) though notably, the latter was at risk of model overfit, with a sample size of n=100 and an EPV ratio of 3.1. The mean C statistic across all included studies was 0.77, with 54 studies (49%) scoring above 0.70, suggesting 'good' discrimination. The majority of studies that reported calibration statistics used the Hosmer-Lemeshow goodness-of-fit chi^2 test. Seventeen studies (15%) used the preferred (Collins et al., 2015) method of calibration plots. See Appendix D Table 5 for a detailed description of algorithm performance of included studies.

Figure 18: Range and Frequency of Predictors Used In Algorithms Included In The Systematic Review



SysBP = Systolic Blood Pressure; T2DM = Type 2 Diabetes Mellitus; Tot Chol = Total Cholesterol; HTN = Diagnosis of Hypertension; HDL = High-Density Lipoprotein; FHx T2DM = Family history of Type 2 Diabetes; BP Meds = Prescribed Antihypertensive Medication; WC = Waist Circumference; FHx CVD = Family History Cardiovascular Diseases; Phys Act = Physical Activity; FPG = Fasting Plasma Glucose; Tri = Triglycerides; Chol:HDL = Cholesterol:HDL Ratio; LDL = Low-Density Lipoprotein; ETOH = Alcohol Use; ECG = Electrocardiogram Findings; CVD Event = Personal History of Cardiovascular Diseases; HbA1C = Glycated Haemoglobin; WHR = Waist:Hip Ratio; Genetic = Genotype Data; DBP = Diastolic Blood Pressure; Gest DM = Gestational Diabetes Mellitus; RA = Rheumatoid Arthritis; Renal Dis = Renal Disorders; HR = Heart Rate; SMI = Diagnosis of Serious Mental Illness; eGFR = Glomerular Filtration Rate; IFG = Impaired Fasting Glucose; ALT = Alanine Aminotransferase; Atyp Antipsych = Prescribed Antipsychotic Medication; T1DM = Type 1 Diabetes Mellitus; WCC = White Cell Count; Chron Dis = Personal History of Chronic Disease; ApoA/ApoB = Apolipoprotein A/B Levels; *not counted as a predictor in studies that developed sex-specific algorithms

6.4.1.5 Potential Applicability of Existing Cardiometabolic Risk Algorithms for Young People with Psychosis

Psychiatric disorders and treatment were taken into account in three studies (Osborn et al., 2015, Hippisley-Cox and Coupland, 2017, Hippisley-Cox et al., 2017) predicting risk of CVD (Hippisley-Cox et al., 2017, Osborn et al., 2015) or T2D (Hippisley-Cox and Coupland, 2017). Two of these studies (QRISK3 and QDiabetes) were conducted on large general-population samples, and one (PRIMROSE) was conducted in people with severe mental illness. QRISK3 and QDiabetes included a diagnosis of severe mental illness as a single predictor, whereas PRIMROSE included separate predictors for bipolar disorder and psychosis. QRISK3 and QDiabetes included the presence of any atypical antipsychotic as a predictor; PRIMOSE included first- or second-generation antipsychotics as separate predictors, along with antidepressants as another predictor. All three studies were conducted on middle- to older-aged adults (mean ages QDiabetes: 42.9 years QRISK3: 44.9 years, PRIMROSE: 49.5 years). In PRIMROSE, age was applied as a non-linear term with a log transformation and was weighted heavily compared with other risk factors. See Figure 19. In QRISK3 and QDiabetes, age was applied as a fractional polynomial, implying a non-linear impact on risk. QRISK3 and QDiabetes included interactions between age and other predictors, further amplifying the relative importance of age in the algorithms.

Figure 19: The Relative Weighting of Age vs Other Predictors in The PRIMROSE Algorithm

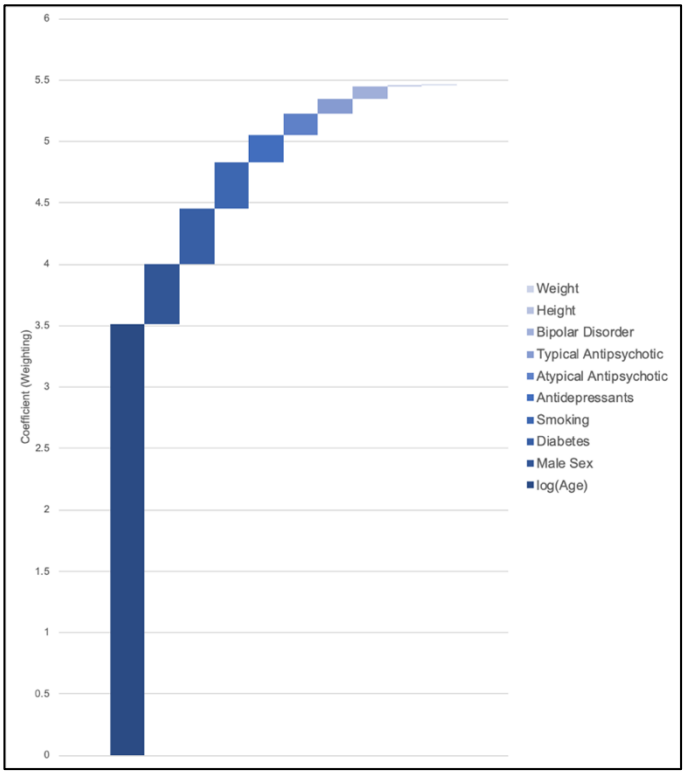


Figure illustrates the coefficients of predictors stacked upon one another cumulatively to show the relative weighting of age (presented at the bottom-left) compared with other predictors in the algorithm; QRISK3, QDiabetes and PRIMROSE were taken forward for the exploratory validation analysis, on the basis of: large samples used in development and validation; strong performance statistics; low risk of bias in three domains; and inclusion of psychiatric predictors / development in a psychiatric sample.

6.4.2 Exploratory Validation Analysis

6.4.2.1 Baseline Characteristics

The six-year observed risk of metabolic syndrome at age 24 years in the sample of participants with, or at risk of developing psychosis was 14.21% in males and 11.88% in females. In the sensitivity analysis (all available ALSPAC participants), the six-year observed risk was 7.54% for females and 5.76% for males. In the primary analysis, I included 3,030 person-years of observation. In the sensitivity analysis, I included 19,020 person-years of observation. Characteristics of included participants for both the primary and sensitivity analyses are presented in Table 31.

Table 31: Characteristics of Participants Included in Exploratory Validation Analysis

Characteristic (N, % unless stated)	Psychosis Risk Sample		Whole Sample Sensitivity Analysis	
	Females	Males	Females	Males
Number of Participants	323 (63.9)	182 (36.1)	1,909 (55.0)	1,561 (45.0)
Total person-years of observation	1,938	1,092	11,454	7,566
Ethnicity – White / Not-recorded	315 (97.5)	176 (96.7)	1,861 (97.5)	1,519 (97.3)
Systolic BP (mmHG), Mean (SD)	109.88 (8.28)	118.90 (9.67)	109.98 (7.98)	119.99 (9.09)
HDL (mmol/L), Mean (SD)	1.29 (0.36)	1.18 (0.33)	1.34 (0.31)	1.21 (0.24)
FPG (mmol/L), Mean (SD)	4.88 (0.36)	5.19 (0.66)	4.92 (0.49)	4.16 (0.24)
Total Cholesterol (mmol/L), Mean (SD)	3.86 (0.68)	3.55 (0.63)	3.94 (0.69)	3.56 (0.62)
Chol:HDL Ratio, Ratio SD	3.04 (0.85)	3.08 (0.85)	3.07 (1.01)	3.16 (0.96)
BMI (kg/m^2), Mean (SD)	23.75 (3.55)	23.62 (4.50)	23.06 (4.48)	22.14 (3.87)
FHx Cardiometabolic/Cardiovascular Disorders	194 (60.1)	117 (64.3)	603 (31.6)	448 (28.7)
Smoking (≥1 cigarette daily)	173 (53.6)	100 (54.9)	840 (44)	704 (45.1)
Depression	90 (27.9)	28 (15.4)	270 (14.1)	90 (5.7)
Alcohol Use	47 (15.4)	31 (16.7)	477 (6.5)	534 (6.9)
Antidepressant Medication	45 (14.7)	16 (8.6)	186 (2.5)	57 (0.7)
Antipsychotic Medication	48 (14.8)	29 (15.9)	13 (0.2)	6 (0.1)

BP=blood pressure; HDL=high-density lipoprotein; FPG=fasting plasma glucose; Chol=cholesterol; BMI=body mass index; FHx=family history

6.4.2.2 Primary Analysis – Psychosis Risk Sample

6.4.2.2.1 Discrimination

Discrimination C Statistics were: QDiabetes males C=0.75 (95% C.I., 0.72-0.78) and females C=0.78 (95% C.I., 0.73-0.84); QRISK3 males C=0.58 (95% C.I., 0.52-0.65) and females C=0.61 (95% C.I., 0.55-0.66); PRIMROSE C=0.73 (95% C.I., 0.70-0.78). After substituting participant ages to the mean

age of the original studies, C statistics mildly improved for each algorithm. Similarly, at age 18y, R^2 statistics were marginally higher in females than males in QDiabetes and QRISK3 and improved mildly after substituting participant ages to the mean age of the original studies. See Table 32.

Table 32: Discrimination Statistics for Algorithms Tested on ALSPAC Psychosis-Risk Sample at Age 18 Years and Mean Age of Original Study

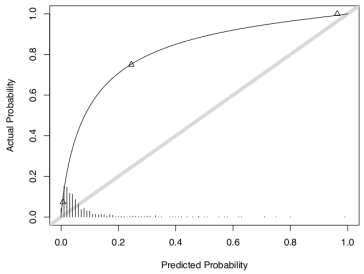
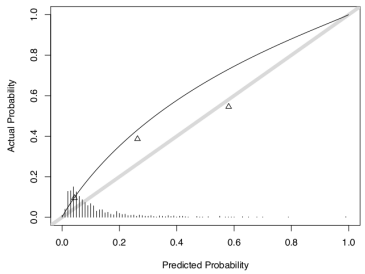
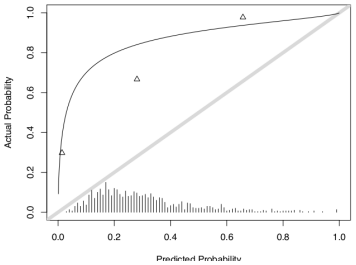
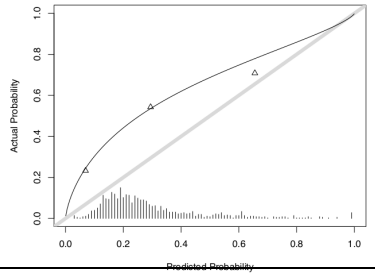
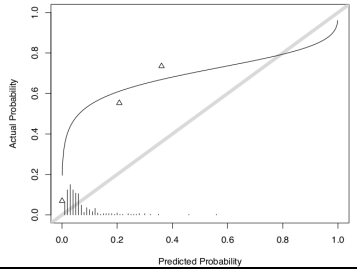
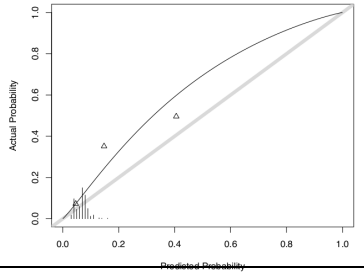
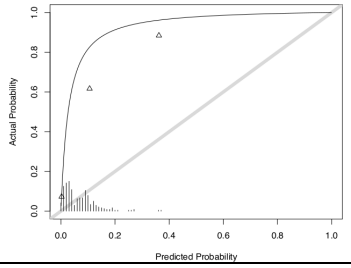
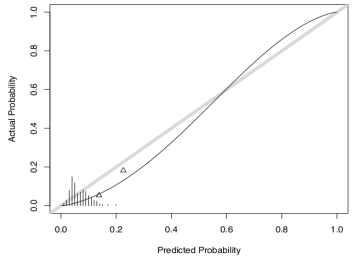
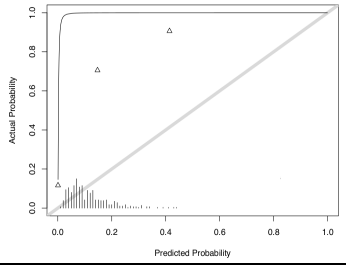
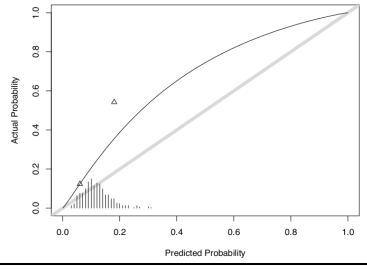
Algorithm	C Statistic (95% CI); R^2 Statistic			
	Age 18 Years		Mean Age Original Study	
	Male	Female	Male	Female
QDiabetes FPG	C=0.70 (0.65-0.74) R^2=0.13 (0.09-0.19)	C=0.78 (0.73-0.84) R^2=0.16 (0.10-0.24)	C=0.78 (0.75-0.80) R^2=0.21 (0.14-0.27)	C=0.83 (0.80-0.87) R^2=0.25 (0.19-0.31)
QRISK3	C=0.58 (0.52-0.65) R^2=0.09 (0.05-0.16)	C=0.61 (0.55-0.66) R^2=0.10 (0.03-0.18)	C=0.63 (0.58-0.69) R^2=0.11 (0.07-0.16)	C=0.66 (0.59-0.72) R^2=0.13 (0.05-0.20)
PRIMROSE Lipid	0.73 (0.70-0.78) R^2=0.13 (0.10-0.0.17)		0.75 (0.69-0.79) R^2=0.16 (0.12-0.22)	

FPG=fasting plasma glucose

6.4.2.2.2 Calibration

Calibration was poor across all three algorithms, with observed risk estimates consistently higher than predicted risk estimates, indicating a significant underprediction of risk. After substituting participant ages to the mean age of the original studies, calibration improved markedly in all three algorithms. See Figure 20.

Figure 20: Calibration Plots of Algorithms Tested on ALSPAC Psychosis-Risk Sample at Age 18 Years and at The Mean Age of Original Study

	Age 18 Years	Mean Age Original Study
QDiabetes (Female)		
QDiabetes (Male)		
QRisk3 (Female)		
QRisk3 (Male)		
PRIMROSE		

Perfect calibration (dark grey) would follow the diagonal (light grey) line, indicating perfect agreement between observed/expected risk. Grouped observations were split at each 0.2 of predicted risk.

6.4.2.3 Sensitivity Analysis – ALSPAC Whole Sample

6.4.2.3.1 Discrimination

QDiabetes and QRISK3 performed better in the whole sample than the psychosis-risk sample. PRIMROSE performed better in the psychosis-risk sample. Harrell's C Statistics were: QDiabetes males C=0.72 (95% C.I., 0.70-0.73) and females C=0.82 (95% C.I., 0.79-0.84); QRISK3 males C=0.64 (95% C.I., 0.62-0.66) and females C=0.62 (95% C.I., 0.59-0.65); PRIMROSE C=0.68 (95% C.I., 0.67-0.70) . Similarly, at age 18y, R^2 statistics were marginally higher in females than males in QDiabetes, but marginally higher in males in QRISK3. After substituting age to the mean age of the original studies, Harrell's C statistics and R^2 improved in all three algorithms. See Table 33.

Table 33: Discrimination Statistics for Algorithms Tested on ALSPAC Whole Sample at Age 18 Years and Mean Age of Original Study

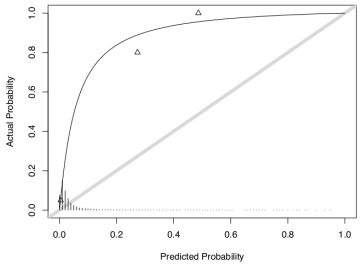
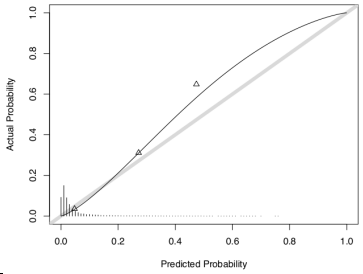
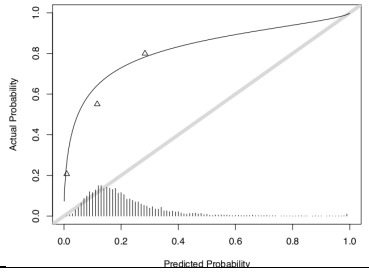
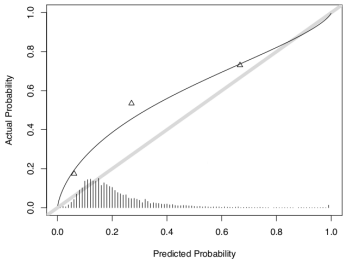
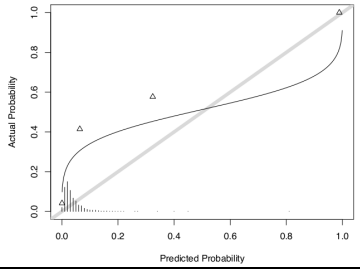
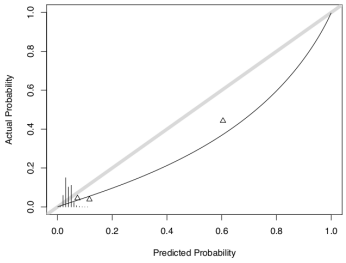
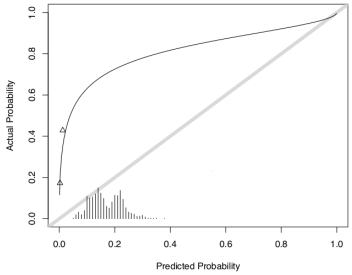
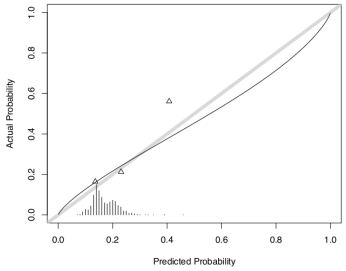
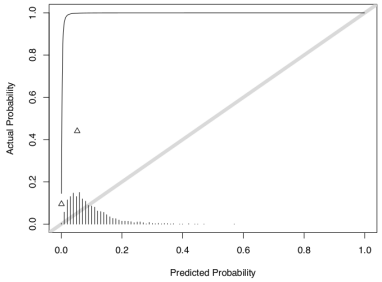
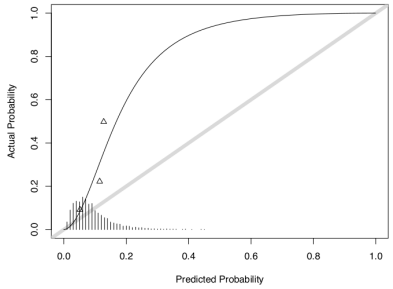
Algorithm	**C Statistic (95% CI); R^2 Statistic**			
	Age 18 Years		**Mean Age Original Study**	
	Male	**Female**	**Male**	**Female**
QDiabetes FPG	C=0.72 (0.70-0.73) R^2=0.14 (0.09-0.20)	C=0.82 (0.79-0.84) R^2=0.17 (0.10-0.26)	C=0.74 (0.72-0.77) R^2=0.19 (0.13-0.26)	C=0.81 (0.78-0.83) R^2=0.23 (0.17-0.28)
QRISK3	C=0.64 (0.62-0.65) R^2=0.11 (0.06-0.16)	C=0.62 (0.59-0.65) R^2=0.10 (0.06-0.15)	C=0.65 (0.64-0.67) R^2=0.11 (0.06-0.17)	C=0.72 (0.69-0.75) R^2=0.12 (0.07-0.18)
PRIMROSE Lipid	0.68 (0.67-0.70) R^2=0.11 (0.05-0.17)		0.68 (0.66-0.69) R^2=0.13 (0.07-0.19)	

FPG=fasting plasma glucose

6.4.2.3.2 Calibration

In a similar pattern to the psychosis sample, calibration was poor across all three algorithms with observed risk estimates consistently higher than predicted risk estimates, indicating a significant underprediction of risk. After substituting participant ages to the mean age of the original studies, calibration improved markedly in all three algorithms. See Figure 21.

Figure 21: Calibration Plots of Algorithms Tested on ALSPAC Whole Sample at Age 18 years and at Mean Age of Original Study

	Age 18 Years	Mean Age Original Study
QDiabetes (Female)		
QDiabetes (Male)		
QRisk3 (Female)		
QRisk3 (Male)		
PRIMROSE		

Perfect calibration (dark grey) would follow the diagonal (light grey) line, indicating perfect agreement between observed/expected risk. Grouped observations were split at each 0.2 of predicted risk.

6.5 Discussion

I performed a systematic review of cardiometabolic risk prediction algorithms developed either for the general or psychiatric populations and considered their potential suitability for young people with psychosis. I also used data from a sample of relatively young adults to first explore whether existing cardiometabolic risk prediction algorithms may be suitable for young people with or at risk of psychosis, and secondly to examine the impact of how age is weighted in existing cardiometabolic risk prediction algorithms.

Regarding the systematic review, I identified a substantial number of cardiometabolic risk prediction algorithms, yet most have not been integrated into clinical practice. Only one included algorithm (PRIMROSE) was developed in a population of people with severe mental illness (Osborn et al., 2015). Two (QRISK3, QDiabetes) were developed in the general population and included psychiatric predictors (Hippisley-Cox et al., 2017, Hippisley-Cox and Coupland, 2017).

All included algorithms were developed in samples of middle- to older-age adults. One might traditionally consider this proportionate since cardiometabolic disorders are traditionally regarded as diseases of advancing age. Yet, cardiometabolic risk still exists in the absence of advancing age. Even in the general population, there is an increasing prevalence of early-onset T2D (Wilmot and Idris, 2014) and childhood obesity (Skinner et al., 2016), likely related to the shift toward a more sedentary lifestyle and unhealthy diet in recent decades. The absence of an algorithm developed for younger populations is an important finding since early intervention may reduce the risk of young people forming part of a future generation of patients with chronic CVD (Chrysant, 2011). This finding suggests the need for either new or recalibrated versions of cardiometabolic risk algorithms tailored for younger generations.

Primary prevention is the best means to address the personal and societal burden attributed to T2D, CVD and its associated morbidity and mortality (Weintraub et al., 2011). While this message is important for the general population, it is crucial for young people with psychosis who are at a substantially higher risk of precipitant cardiometabolic disorders. This population may be more likely to smoke (Sagud et al., 2018), exercise less (Heald et al., 2017), and eat a more unhealthy diet (Heald et al., 2017) than their peers, and yet may also be prescribed medication that in itself can adversely and severely impact cardiometabolic indices (Leucht et al., 2013). Further, they may be faced with inappropriate barriers to accessing healthcare (Lawrence and Kisely, 2010), diagnostic overshadowing (Jones et al., 2008), and may have an intrinsic biological propensity for altered cardiometabolic function (Perry et al., 2018). Meta-analyses featuring mostly antipsychotic-naïve young people with first-episode psychosis have consistently reported an increased incidence of

insulin resistance, impaired glucose tolerance (Perry et al., 2016, Pillinger et al., 2017a) and dyslipidaemia (Pillinger et al., 2017b, Perry et al., 2016, Misiak et al., 2017) compared with matched controls from the general population, after adjusting for anthropometric and sociodemographic factors. Each are predeterminants of cardiometabolic disorders such as T2D and obesity. Existing algorithms may not adequately capture these factors. Additionally, meta-analyses of cross-sectional studies suggest that psychosis is associated with higher levels of circulating inflammatory markers (Upthegrove et al., 2014, Miller et al., 2011, Miller et al., 2014, Fernandes et al., 2016). Evidence from longitudinal studies suggests an association between inflammatory markers at baseline and psychosis at follow-up (Khandaker et al., 2014, Metcalf et al., 2017, Goldsmith et al., 2019). Inflammatory states are also associated with cardiometabolic disorders (Rethorst et al., 2014, Monteiro and Azevedo, 2010, Hermsdorff et al., 2011, Calabro and Yeh, 2008). While 15 relatively newer algorithms from the systematic review did include inflammatory predictors, none also included psychiatric predictors.

Each of the three algorithms that did include psychiatric factors featured an antipsychotic-related predictor. Antipsychotic associated weight gain can occur relatively quickly after initiation (Spertus et al., 2018) and is associated with altered eating behaviours (Sentissi et al., 2009) and sedentariness (Vancampfort et al., 2017). However, whilst there are some efficacy differences between antipsychotics, these are gradual rather than discrete (Huhn et al., 2019). Differences in side-effects are more marked, and each has an inherently different impact upon cardiometabolic risk (Vancampfort et al., 2015). This may be explained by differing affinities to receptors other than the dopamine-2 (D2) receptor, for example, the histamine-1 (H1) receptor, serotonin-2c (5-HT2c) and adrenergic receptors (a2 and b3) (Starrenburg and Bogers, 2009), which may have a role in the regulation of food intake (Kroeze et al., 2003). The varied impact upon cardiometabolic risk by different antipsychotics does not abide by the traditional distinctions of either typical/atypical or first/second generation, which were the binary distinctions of the included algorithms. A more appropriate antipsychotic predictor may instead model antipsychotics based on their relative cardiometabolic risk.

I used the PROBAST tool (Wolff et al., 2019) to examine the risk of bias of included studies in the systematic review. Only two studies were rated as low risk of bias, with all others rated as either unclear or high risk of bias. This may reflect the relatively recent introduction of the TRIPOD guidelines for prediction model studies (Collins et al., 2015). Nevertheless, the results suggest that the results and therefore clinical validity of most included studies should be accepted with caution.

The EPV ratio also varied widely between studies. A low EPV ratio can be an indicator of model-overfit (Pavlou et al., 2015), which can bias results. I identified 20 studies with an EPV ratio of likely

<10, and therefore the performance reported in those studies should be interpreted with caution. Finally, it is striking that whilst many included studies promoted the use of their algorithms in clinical practice, there appears to have been relatively little follow-up to assess either clinical or economic impact. A notable exception was PRIMROSE (Osborn et al., 2015), which was the only algorithm developed and validated on a sample of people with mental illness. A cost-effectiveness analysis (Zomer et al., 2017) found it improved quality of life and reduced healthcare-related costs compared to using no algorithm.

A previously published systematic review (Damen et al., 2016) examining cardiovascular risk prediction algorithms in the general population also identified an abundance of studies. The review similarly concluded the methodological shortcomings of most risk prediction algorithms likely limit their suitability for clinical practice. The previous review differs from this review since it aimed to identify algorithms and assess their suitability for young people with psychosis. Therefore, I did not include studies reporting new validations in a similar population to already validated algorithms. The previous review also presented sex-stratified algorithms as distinct entities, increasing the apparent number of algorithms they reported. For simplicity and in consideration of the overarching research question, I did not take this step. Finally, many new algorithms have been developed since the previous review, which I was able to include in this review.

Regarding the exploratory validation analysis, I considered three algorithms for this step; QRISK3, QDiabetes and PRIMROSE. These were selected due to the large sample sizes in model development and validation, favourable model performance statistics, relatively low risk of bias, and the inclusion of psychiatric predictors/development in a psychiatric population.

I found that discrimination statistics were relatively good at age 18 years for QDiabetes and PRIMROSE and improved further when substituting to the mean age of original studies. This means that QDiabetes and PRIMROSE could predict higher risks in 'cases' than 'non-cases', even in relatively young adults. This did not apply to QRISK3, particularly in males, where the algorithm was little better than chance at discriminating higher and lower cardiometabolic risk in young adults with or at risk of developing psychosis.

For all three algorithms included in the validation analysis, the discriminative ability was attenuated compared with the original published studies (Hippisley-Cox and Coupland, 2017, Hippisley-Cox et al., 2017, Osborn et al., 2015). This may be because the present analysis included younger participants than the original studies. For example, both QRISK3 and QDiabetes were developed and validated in participants aged 25 and over, and PRIMROSE was developed and validated in participants aged 30 and over. QRISK3 and QDiabetes define a minimum age of 25 when using their online calculators,

although PRIMROSE sets a minimum of age 18 years. Additionally, in the primary analysis, I tested a sample of participants with or at risk of developing psychosis, whereas QDiabetes and QRISK3 were designed for use in the general population.

Furthermore, I tested a different outcome compared with the original algorithms. I tested metabolic syndrome since it is an established precursor of both T2D and CVD (Wilson et al., 2005, Shin et al., 2013) and is a more suitable outcome for younger populations. Nevertheless, the improvement in discrimination statistics after substituting age provides some face validity to the choice of outcome.

However, discriminative ability is only half the story regarding predictive performance since discrimination statistics cannot assess the accuracy of the amount of risk apportioned by a model; this represents a test of *absolute risk* estimates and is examined with a measure of calibration. The calibration plots showed that observed risk was systematically greater than predicted risk in all models, indicating substantial underprediction of risk in younger participants. Calibration plots improved markedly in all algorithms when age was artificially increased to the mean age of the original studies. This suggests that the manner with which age is modelled in current algorithms is a major limiting factor in applying them to younger populations. This is likely because many cardiometabolic risk factors are cumulative over time (Reinikainen et al., 2015), thus, age becomes an increasingly important contributor to cardiometabolic risk as one gets older. This notion is elegantly painted by all three algorithms, which modelled age as either a non-linear function, included interactions between age and other predictors, or both.

Strengths of this systematic review include following PRISMA reporting guidelines (Moher et al., 2009) and the ability to complement the findings with an exploratory validation analysis using data from a large birth cohort of young adults. I was able to test three validated cardiometabolic risk prediction algorithms which are commonly used in clinical practice in the UK, on a different population who are in clear and crucial need of a suitable tool.

Limitations of the study first and foremost relate to the exploratory validation analysis. The three algorithms I tested were not designed for use in young adults, though this in itself should not be a barrier to explore potential suitability in a different population. Nevertheless, the results should not be seen to cast doubt on the predictive ability of such algorithms when applied to the populations they were developed for. I could not include every predictor from the algorithms I tested, which may have impacted performance statistics. That said, the impact of this limitation on the results is unlikely to have been uniform for each predictor I could not include. For example, even if data were available, it is doubtful that many participants in the relatively young cohort would have diagnosed CVD or chronic kidney disease, a history of gestational diabetes, or be prescribed statins. Also, the measured

outcome differed from the outcome of the algorithms I tested. While three algorithms included in the systematic review did aim to predict risk of metabolic syndrome, I did not consider them for the exploratory validation analysis. This is because they did not include psychiatric predictors; were at relatively high risk of bias; and study authors did not publish their fully specified algorithm equations. Nevertheless, metabolic syndrome is a precursor of T2D (Shin et al., 2013) and CVD (Wilson et al., 2005), and the relatively good performance of the algorithm when I artificially substituted age to the mean age of the original study suggests face validity of the metabolic syndrome outcome.

Other limitations relate to the systematic review. I was unable to follow a meta-analytic approach to the synthesis of results due to study heterogeneity. The lack of a meta-analytic approach meant I could not examine the risk of publication bias, which may have played a part in the configuration of studies included since only three studies were not published in peer-reviewed journals.

In conclusion, young people with psychosis are at higher risk of developing cardiometabolic disorders than the general population. A suitable cardiometabolic risk prediction algorithm for this population would be highly beneficial for healthcare professionals to help them tailor treatment plans to reduce long-term physical and psychiatric morbidity. Existing cardiometabolic risk algorithms cannot be recommended for this purpose since they likely underestimate the cardiometabolic risk of all young people, let alone a group already at significantly higher risk than the general population. Existing algorithms require recalibration to suit younger populations, and, better still, a new cardiometabolic risk prediction algorithm is required which is specifically developed for young people with psychosis. A well-designed algorithm may include a more appropriate distinction of metabolically-active antipsychotics; should more appropriately weight the predictors for the specific characteristics of young people with psychosis; and may include a more age-appropriate outcome, such as metabolic syndrome. Further, particular attention should be paid to patient acceptability to ensure the algorithm is used in clinical practice rather than simply buried in a research database.

Chapter 7

Development and External Validation of The Psychosis Metabolic Risk Calculator (PsyMetRiC): A Cardiometabolic Risk Prediction Algorithm for Young People with Psychosis

7.1 Introduction

Young people with psychosis are at significantly higher cardiometabolic risk than the general population. Insulin resistance and dyslipidaemia are detectable from the onset of psychosis in relatively young patients (Perry et al., 2016, Pillinger et al., 2017b) and, left unchecked, contribute to a higher risk of more chronic conditions such as T2D and CVD (Firth et al., 2019), and a shortened life expectancy of up to 15 years (Plana-Ripoll et al., 2019). Since some treatments for psychosis can exacerbate cardiometabolic risk (e.g., certain antipsychotic medications), young patients who are most at risk of adverse cardiometabolic outcomes must be identified at the outset. Then, interventions can be tailored to reduce the risk of longer-term cardiovascular morbidity/mortality.

Prognostic risk prediction algorithms are a valuable means to encourage personalised, informed healthcare decisions. In the general population, cardiometabolic risk prediction algorithms such as QRISK3 (Hippisley-Cox et al., 2017) are commonly used to predict CVD risk from baseline demographic, lifestyle, and clinical information to identify higher-risk individuals for tailored interventions. In Chapter 6, I performed a systematic review of cardiometabolic risk prediction algorithms developed in for the general or psychiatric populations. I found that all algorithms were developed in samples of comparatively older adults and most didn't include relevant predictors such as antipsychotic medication. In the accompanying exploratory validation analysis, I found that existing algorithms significantly underpredict cardiometabolic risk in young people with or at risk of developing psychosis. Therefore, I concluded that no existing algorithm is likely to be suitable for young people with psychosis. See Chapter 6.

7.2 Aims and Objectives

Therefore, I aimed to develop and externally validate the Psychosis Metabolic Risk Calculator (PsyMetRiC) to predict up to six-year risk of metabolic syndrome, an age-appropriate precursor of CVD and early mortality, in young people with psychosis. I aimed to prioritise clinical usefulness and patient acceptability via input from a young person's advisory group and by developing two PsyMetRiC versions, one with and one without biochemical results. I followed TRIPOD reporting guidelines (Collins et al., 2015). See Appendix D Table 6 for the completed reporting guidelines.

7.3 Methods

7.3.1 Data Sources

7.3.1.1 Algorithm Development in an EIS Patient Sample

I developed PsyMetRiC in pooled retrospective data from patients aged 16-35 years enrolled in Birmingham EIS (sample frame *n*=391) or Cambridgeshire and Peterborough NHS Foundation Trust EIS (CAMEO) (sample frame *n*=1,113). Anonymised Birmingham data were collected between 2014-2018 as part of the National Clinical Audit of Psychosis Quality Improvement program, enhanced locally with medication data, conforming to HRA definition of Service Evaluation, confirmed by Birmingham Women's and Children's Hospital NHS Foundation Trust. CAMEO data were collected by conducting an anonymised search of EIS patients enrolled since 2013 using the Clinical Records Anonymisation and Text Extraction (CRATE) tool (Cardinal, 2017) (NHS National Research Ethics Service references 12/EE/0407; 17/EE/0442). Assistance in accessing and processing the data was provided by Professor Rachel Upthegrove (University of Birmingham) and Dr Emanuele Osimo (Imperial College London). Predictors were assessed at the closest point (+/- 100 days) to EIS enrolment, and outcomes were assessed up to six years later. I excluded patients who: had <1 year follow-up; had the outcome at baseline; or had missing data on all predictor or outcome variables, resulting in a final sample of *n*=651. See Table 34. See Appendix D Table 8 for a missing sample analysis for the pooled development sample.

7.3.1.2 External Validation in in EIS Patient Sample

I used the Clinical Records Interactive Search (CRIS) resource to capture anonymised data from South London and Maudsley NHS Foundation Trust EIS (SLaM) (NIHR Biomedical Research Centre CRIS Oversight Committee reference: 20-005)). The sample frame included 2,985 EIS patients aged 16-35 years enrolled since 2012. Assistance in accessing and processing this data was provided by Dr Emanuele Osimo (Imperial College London). Predictors and outcomes were assessed in the same manner as described above. I excluded participants as described above, resulting in a final sample of *n*=510. See Table 34. Please see Appendix D Table 9 for a detailed analysis of the missing sample for the validation sample.

7.3.1.3 External Validation Sensitivity Analysis in a General Population Sample

I examined the performance of PsyMetRiC in young adults who had or were at risk of developing psychosis using ALSPAC data. See Section 2.3.1 for a full description of the ALSPAC cohort. The

sample frame included 527 participants identified as having experienced definite psychotic symptoms at either age 18 or 24 years, assessed via the semi-structured Psychosis-Like Symptom Interview. See Section 2.3.3 and Section 3.3.2 for detail on assessment of PEs and psychotic disorder at age 18 and 24 years. Predictors were assessed at age 18 years, and the outcome was assessed at age 24 years. I excluded participants as described above, resulting in a final sample of *n*=505. See Table 34. ALSPAC Ethics and Law Committee and Local Research Ethics Committees provided ethical approval. Informed consent was obtained from participants following the recommendations of the ALSPAC Ethics and Law Committee at the time.

Table 34: Predictor Comparisons between Samples Used in Algorithm Development and Internal/External Validation

Predictor	**Sample**				
	Development			**External Validation / Sensitivity Analysis**	
	Birmingham EIS	**CAMEO EIS**	**Pooled Development Sample**	**SLaM EIS Validation Sample**	**ALSPAC Risk of Psychosis**
Final Included Sample size, *N*.	352	299	651	510	505
Age in Years, mean (SD)	23.76 (4.90)	25.42 (4.77)	24.52 (4.91)	24.45 (4.75)	17.81 (0.43)
White/NA Ethnicity, *N*. (%)	110 (31.25)	250 (83.61)	360 (55.30)	154 (30.20)	491 (97.2)
Black/African-Caribbean Ethnicity, *N*. (%)	94 (26.70)	15 (5.01)	109 (16.74)	250 (49.02)	<5 (<1.00)[a]
Asian Ethnicity, *N*. (%)	147 (41.76)	34 (11.37)	181 (27.80)	106 (20.78)	<5 (<1.00)[a]
Male Sex, *N*. (%)	232 (65.90)	208 (69.57)	440 (67.59)	351 (68.82)	182 (36.1)
HDL, mmol/L, mean (SD)	1.76 (0.35)	2.08 (0.49)	1.88 (0.57)	1.57 (0.37)	1.21 (0.31)
Triglycerides, mmol/L, mean (SD)	1.46 (1.18)	1.30 (0.89)	1.39 (1.06)	1.23 (0.71)	1.06 (0.77)
BMI, kg/m^2, mean (SD)	22.06 (5.13)	24.01 (5.73)	23.63 (5.43)	22.96 (6.94)	23.68 (3.55)
FPG (mmol/L), mean (SD)	5.20 (1.02)	5.17 (1.45)	5.19 (1.28)	5.03 (1.10)	5.01 (0.49)
Systolic BP (mmHg), mean (SD)	121.18 (11.04)	119.88 (12.25)	120.65 (11.68)	119.96 (13.70)	115.10 (11.88)
Metabolically-Active Antipsychotics[b], *N*. (%)	239 (67.90)	216 (72.24)	455 (69.89)	472 (92.55)	58 (11.26)
Smoking, *N*. (%)	182 (51.70)	133 (44.48)	315 (48.39)	469 (91.96)	273 (54.05)
Follow-up time, years, mean (SD)	2.44 (1.54)	1.43 (1.03)	1.86 (1.32)	2.73 (1.76)	5.18 (0.39)
Mean time of predictor assessment from EIS enrolment, mean days (SD)	23.55 (25.44)	21.93 (29.84)	16.71 (26.38)	3.05 (36.01)	*
Metabolic Syndrome at baseline, *N*. (%[c])	31 (7.90)	18 (5.11)	49 (6.58)	30 (5.64)	19 (4.17)
Metabolic Syndrome at Follow-up, *N*. (%)	74 (21.04)	35 (11.71)	109 (16.74)	86 (16.86)	76 (14.75)

HDL=high-density lipoprotein; BMI=body mass index; FPG=fasting plasma glucose; BP=blood pressure; ALSPAC=Avon Longitudinal Study of Parents and Children; SLaM=South London and Maudsley NHS Foundation Trust; EIS=Early Intervention Service; CAMEO=Cambridgeshire and Peterborough Foundation NHS Trust; [a]Reported as <5 due to ALSPAC reporting guidelines; [b]Metabolically-active antipsychotics are listed in Table 36; [c]Corresponds to percentage of sample before exclusion; *Health record / service use data is not currently available in ALSPAC

7.3.2 Outcome

I used the harmonized definition (Alberti et al., 2009) of the metabolic syndrome as a binary outcome. See Table 2.

7.3.3 Predictor Variables

7.3.3.1 Method of Predictor Selection

Predictors were included on a balance of clinical knowledge, prior research, and likely clinical usefulness/patient acceptability after discussion of the work with the McPin Foundation Young Persons Advisory Group (YPAG), comprising volunteers aged <24years with lived experience of mental health difficulties. I attended three meetings of the YPAG to discuss and shape the work as it progressed. Please see Table 35 for quotes and comments from the YPAG regarding PsyMetRiC.

7.3.3.2 Rationale and Coding of Predictors Selected for Inclusion In PsyMetRiC

7.3.3.2.1 Age

Age is frequently included in existing cardiometabolic risk prediction algorithms (see Chapter 6), and I also included it in PsyMetRiC as a continuous variable. Whilst some previous large-scale general population risk-prediction algorithms have considered age either as a non-linear term or as an interaction term with other predictors (see Chapter 6), I did not take this step to limit potential model complexity and thus reduce the risk of model-overfit given the available sample size. Considering age as an interaction term with other predictors would have added the requirement for a variable selection technique such as backward selection or more automatic penalized methods such as lasso regression with nested cross-validation. Given the sample size available, I chose not to proceed with such methods since they increase the risk of model overfit in smaller samples compared with forced-entry (Subramanian and Simon, 2013, Harrell, 2001, Steyerberg et al., 2000), and thus may have hampered external validation performance (Lever, 2016).

Table 35: Comments From McPin Young Person's Advisory Group (YPAG)

Question Asked To The YPAG	Responses From The YPAG
"Does it surprise you that despite many calculators for diabetes/obesity[a] have been made, none of them have been made for younger people? What do you think about that?"	It is quite worrying because there is strong research evidence that these conditions can develop in young people who have emerging mental health problems. Could be prevented if such a scale was made to lower risk of health issues in later life.
	The calculator could help bring awareness to doctors and young people about the risk.
	Because of the link found with mental health issues which affect all ages, it is important that this calculator is being made.
"On a scale of 1 (not important at all) to 10 (really important), how important do you think it is to know your chance of getting diabetes /obesity[a] in the next 6 years? Why/why not?"	**9** - Because it could help people to make changes to their lifestyle that would prevent them from getting these diseases in the future which would help them to live a longer life. The only reason I didn't put 10 is that some people may not want to know if they are destined to get a disease, even if this is not true, it may not be helpful to some people.
	5 - It's useful because some people will want to make changes such as exercise more or sleep more to prevent getting these conditions. However, some may find these pointless and counterproductive as the calculator works only by chance.
	9 – more likely to make those changes if they receive this information
From the information that is asked by the calculator, how happy do you think a young person would be to give that information to a doctor today?	Most people won't have a problem with sharing their height however a lot of people might be uncomfortable sharing their weight because they are unhappy with it
	I don't think that anyone would have a problem sharing this information *[smoking]* unless they are ashamed of how much they smoke
	If there was an option not to have a blood test, it's likely that not many people would opt out
	Weight & sex are quite sensitive subjects

[a]The phrase diabetes/obesity was used in place of metabolic syndrome at YPAG meetings since the former terms are more commonly used in common parlance, and thus more widely understood by non-healthcare professionals.

7.3.3.2.2 Ethnicity

Ethnicity is one of the most frequently included predictors in existing cardiometabolic risk prediction algorithms (see Chapter 6), and I included it in PsyMetRiC. Non-White ethnicity is an important risk factor for metabolic syndrome (Deboer, 2011) and predicts antipsychotic-induced metabolic dysfunction (Pillinger et al., 2020). In the development and validation samples, ethnicity was recorded inconsistently, with the majority of included records classified in relatively simple terms, for example "White" or "Asian". However, these simplified classifications do not recognise the heterogeneity within these groupings, therefore potentially incorrectly inferring that the populations are homogeneous (Lear and Gasevic, 2019). Nevertheless, to strike an appropriate balance between the

available sample size, the case-mix of the development and validation samples, and with a consideration to maximise coding harmonisation between datasets, I proceeded with a categorical nominal variable with as much granularity as the data permitted, and so the variable consisted of White European/not stated (reference category), Black/African-Caribbean ethnicity, and Asian/Other ethnicity.

7.3.3.2.3 Sex

Sex is frequently considered in cardiometabolic risk prediction algorithms, either as a predictor or a stratification variable (see Chapter 6). There are notable sex differences in the epidemiology, aetiology, biology and clinical expression of metabolic syndrome (Pradhan, 2014). For example, before the menopause, increased adiposity is more commonly precipitated in females than males (Kuk and Ardern, 2010), whereas hypertension and disrupted biochemical indices are more common in males (Kim and Reaven, 2013), possibly due to a metabolically-active effect of oestrogen (Gupte et al., 2015). Longer-term cardiovascular outcomes such as CVD affect both sexes but show differences in presentation and clinical course (Beale et al., 2018). Recent meta-analytic reports have suggested that male sex is an important risk factor for antipsychotic-induced biochemical disruption (Pillinger et al., 2020). Considering the available sample size, I did not consider separate algorithms for males and females and chose to model sex as a binary variable.

7.3.3.2.4 Body Mass Index

BMI is frequently included in cardiometabolic risk prediction algorithms (see Chapter 6), and overweight/obesity is a reliable predictor of adverse cardiometabolic and cardiovascular outcomes (Van Gaal et al., 2006). Weight gain is also a common side-effect of certain antipsychotic medications (Leucht et al., 2013) and can precipitate relatively quickly after initiation (Spertus et al., 2018). While BMI may be less accurate at classifying adiposity than laboratory or research-based measures such as dual-energy x-ray absorptiometry or bio-impedance analysis (Shah and Braverman, 2012), it is commonly recorded in clinical practice and correlates well with other measures of obesity (Barreira et al., 2011). Therefore, I included BMI as a continuous variable. I did not consider interactions of BMI with other predictors (including but not limited to, for example, antipsychotic medication) to limit model complexity and thus reduce the risk of model overfit in the available sample.

7.3.3.2.5 Smoking

Smoking is frequently included in cardiometabolic risk prediction algorithms (see Chapter 6) and is strongly associated with adverse cardiometabolic and cardiovascular outcomes (Banks et al., 2019). The impact of smoking on cardiometabolic and cardiovascular risk is dose-dependent, yet, in previous large-scale general population algorithms developed for older adult populations, smoking is usually classified as a categorical variable including 'current smoker', 'ex-smoker' and 'never-smoked'. The lack of consideration of dosage in previous algorithms (i.e., the number of cigarettes smoked per day and for how long) is likely due to the highly variable reporting of smoking history in electronic health record datasets (Polubriaginof et al., 2017). However, whilst a prolonged smoking history increases cardiometabolic and cardiovascular risk compared with 'never smoked' (Duncan et al., 2019), particularly in older adults (Mons et al., 2015), some research suggests that smoking cessation in young people can reduce cardiometabolic and cardiovascular risk to baseline in as little as five years (Lloyd-Jones et al., 2017). This is relevant since PsyMetRiC was developed for younger populations. Therefore, for this reason, and to assist in harmonisation across the development and validation datasets, I included smoking as a binary variable (yes/no). For the SLaM external validation sample, smoking status was derived using the 'CRIS-IE-Smoking' application, which sits within the General Architecture for Text Engineering (GATE) natural language processing software to extract smoking status information from open-text fields (Wu et al., 2013). For all other samples, smoking was captured as current smoking status from clinical interview.

7.3.3.2.6 Prescription of a Metabolically-Active Antipsychotic

Antipsychotic medication is an important contributor to cardiometabolic risk in young people with psychosis, and so it was crucial to include in PsyMetRiC. Antipsychotic medications are rarely included in existing cardiometabolic risk prediction algorithms. Three more recent algorithms (QRISK3, QDiabetes, PRIMROSE) have included antipsychotics as predictors, grouped as binary variables based on the traditional distinctions of typical/atypical or first/second-generation. See Chapter 6. However, the differential cardiometabolic effects of antipsychotics do not necessarily abide by these distinctions.

Therefore, I instead grouped antipsychotics based on existing evidence (Leucht et al., 2013, Pillinger et al., 2020) as 'metabolically-active' or not (Table 36). This is a notable advance over previous risk prediction algorithms. Therefore, I classified all individuals prescribed a metabolically-active antipsychotic as "1" and all participants who were not prescribed a metabolically-active antipsychotic (including participants who were not prescribed any antipsychotic) as "0". However, I could not consider dosage or a more granular categorical antipsychotic medication variable for several reasons.

First, interactions of dosage with antipsychotic choice would have added significant complexity to the model and may have increased the risk of overfit, given the available sample size. It would also have been challenging to capture the effect of dosage change on cardiometabolic risk from a single baseline measure of predictor assessment. This is important because antipsychotics are usually commenced at a low dose and upwardly titrated over time, depending on treatment response. Second, with increasing numbers of risk-distinguishing categories comes increased subjectivity of group classification for some antipsychotics. In future, when development and validation samples of young people with psychosis are large enough, it would be most appropriate to model the cardiometabolic risk associated with each antipsychotic medication individually.

Table 36: Classification of Metabolically-Active Antipsychotics

More Metabolically Active Antipsychotics	Less Metabolically Active Antipsychotics
Olanzapine (Leucht et al., 2013)*	Aripiprazole (Leucht et al., 2013)*
Quetiapine (Leucht et al., 2013)*	Amisulpiride (Leucht et al., 2013)*
Risperidone (Leucht et al., 2013)*	Haloperidol (Leucht et al., 2013)
Paliperidone (Leucht et al., 2013)	Sulpiride (Bak et al., 2014)
Clozapine (Leucht et al., 2013)	Periciazine (Matar et al., 2014)†
Chlorpromazine (Leucht et al., 2013)	Lurasidone (Leucht et al., 2013)†
Asenapine (Pillinger et al., 2020)†	Ziprasidone (Leucht et al., 2013)†
Pimozide (Bak et al., 2014)†	Flupentixol (Pillinger et al., 2020)†
Levomepromazine (Bak et al., 2014)†	Fluphenazine (Pillinger et al., 2020)†
Prochlorperazine (Leucht et al., 2013)†	Zuclopenthixol (Bak et al., 2014)†
Trifluoperazine (Alonso-Pedrero et al., 2019)†	
Pipotiazine (Alonso-Pedrero et al., 2019)†	

This table comprises all antipsychotics prescribed for participants/patients in all samples; *indicates the five most commonly prescribed antipsychotics across all samples; †indicates antipsychotics rarely prescribed (<3 participants/patients in total across all samples)

7.3.3.2.7 Blood-based Predictors: HDL and Triglycerides

Blood-based predictors feature less often in cardiometabolic risk prediction algorithms (see Chapter 6). However, meta-analytic evidence suggests abnormal triglyceride and HDL levels are detectable at FEP (Misiak et al., 2017), even in individuals with limited exposure to antipsychotic medication.

A raised triglyceride:HDL ratio is a hall-mark of insulin resistance (Murguia-Romero et al., 2013), which is also associated with antipsychotic-naïve FEP (Perry et al., 2016), whereas meta-analytic evidence suggests that other measures of glucose-insulin homeostasis (e.g. FPG, HbA1C) are not associated with antipsychotic-naïve FEP (Perry et al., 2016). Abnormal HDL (Rader and Hovingh, 2014) and triglycerides (Nordestgaard and Varbo, 2014) are longitudinally associated with cardiometabolic outcomes. Therefore, I chose to include HDL and triglycerides as continuous variables because they are associated with dyslipidaemia in FEP, are associated with long term cardiometabolic outcomes, and are also a useful risk-marker for insulin resistance considering that gold-standard measures for insulin resistance (e.g. HOMA-IR (Levy et al., 1998)) are rarely carried out in current psychiatric clinical practice. I also developed a PsyMetRiC 'partial-model' (without HDL and triglycerides) to cover eventualities where biochemical results are not available.

7.3.4 Statistical Analysis

7.3.4.1 Algorithm Development and Internal Validation

I developed PsyMetRiC using the forced entry method, after ruling out predictor multi-collinearity, to minimize risk of overfitting and as recommended for smaller datasets (Steyerberg et al., 2000). I performed a formal sample size calculation. See Section 7.3.4.2 below. I did not consider non-linear terms or interactions to reduce risk of overfitting. I used MICE for missing data and estimates were pooled using Rubin's rules. See Section 7.3.4.3 below. An initial internal validation step (500 bootstraps) was performed, and coefficients were shrunk for optimism using the pooled corrected C-slope as a shrinkage factor. After this step, predictive performance was assessed.

7.3.4.2 Sample Size Calculation

Riley and colleagues (Riley et al., 2019) proposed a set of criteria that sample size should meet for development of a prediction algorithm with a binary outcome, in order to minimise the risk of overfitting and to ensure precise estimation of key parameters in the prediction algorithm. The sample size calculation requires the user-specified anticipated R^2 of the algorithm, and the average outcome value and standard deviation of outcome values in the population of interest. The three criteria are: a) small overfitting defined by an expected shrinkage of predictor effects by 10% or less; b) small absolute difference of 0.05 in the algorithm's apparent and adjusted Nagelkerke's R-squared value; c) precise estimation (within +/- 0.05) of the average outcome risk in the population.

Three calculations of sample size are made based upon these criteria. The final recommended sample size is taken as the largest of the three individual calculations (Riley et al., 2019). The above criteria have been developed into a statistical package, *pmsampsize* (Ensor, 2019) for R (R Core Team, 2017), which I used for sample size calculation. The user-specified arguments were:

1) Outcome prevalence = 20% based on meta-analytic prevalence estimates of unmedicated psychosis patients (Mitchell et al., 2013b).

2) R^2 = 0.15, selected as a conservative estimate since there is no equivalent risk prediction algorithm developed in the same population with which to base the calculation. I did not consider using the one previous cardiovascular risk prediction algorithm developed for people with serious mental illness (PRIMROSE) (Osborn et al., 2015) to derive the calculation since PRIMROSE was developed in an older population, and with a different outcome. Should I have used that estimate (C=0.80, converted using Table 2 from Riley and colleagues (Riley et al., 2019) to R^2=0.47), the sample size requirement would have been significantly smaller.

3) Shrinkage = 0.9 (as recommended (Ensor, 2019)).

After applying the above criteria, the minimum required sample size based on the number of included predictors was *n*=494 for the full-model, and *n*=384 for the partial-model. See Table 37.

Table 37: Results of Sample Size Calculations for PsyMetRiC

Criteria	Sample Size	Shrinkage	Parameters	R^2	EPV Ratio
Full-Model					
Criteria 1	494	0.90	9	0.15	10.98
Criteria 2	259	0.83	9	0.15	5.76
Criteria 3	246	0.90	9	0.15	5.47
Final	494	0.90	9	0.15	10.98
Partial-Model					
Criteria 1	384	0.90	7	0.15	10.97
Criteria 2	201	0.83	7	0.15	5.74
Criteria 3	246	0.90	7	0.15	7.03
Final	384	0.90	7	0.15	10.97

EPV=events per variable

7.3.4.3 Missing Data

I used MICE (Buuren, 2011) for missing data in all samples for predictors which were <40% missing (Lee, 2011) and had suitable auxiliary variables available for use as 'indicators of missingness' to

reduce the impact of 'missing not at random' bias (Dong and Peng, 2013). I imputed 100 datasets. Auxiliary variables were selected based upon minimizing the fraction of missing information (Madley-Dowd et al., 2019). Box-and-Whisker and Density plots were used to check similarities of observed and imputed data. Estimates were pooled using Rubin's rules. See Table

Table 38: Proportion of Missing Data per Variable for Model Development and External Validation

Predictor	Model Development Sample	External Validation Sample
Sex	0	0
Ethnicity	0	0
Smoking Status	0	0
Age	0	0
Antipsychotic Prescription	0	0
SBP – Baseline	0.11	0.09
SBP – Follow-up	0.38	0.09
BMI – Baseline	0.32	0.17
BMI – Follow-up	0.31	0.13
Triglycerides – Baseline	0.33	0.16
Triglycerides – Follow-up	0.37	0.20
HDL – Baseline	0.33	0.16
HDL – Follow-up	0.37	0.20

SBP=systolic blood pressure; BMI=body mass index; HDL=high-density lipoprotein

7.3.4.4 External Validation and Tests of Algorithm Performance

The algorithms were applied to the external validation sample. The distribution of predicted outcome probabilities was inspected using histograms. Algorithm performance was primarily assessed with measures of discrimination (concordance (c-) statistic), and calibration (calibration plots). The C-statistic is derived from the area under the curve and estimates the probability that a randomly selected 'case' will have a higher predicted probability for incident metabolic syndrome than a randomly selected non-case. Scores of 1.0 indicate perfect discrimination; scores of 0.5 indicate that the algorithm is no better than chance; scores of >0.7 are generally considered acceptable (Fukuma et al., 2018). Calibration plots estimate the accuracy of absolute-risk estimates (i.e., agreement between observed and predicted risk). I also recorded the Nagelkerke-Cox-Snell-Maddala-Magee R^2 index, the calibration intercept (ideally close to 0), C-slope (ideally close to 1), and the Brier score which is an overall measure of algorithm performance (ideally close to 0, with scores >0.25 indicating a poor model).

7.3.4.5 Clinical Usefulness and Potential Cut-offs

Decision curve analysis (Vickers and Elkin, 2006) was used to assess the clinical usefulness of PsyMetRiC by estimating net benefit. Net benefit is a metric of true positives minus false positives, and is calculated as:

$$\text{sensitivity} \times \text{prevalence} - (1 - \text{specificity}) \times (1 - \text{prevalence}) \times w$$

where w is the outcome odds at a given risk threshold (Vickers et al., 2019). The risk threshold is the amount of tolerable risk before an intervention is deemed necessary. Net benefit incorporates the consequences of the decisions made on the basis of an algorithm, and is therefore preferable to related measures such as sensitivity and specificity alone (Vickers et al., 2019). I also reported the standardized net benefit (net benefit / outcome prevalence) and related metrics (sensitivity and specificity) across a range of reasonable risk thresholds. I drew a decision curve plot which visualised the net benefit of both PsyMetRiC versions over varying risk-thresholds compared with intervening in all or intervening in none. Classical decision theory proposes that at a chosen risk-threshold, the choice with the greatest net-benefit should be preferred (Vickers et al., 2019).

7.3.5 Visual Representation of PsyMetRiC

I simulated two case histories applying the PsyMetRiC algorithms. Additionally, I developed an online data-visualisation app using shiny (Chang, 2020) for R (R Core Team, 2017), which allows an interactive exploration of the impact of modifiable and non-modifiable risk factors and their combinations on cardiometabolic risk in young people with psychosis, based on PsyMetRiC scores.

7.4 Results

7.4.1 Model Development and Coefficient Shrinkage

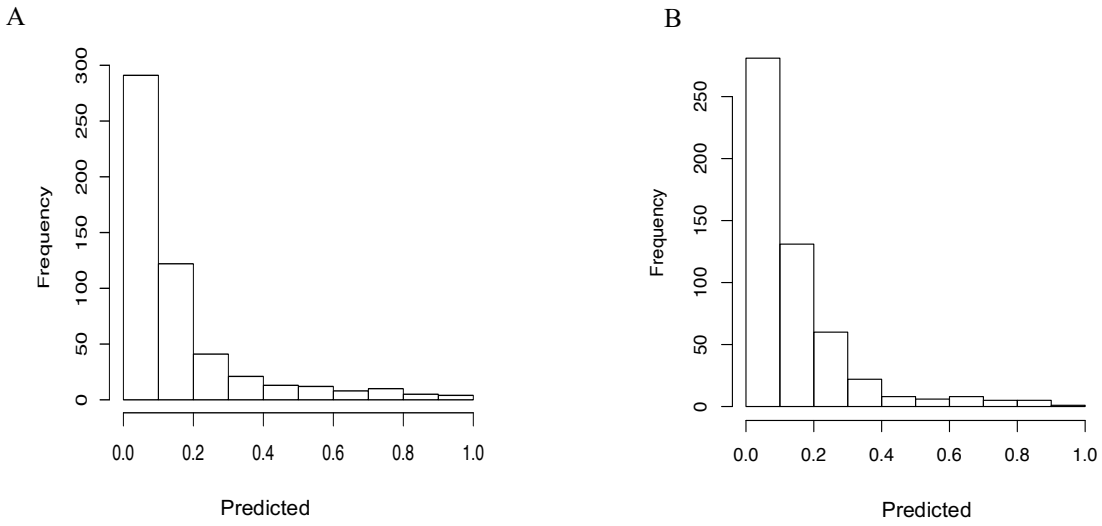
After 500 bootstraps, the pooled corrected C-slopes were: full-model: 0.90; partial-model: 0.93, which were used as shrinkage factors. Final PsyMetRiC coefficients are presented in Table 39. See Figure 22 for histograms showing the distribution of predicted outcome probabilities in the model development sample.

Table 39: Final Coefficients for PsyMetRiC Algorithms After Shrinkage

Predictor	Full-Model	Partial-Model
Intercept	-6.439813	-6.973829
Age	0.006233226	0.00633115
Black/African-Caribbean Ethnicity	0.004258861	0.07548129
Asian / Other Ethnicity	0.211217746	0.29285950
Male Sex	0.222300765	0.31460036
Body Mass Index	0.141186241	0.16912161
Smoking	0.153691193	0.24751854
Prescribed Metabolically-Active Antipsychotic	0.497552758	0.60013558
High-Density Lipoprotein (mmol/L)	-0.399013329	[a]
Triglycerides (mmol/L)	0.343528440	[a]

[a]Variable not included in model

Figure 22: Histograms of Predicted Outcome Probabilities in PsyMetRiC Development Sample after Coefficient Shrinkage

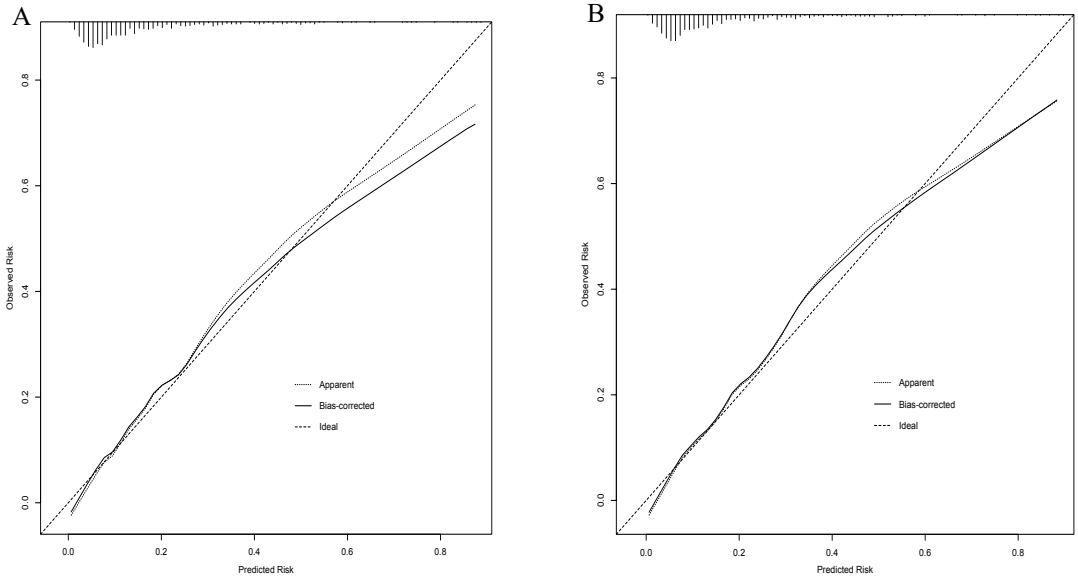


A=Full-Model; B=Partial-Model

7.4.2 Results of the Internal Validation Analysis

At internal validation, the pooled performance statistics were: full-model: C=0.80 (95% C.I., 0.74-0.86); r^2=0.25 (95% CI 0.22-0.28); Brier score=0.07 (95% C.I., 0.05-0.09); intercept=-0.05 (95% C.I., -0.08, -0.02); partial-model: C=0.79 (95% C.I., 0.73-0.84); r^2=0.19 (95% C.I., 0.14-0.24); Brier score=0.10 (95% C.I., 0.07-0.13); intercept=-0.07 (95% C.I., -0.10, -0.04). Calibration plots showed good agreement between observed and expected risk at most predicted probabilities, although in both PsyMetRiC versions there was evidence of slight over-prediction of risk at higher predicted probabilities. See Figure 23.

Figure 23: Internal Validation Calibration Plots for PsyMetRiC in Development Sample



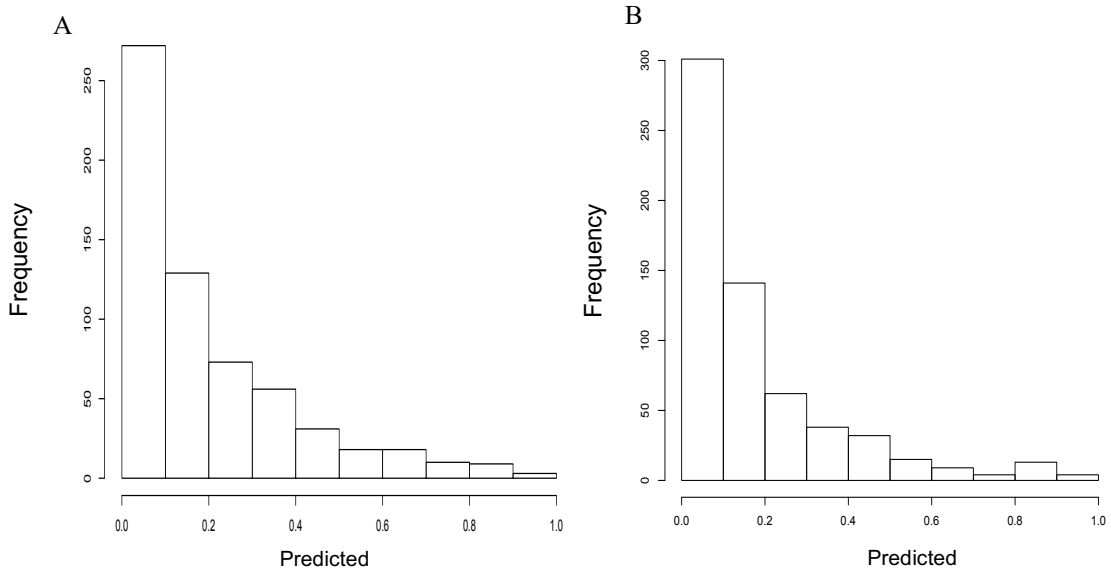
A=Full-Model; B=Partial-Model
Calibration plots illustrate agreement between observed risk (y axis) and expected risk (x axis). Perfect agreement would trace the dotted "ideal" line. Algorithm calibration is illustrated by the dotted (Apparent) and solid (Bias Corrected) lines.

7.4.3 Results of the External Validation

See Figure 24 for histograms of predicted outcome probabilities for the PsyMetRiC algorithms when applied to the SLaM EIS sample. Performance statistics were: full-model: C=0.75 (95% C.I., 0.69-0.80; r^2=0.21 (95% CI., 0.18-0.25); Brier score=0.07 (95% C.I., 0.04-0.10); intercept=-0.05 (95% C.I., -0.08, -0.02); partial-model: C=0.74 (95% C.I., 0.67-0.79); r^2=0.17 (95% C.I., 0.14-0.20); Brier score=0.08 (95% C.I., 0.05-0.11); intercept=-0.07 (95% C.I., -0.11, -0.03). Calibration plots (Figure 25) show good agreement between observed and expected risk in the full-model; but in the partial model there was evidence of slight miscalibration (under-prediction of risk at lower predicted probabilities, and over-prediction of risk at higher predicted probabilities). In both models, confidence

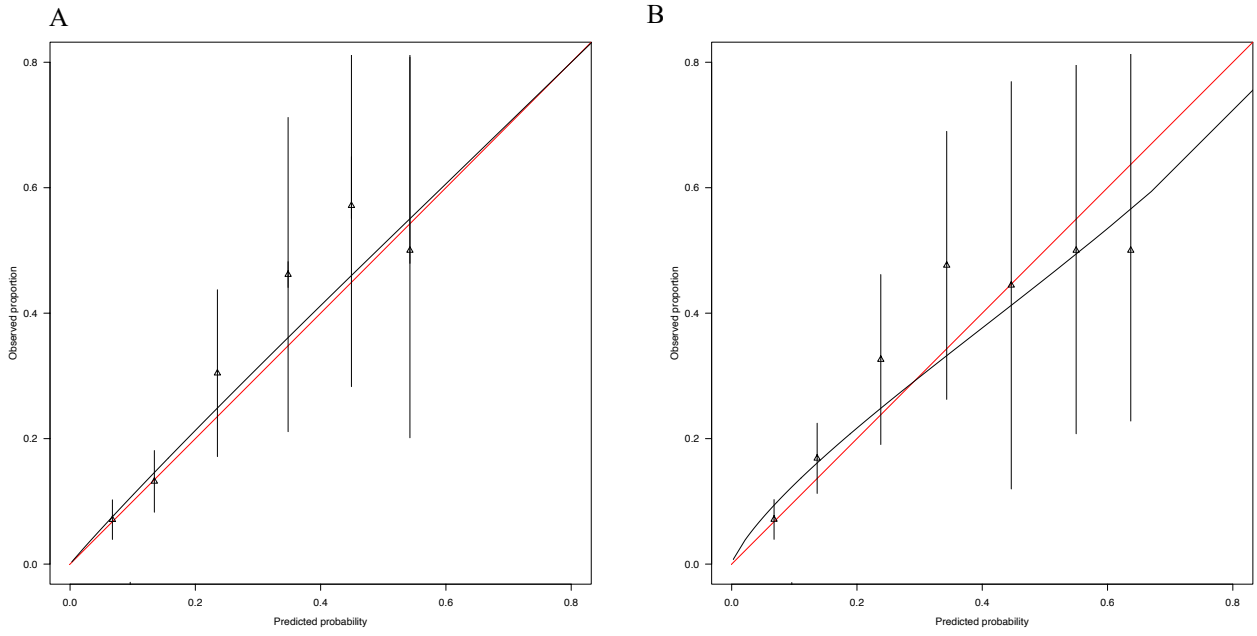
intervals widened as predicted probabilities became more extreme due to lower numbers of participants with more extreme predicted probabilities.

Figure 24: Histograms of Predicted Outcome Probabilities in External Validation Sample



A=Full-Model; B=Partial-Model

Figure 25: External Validation Calibration Plots for PsyMetRiC

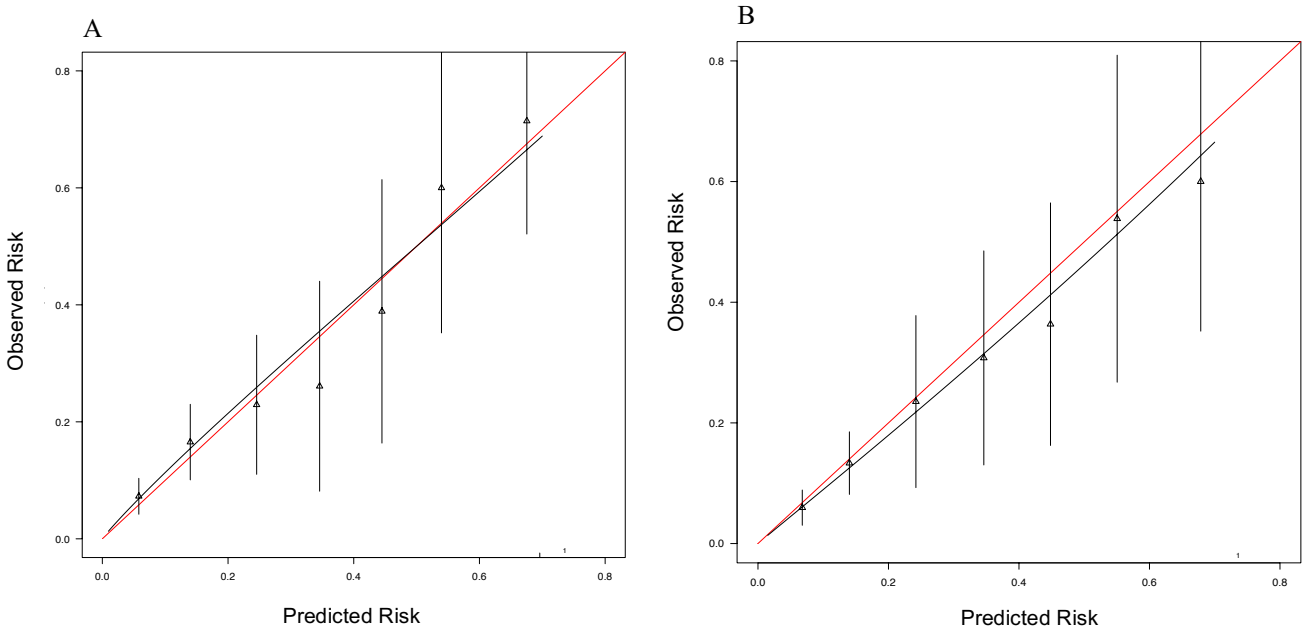


Calibration plots illustrate agreement between observed risk (y axis) and predicted risk (x axis). Perfect agreement would trace the red line. Algorithm calibration is illustrated by the black line. Triangles denote grouped observations for participants at deciles of predicted risk, with 95% C.I.'s indicated by the vertical black lines. Axes range between 0-0.8 since very few individuals received predicted probabilities greater than 0.8

7.4.4 Results of the External Validation Sensitivity Analysis

In the ALSPAC sample, performance statistics were: full-model: C=0·73 (95% C.I., 0·66-0·79; r^2=0·20 (95% CI., 0·17-0·23); Brier score=0·08 (95% C.I., 0·04-0·11); intercept=-0·03 (95% C.I., -0·07, 0·01); partial-model: C=0·71 (95% C.I., 0·64-0·77); r^2=0·17 (95% C.I., 0·13-0·22); Brier score=0·09 (95% C.I., 0·05-0·13); intercept=-0·03 (95% C.I., -0·07, 0·00). Calibration plots (Figure 26) show relatively good agreement between observed and expected risk in the full-model albeit with some minor evidence of miscalibration (slight under-prediction of risk at lower predicted probabilities, and over-prediction of risk at higher predicted probabilities). The same pattern of slight miscalibration was marginally more pronounced in the partial-model.

Figure 26: Calibration Plots in ALSPAC Sensitivity Analysis Sample



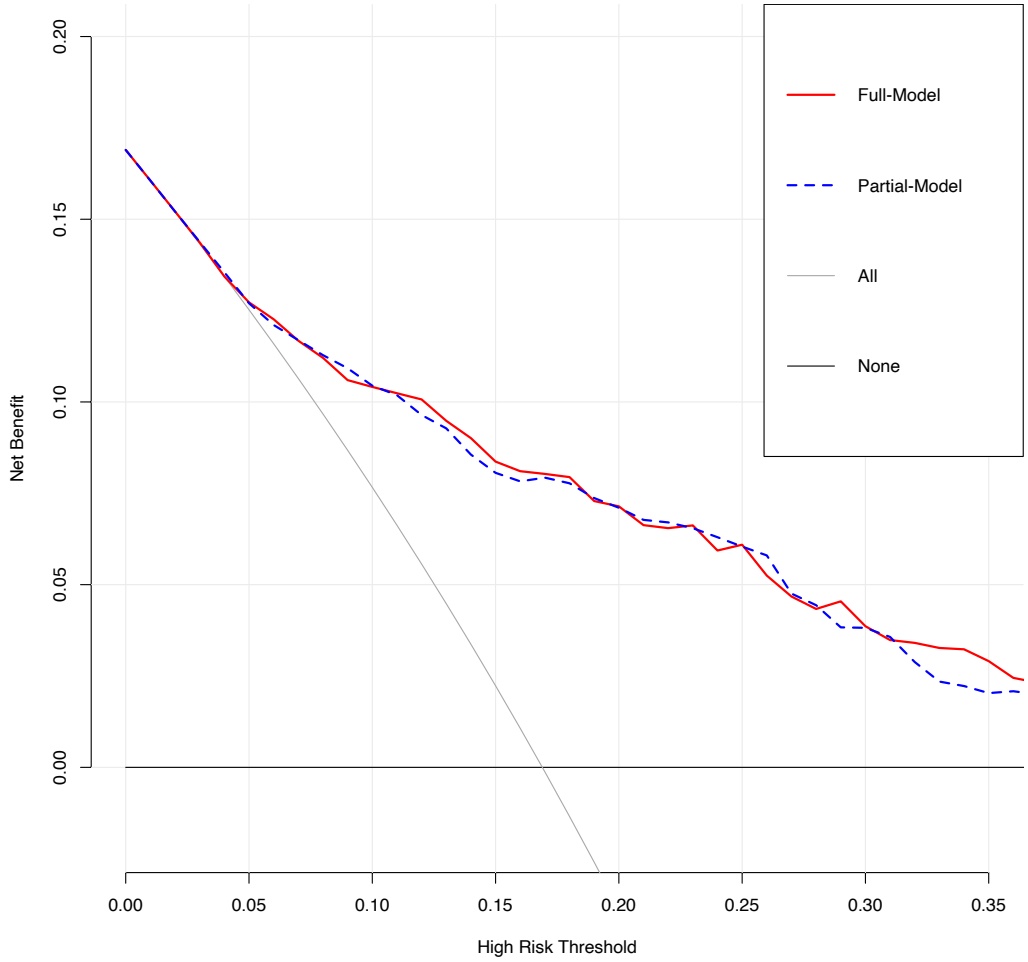
A=Full-Model; B=Partial-Model.
Calibration plots illustrate agreement between observed risk (y axis) and predicted risk (x axis). Perfect agreement would trace the red line. Algorithm calibration is illustrated by the black line. Triangles denote grouped observations for participants at deciles of predicted risk, with 95% C.I.'s indicated by the vertical black lines.

7.4.5 Results of The Decision Curve Analysis

Decision curve analysis (Figure 27) suggested that at predicted probability cut-offs >0.05, both PsyMetRiC algorithms provided greater net benefit than intervening in all or none. At most risk thresholds >0.05, the full-model provided slight improvement in net benefit compared with the partial model. See Tables 40-41 for numerical decision curve analysis results for both PsyMetRiC versions (net benefit, standardized net benefit, sensitivity, specificity) across a range of reasonable risk

thresholds. For example, if an intervention were considered necessary above a risk score of 0.18, the full-model would provide a net benefit of 7.95% (95% C.I., 5.37-10.82%) with a sensitivity and specificity of 0.75 (95% C.I., 0.66-0.82) and 0.74 (95% C.I., 0.71-0.78) respectively, meaning that an additional 47% of metabolic syndrome cases could be prevented. At the same risk-threshold, the partial-model would provide a net benefit of 7.74% (95% C.I., 4.79-10.36%) with a sensitivity and specificity of 0.75 (95% C.I., 0.65-0.81) and 0.74 (95% C.I., 0.70-0.77) respectively, meaning that an additional 46% of metabolic syndrome cases could be prevented. For both models this equates to around an additional eight cases of metabolic syndrome that could be prevented per 100 individuals, without any increase in false positives.

Figure 27: Decision Curve Analysis Plot for PsyMetRiC Full- and Partial-Models



The plot reports net benefit (y axis) of PsyMetRiC full- (red line) and partial- (blue line) models across a range of risk thresholds (x axis) compared with intervening in all (grey line) or intervening in none (black line). In decision curve analysis, it is customary to consider only the range of risk-thresholds that may reasonably be considered in clinical practice. The upper bound of 0.35 represents a greater than one in three chance of developing metabolic syndrome should nothing change, and it is unlikely that risk thresholds greater than this would be tolerated. Net harm (i.e., more false positives than true positives exposed to an intervention at a selected risk threshold) is indicated when a proposed intervention is plotted at y<0.

Table 40: Decision Curve Analysis Results at Different Thresholds – PsyMetRiC Full-Model

Net Benefit Performance Measure (95% C.I.)				
Risk Threshold[a]	**Sensitivity**	**Specificity**	**Net Benefit**	**Standardized Net Benefit[b]**
0.02	1.00 (1.00-1.00)	0.01 (0.00-0.02)	0.15 (0.13-0.18)	0.90 (0.88-0.92)
0.04	0.99 (0.97-1.00)	0.04 (0.03-0.06)	0.13 (0.11-0.16)	0.80 (0.75-0.83)
0.06	0.99 (0.97-1.00)	0.16 (0.12-0.19)	0.12 (0.09-0.15)	0.73 (0.67-0.77)
0.08	0.96 (0.92-1.00)	0.30 (0.26-0.34)	0.11 (0.09-0.14)	0.66 (0.58-0.72)
0.10	0.94 (0.88-0.98)	0.41 (0.37-0.46)	0.10 (0.08-0.13)	0.62 (0.52-0.69)
0.12	0.92 (0.86-0.97)	0.52 (0.47-0.57)	0.10 (0.07-0.13)	0.60 (0.50-0.68)
0.14	0.85 (0.77-0.91)	0.61 (0.57-0.65)	0.09 (0.06-0.12)	0.53 (0.44-0.62)
0.16	0.76 (0.69-0.83)	0.70 (0.66-0.74)	0.08 (0.06-0.11)	0.48 (0.38-0.59)
0.18	0.75 (0.66-0.82)	0.74 (0.71-0.78)	0.08 (0.05-0.11)	0.47 (0.37-0.58)
0.20	0.68 (0.59-0.77)	0.79 (0.75-0.83)	0.07 (0.05-0.10)	0.42 (0.31-0.53)
0.22	0.62 (0.52-0.70)	0.83 (0.80-0.87)	0.07 (0.04-0.09)	0.39 (0.27-0.49)
0.24	0.56 (0.47-0.65)	0.86 (0.83-0.89)	0.06 (0.04-0.08)	0.35 (0.22-0.49)
0.26	0.52 (0.43-0.62)	0.88 (0.85-0.91)	0.05 (0.03-0.07)	0.31 (0.19-0.43)
0.28	0.45 (0.37-0.54)	0.90 (0.87-0.92)	0.04 (0.02-0.07)	0.26 (0.15-0.38)
0.30	0.40 (0.31-0.50)	0.92 (0.89-0.94)	0.04 (0.02-0.06)	0.23 (0.12-0.36)
0.32	0.37 (0.28-0.47)	0.93 (0.90-0.95)	0.03 (0.02-0.06)	0.20 (0.10-0.32)
0.34	0.34 (0.24-0.43)	0.94 (0.92-0.96)	0.03 (0.01-0.05)	0.19 (0.08-0.30)
0.36	0.27 (0.19-0.36)	0.95 (0.94-0.97)	0.02 (0.01-0.04)	0.14 (0.04-0.26)

[a]Different risk thresholds may be selected depending on the proposed intervention, as well as patient or clinician preference; [b]Standardized net benefit is calculated as the net benefit / outcome prevalence, showing the proportion of improvement in net benefit at the selected risk threshold.

Table 41: Decision Curve Analysis Results at Different Thresholds – PsyMetRiC Partial-Model

Net Benefit Performance Measure (95% C.I.)				
Risk Threshold[a]	**Sensitivity**	**Specificity**	**Net Benefit**	**Standardized Net Benefit[b]**
0.02	1.00 (1.00-1.00)	0.01 (0.00-0.01)	0.15 (0.12-0.18)	0.90 (0.88-0.92)
0.04	1.00 (1.00-1.00)	0.03 (0.02-0.05)	0.14 (0.11-0.16)	0.80 (0.75-0.83)
0.06	0.99 (0.96-1.00)	0.13 (0.10-0.15)	0.12 (0.09-0.15)	0.72 (0.64-0.77)
0.08	0.99 (0.96-1.00)	0.24 (0.21-0.28)	0.11 (0.08-0.14)	0.67 (0.58-0.73)
0.10	0.95 (0.91-0.99)	0.38 (0.34-0.43)	0.10 (0.07-0.13)	0.62 (0.53-0.69)
0.12	0.91 (0.86-0.96)	0.50 (0.46-0.54)	0.10 (0.07-0.12)	0.57 (0.47-0.65)
0.14	0.85 (0.78-0.91)	0.58 (0.53-0.62)	0.09 (0.06-0.11)	0.51 (0.38-0.59)
0.16	0.78 (0.71-0.86)	0.66 (0.62-0.70)	0.08 (0.05-0.11)	0.46 (0.33-0.55)
0.18	0.75 (0.65-0.83)	0.74 (0.70-0.77)	0.08 (0.05-0.10)	0.46 (0.33-0.56)
0.20	0.67 (0.60-0.75)	0.79 (0.76-0.83)	0.07 (0.04-0.09)	0.42 (0.30-0.51)
0.22	0.65 (0.56-0.72)	0.82 (0.79-0.86)	0.07 (0.04-0.09)	0.40 (0.27-0.50)
0.24	0.59 (0.50-0.67)	0.86 (0.83-0.90)	0.06 (0.04-0.08)	0.37 (0.25-0.48)
0.26	0.56 (0.47-0.65)	0.87 (0.85-0.91)	0.06 (0.03-0.08)	0.34 (0.23-0.44)
0.28	0.48 (0.40-0.57)	0.89 (0.86-0.92)	0.04 (0.02-0.07)	0.26 (0.13-0.37)
0.30	0.41 (0.34-0.50)	0.91 (0.89-0.94)	0.04 (0.02-0.06)	0.23 (0.11-0.33)
0.32	0.35 (0.28-0.44)	0.92 (0.90-0.94)	0.03 (0.01-0.05)	0.17 (0.06-0.27)
0.34	0.29 (0.21-0.38)	0.94 (0.92-0.96)	0.02 (0.00-0.04)	0.13 (0.02-0.24)
0.36	0.28 (0.20-0.36)	0.94 (0.92-0.96)	0.02 (0.00-0.04)	0.12 (0.01-0.22)

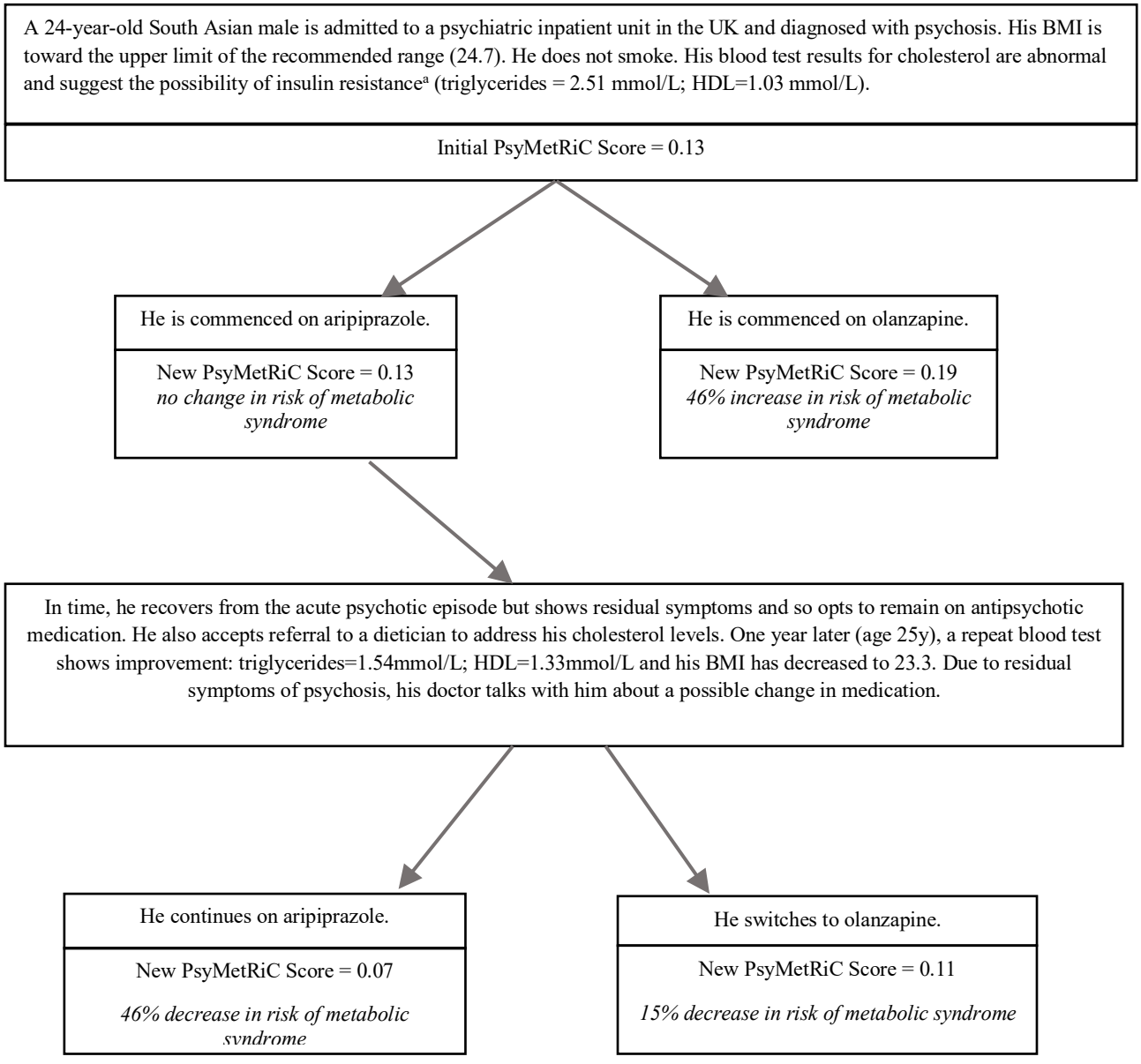
[a]Different risk thresholds may be selected depending on the proposed intervention, as well as patient or clinician preference; [b]Standardized net benefit is calculated as the net benefit / outcome prevalence, showing the proportion of improvement in net benefit at the selected risk threshold.

7.4.6 Visual Representation of PsyMetRiC

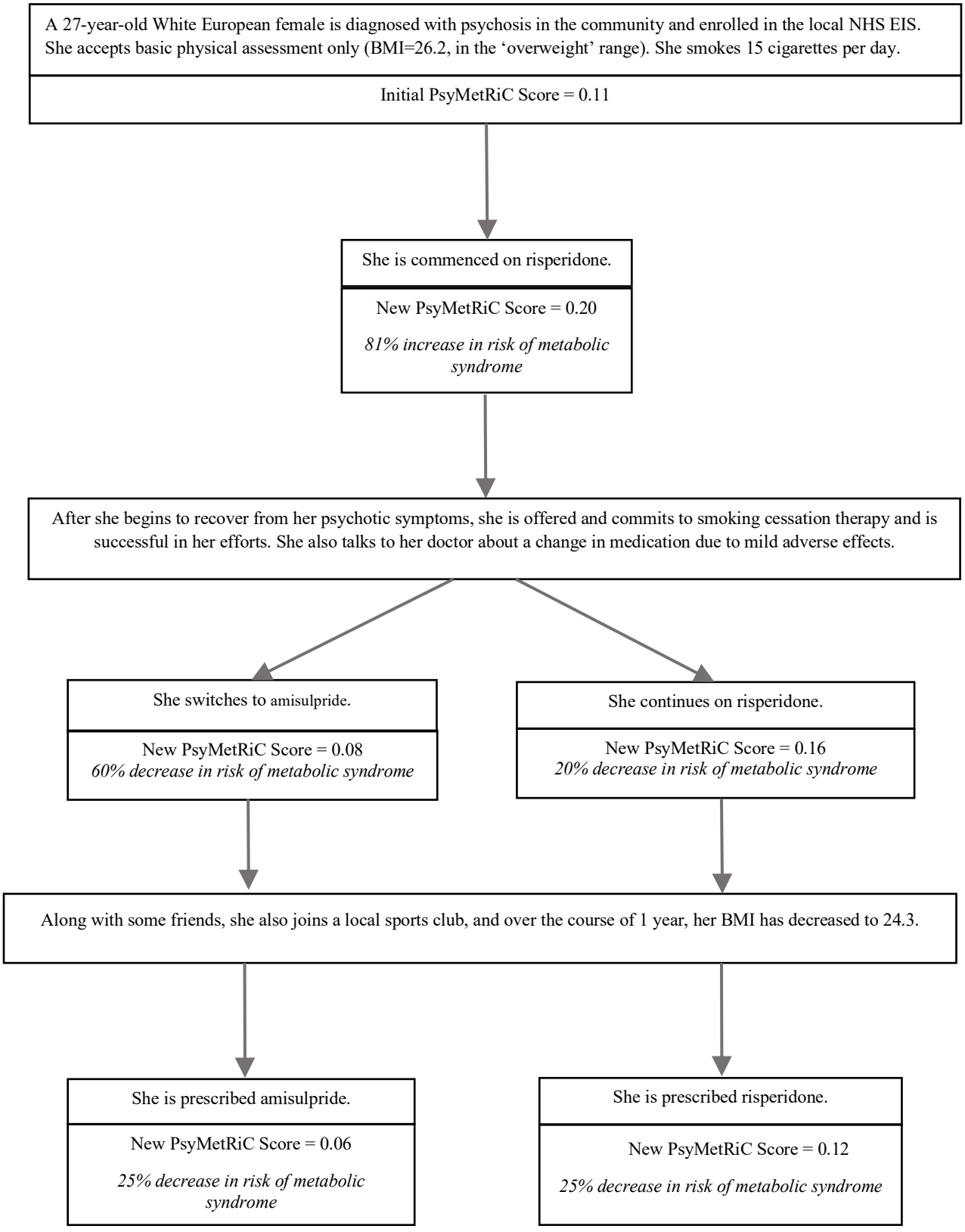
Figure 28 shows decision trees outlining two simulated case scenarios to visualise the impact of modifiable and non-modifiable risk factors in young people with psychosis, as calculated from PsyMetRiC full- and partial-models. Visit http://psymetric.shinyapps.io/psymetric/ for an online data visualisation app for both PsyMetRiC versions, which allows the user to interactively explore the impact of modifiable and non-modifiable risk factors and their combinations on cardiometabolic risk in young people with psychosis, based on PsyMetRiC scores.

Figure 28: Simulated Case Scenarios to Visualize Impact of Modifiable and Non-Modifiable Risk Factors on Cardiometabolic Risk in Young People with Psychosis as Calculated from PsyMetRiC Full- and Partial Models

A. PsyMetRiC Full-Model



B. PsyMetRiC Partial-Model



PsyMetRiC scores presented as predicted probabilities, which can be converted to %chance of incident metabolic syndrome by multiplying by 100. [a]A raised triglyceride:HDL ratio is indicative of insulin resistance
EIS=psychosis early intervention service; BMI=body mass index; HDL=high-density lipoprotein.

7.5 Discussion

I have developed and externally validated PsyMetRiC, which is, to the best of my knowledge, the first cardiometabolic risk prediction algorithm specifically tailored for young people with psychosis. PsyMetRiC can predict up to six-year risk of incident metabolic syndrome from commonly recorded clinical information, highlighting modifiable risk factors that could be addressed to reduce risk. Metabolic syndrome is a precursor to CVD and early mortality (Isomaa et al., 2001) and is a suitable outcome for younger populations. Both PsyMetRiC versions externally validated well, with C-statistics >0·70. Calibration of the full-model was good, but there was evidence of slight miscalibration of the partial-model. Therefore, the partial model may benefit from recalibration in larger samples. Both PsyMetRiC versions displayed greater net benefit than alternative strategies across a range of feasible risk thresholds. However, at most risk thresholds, the results show that the full-model should be used preferentially.

The data visualisations in Section 7.4.5 help to illustrate three things: First, antipsychotic medication choice imparts a substantial influence on cardiometabolic risk; second, addressing lifestyle factors can effectively reduce cardiometabolic risk even in the presence of antipsychotic medication; third; advancing age in relatively-young adults does not substantially influence cardiometabolic risk relative to other risk factors. While PsyMetRiC will benefit from future validation in larger samples, it has the potential to become a valuable resource to promote better management of physical health in young people with psychosis. PsyMetRiC could be used to highlight malleable risk factors and encourage clinicians to make more personalized, informed decisions such as with the choice of antipsychotic medication and lifestyle interventions.

Over 100 studies were included in my systematic review that explored the suitability of existing cardiometabolic risk prediction algorithms for young people with psychosis (see Chapter 6). Yet, few algorithms were externally validated; only one was developed in a sample of people with mental illness; none were conducted in younger populations; and most were rated as being high-risk of bias.

Ethnicity, smoking, and BMI are amongst the most commonly included predictors in existing algorithms (see Chapter 6) and are well-known contributors to cardiometabolic risk (Pillinger et al., 2020), so I included them in PsyMetRiC. Sex is also frequently considered in existing algorithms, and I included it in PsyMetRiC. I found that male sex was a risk factor for incident metabolic syndrome, which aligns with meta-analytic reports that male sex is a risk factor for antipsychotic-induced metabolic dysfunction (Pillinger et al., 2020). Due to the available sample size, I could not consider separate versions of PsyMetRiC for males and females. When larger samples might be

available in future, sex-stratified versions might be studied since existing algorithms developed for the general population commonly take this step.

Age is frequently included in existing algorithms (see Chapter 6), and I included it in PsyMetRiC. However, existing cardiometabolic risk prediction algorithms developed for relatively older-aged adults have weighted age to a much greater extent than other predictors (see Chapter 6). This is likely because most cardiometabolic risk factors contribute cumulative risk over time (Reinikainen et al., 2015), and so age becomes increasingly important as one gets older. In Chapter 6, the accompanying exploratory validation analysis, which examined the predictive performance of existing general population cardiometabolic risk prediction algorithms (QRISK3, QDiabetes and PRIMROSE) in young people who were at risk of developing psychosis, found that each significantly underpredicted risk in the younger population. This is possibly due to the way existing algorithms have modelled age. With PsyMetRiC, age is weighted to a much lesser extent than other predictors, and I achieved favourable calibration in younger populations. While QRISK3, QDiabetes and PRIMROSE are good examples of well-designed algorithms from enormous samples, my results suggest that PsyMetRiC is more appropriate for young people with psychosis.

Blood-based predictors such as HDL and triglycerides feature less often in cardiometabolic risk prediction algorithms (see Chapter 6). Meta-analytic evidence suggests abnormal triglyceride and HDL levels are detectable at FEP (Misiak et al., 2017), and a raised triglyceride:HDL ratio is a hallmark of insulin resistance (Murguia-Romero et al., 2013), which is also associated with FEP (Perry et al., 2016). Guideline recommendations encourage blood-based monitoring pre- and post-antipsychotic exposure (Barnes et al., 2020), and so biochemical data should be available. I found that the inclusion of blood-based predictors improved all predictive performance metrics. However, blood-based monitoring may not always be possible, and I found that the partial-model still provided relatively reliable performance estimates, although it would benefit from recalibration.

Antipsychotic medication is an important contributor to cardiometabolic risk in young people with psychosis yet has rarely been included in existing algorithms. Some more recent algorithms have included antipsychotics as predictors, grouped by the traditional distinctions of typical/atypical or first/second-generation (see Chapter 6). However, the differential cardiometabolic effects of antipsychotics do not abide by these distinctions. Therefore, I instead modelled antipsychotics based on previous research. This is an advance over previous algorithms.

PsyMetRiC cannot yet be recommended for clinical use and requires prospective validation in larger samples, health technology assessment, and regulatory approval. However, PsyMetRiC can become a valuable resource for the better management of physical health in young people with psychosis in

the future. For example, in the presence of a low PsyMetRiC risk score, gentle encouragement to maintain good physical health may be sufficient. For instance, this might include dietary advice, promoting daily physical activity and smoking cessation, if necessary. There is little harm yet much to gain in offering gentle encouragement to live a healthier life, and such conversations need to become part and parcel of psychiatric consultation.

Patients and clinicians might prefer to tolerate a slightly higher risk threshold when the proposed intervention could be deemed more burdensome or may increase the risk of other adverse effects. For example, prescribed lifestyle interventions have shown promise in lowering cardiometabolic risk in young people with psychosis (Fernandez-Abascal et al., 2021), however they may be perceived as burdensome, involving regular appointments that may be difficult to maintain around work or other commitments.

Yet, dietary interventions (Curtis et al., 2016) have shown promise when offered to young people with psychosis but may be less effective in older adults with more chronic illness and ingrained behaviours (Speyer et al., 2016).

Regarding smoking cessation, a systematic review and meta-analysis found relatively strong evidence for pharmacological interventions such as varenicline, a selective nicotine receptor partial agonist, and bupropion, a selective catecholamine reuptake inhibitor. The review found limited evidence for behavioural interventions (Pearsall et al., 2019). A systematic review on the psychosocial barriers to smoking cessation in schizophrenia found that cravings were the main barrier to smoking cessation, followed by a perception that negative symptoms worsened when attempting to quit (Lum et al., 2018).

Regarding physical activity interventions, a Cochrane review of randomized controlled trials (RCTs) found that despite study heterogeneity and small sample sizes, exercise interventions led to an improvement in negative symptoms and quality of life scores as well as weight loss (Gorczynski and Faulkner, 2010). A mixed-methods study found that people with schizophrenia who engaged in regular exercise reported beneficial effects on mood and cognitive symptoms, with improvements in well-being measures and reductions in negative and cognitive symptoms, following an exercise intervention (Ho et al., 2018). Similarly, other qualitative research has shown that people with schizophrenia who regularly engage in exercise reported improved symptom alleviation, improved confidence, and a sense of achievement (Firth et al., 2016).

Other interventions may increase the risk of other adverse effects. For example, my results show that switching from metabolically-active antipsychotics or not prescribing them in the first place is a highly effective means to reduce cardiometabolic risk. This finding is in line with a recent clinical

trial which found that switching to a less metabolically-active antipsychotic significantly reduced BMI in young people with psychosis (Correll et al., 2020). However, the risk of psychosis relapse or other adverse effects may reasonably be worrisome for patient and clinician alike. Moreover, meta-analyses suggest that metabolically-active antipsychotics could be associated with favourable psychosis treatment response (Pillinger et al., 2020), though this may be an artefact of treatment adherence. Nevertheless, antipsychotic selection must strike an intricate balance between caring for psychiatric and physical health. Finally, trials of treatments such as metformin and statins are scarce in young people with psychosis, but evidence suggests that such medications might benefit both cardiometabolic and psychiatric outcomes (Hayes et al., 2019, Correll et al., 2020, de Silva et al., 2016).

Regarding the strengths of the study, I have developed, to the best of my knowledge, the first cardiometabolic risk prediction algorithm for young people with psychosis, harnessing data from three geographically distinct patient samples and a population-based cohort. PsyMetRiC was developed in consultation with The McPin Foundation YPAG to ensure a balance between clinical practicality and patient acceptability, and I received encouraging comments from the YPAG about PsyMetRiC. I developed an online interactive app permitting a visualization of the impact of different cardiometabolic risk factors in young people with psychosis. I have reported the fully specified algorithm coefficients to encourage future validation and model updating. I developed two versions of PsyMetRiC to maximise clinical utility and both validated well, suggesting that PsyMetRiC is likely to be suitable for use in patients aged 16-35 years from a UK EIS population. From the sensitivity analysis results, PsyMetRiC may also be generalizable to young adults at risk of developing psychosis.

Limitations of the study include missing data. I excluded participants who had the outcome at baseline, as recommended (Wolff et al., 2019). However, because the predictors were measured a short time frame after EIS enrolment, some ‘metabolically-sensitive’ individuals, i.e., inidivduals who developed metabolic syndrome quickly, might have been inadvertently excluded from the analysis. I also excluded participants with data missing on either all exposure or all outcome variables, which may have introduced selection bias. The missing sample was more likely to be older and female and less likely to be prescribed metabolically-active antipsychotics. This may have affected some PsyMetRiC predictor coefficients. Nevertheless, I felt this exclusion step was more appropriate than imputing complete participant data.

Multiple imputation may be biased in instances where data are ‘missing not at random’. However, I included auxiliary variables to reduce the fraction of missing information and limit the impact of this. External validation of PsyMetRiC in larger samples is required since simulation studies have

suggested a minimum of 100 outcome events for an accurate validation analysis (Collins et al., 2016). Larger prospectively collected samples in future may also allow for updating the algorithm with interactions, non-linear terms, and sex-stratification. In addition, larger prospectively collected samples may allow the consideration of other potentially important predictors such as other metabolically-active medications, physical activity, and diet. Prospectively collected data may also allow prediction of longer-term risk since the mean follow-up time in the primary analysis was shorter than the maximum included time frame of six years. While the data-driven classification of metabolically-active antipsychotics is an advance over existing algorithms, the metabolically-active nature of different antipsychotics lies on a continuum rather than across a dichotomy. Larger samples may permit the modelling of antipsychotics individually. Prescriber bias may have downwardly biased the coefficients for antipsychotics since metabolically-active medications may have been withheld from patients considered to be at higher cardiometabolic risk.

In conclusion, I have developed and externally validated PsyMetRiC, an algorithm that can reliably predict the risk of incident metabolic syndrome in young people with psychosis. PsyMetRiC has the potential to become a valuable resource for healthcare professionals working in EIS. PsyMetRiC can aid the informed choice of psychotropic and non-psychotropic medications and non-pharmacological interventions, including lifestyle adjustments, to prevent the future development of cardiometabolic comorbidity and consequent years of life lost.

Section D: Summary of Main Findings and Conclusions

Section D concerned the clinical prediction of cardiometabolic risk in young people with psychosis, consisting of a systematic review (Chapter 6) followed by the development of PsyMetRiC, a cardiometabolic risk prediction algorithm for young people with psychosis (Chapter 7). The systematic review (Chapter 6) identified a considerable number of cardiometabolic risk prediction algorithms developed for the middle- to older-aged general population. One algorithm was developed for a non-specific psychiatric population and was also developed in relatively older-aged adults. I also found that most algorithms were not externally validated, thus calling their potential generalizability into question, and most were rated as high risk of bias. Three identified algorithms (QRISK3, QDiabetes, PRIMROSE) included psychiatric predictors such as antipsychotic medications and were assessed for their predictive performance in a younger psychosis-risk population. All three substantially underpredicted cardiometabolic risk in the younger population. Based on the systematic review and exploratory validation analysis, I concluded that currently, no algorithm can be recommended for young people with psychosis, despite this population being at significantly higher cardiometabolic risk than the general population.

In Chapter 7, I developed the first cardiometabolic risk prediction algorithm tailored for young people with psychosis, the Psychosis Metabolic Risk Calculator (PsyMetRiC). I designed PsyMetRiC to be age-appropriate, clinically useful, and acceptable to patients. I developed PsyMetRiC using patient data from two UK EIS and externally validated it in a geographically distinct UK EIS and a population-based cohort. I developed and validated two versions of this tool, one with and one without blood-based biomarkers, to maximise usefulness in day-to-day clinical practice. Predictive performance for PsyMetRiC was universally good, suggesting that the algorithm is suitable for the UK EIS population. PsyMetRiC represents a valuable future tool for clinical practice, which now requires further testing in clinical settings through prospective validation and updating.

Together, results from Section D show that the prediction of cardiometabolic risk in young people with psychosis has been widely overlooked. Nevertheless, with PsyMetRiC, I have shown that it is possible to predict cardiometabolic risk in this population accurately. Whilst future refinements to PsyMetRiC are required to improve predictive performance further, PsyMetRiC is a valuable step toward improved physical healthcare for young people with psychosis.

SECTION E

DISCUSSION

Chapter 8

General Discussion

8.1 Summary of the Main Findings in This Thesis

In this thesis, using a number of complementary methodological approaches relating to genetic and observational epidemiology along with prognosis research, I present evidence that cardiometabolic dysfunction may predate the onset of psychosis and may be inherent to it. I found that this may be explained, at least in part, by common biological mechanisms such as shared genetic liability and inflammation. I also found that it is possible to accurately predict cardiometabolic risk in young people with psychosis from commonly recorded information. Together, these findings can help to explain why young people with psychosis present with cardiometabolic dysfunction even in the earliest stages of illness. The findings can also pave the way for novel therapeutic and preventative approaches for schizophrenia and its associated cardiometabolic comorbidity.

First, to address the issue of direction of association, I used longitudinal data from the ALSPAC birth cohort to delineate developmental trajectories of cardiometabolic indices from early childhood to early adulthood and then tested associations with adult schizophrenia-spectrum and depression outcomes. I found that persistently high fasting insulin levels from mid-childhood were associated in a dose-response manner with schizophrenia-spectrum phenotypes measured in adulthood. Evidence for the associations remained after adjustment for a range of potential confounders, including sex, ethnicity, social class, smoking, physical activity, calorie intake, alcohol, and substance use. This suggests that the traditional attributions of sociodemographic and lifestyle factors are unlikely to fully explain the comorbidity. The associations of disrupted glucose-insulin homeostasis were not identified with depression, a genetically and clinically similar mental disorder with well-known cardiometabolic comorbidity. Together, these findings suggest that disruptions to glucose-insulin homeostasis may be a specific primary pathophysiological hallmark of schizophrenia and may be detectable decades before the first clinical psychotic episode. Therefore, disrupted glucose-insulin homeostasis could be a cause rather than simply a consequence of psychotic illness or share common pathophysiologic mechanisms.

Second, I examined whether genetic predisposition for T2D and schizophrenia were associated with risk of psychosis and disrupted glucose-insulin homeostasis, respectively, and explored whether genetic influences on childhood inflammation could mediate any evident associations. I found that genetic predisposition to schizophrenia was associated with disrupted glucose-insulin homeostasis in early adulthood and *vice versa*, genetic predisposition to T2D was associated with increased risk of schizophrenia-spectrum phenotypes in early adulthood. I found that this risk was partly mediated by childhood inflammation.

Third, I used a set of complementary genomic methods to rigorously examine for evidence of shared genetic liability for schizophrenia, cardiometabolic and inflammatory traits, using summary data from large-scale GWAS. I found evidence for genetic overlap between schizophrenia, cardiometabolic and inflammatory traits that was confined to relatively lower-frequency genetic variants, was heterogeneous in nature, and could be pinpointed to biologically plausible pathways, for example, BDNF and glucose transport.

Fourth, I used summary GWAS data to examine whether insulin resistance and related cardiometabolic traits may be causally related to schizophrenia or whether inflammation may be a common biological mechanism for the comorbidity. I found consistent evidence supporting that inflammation could be a common cause for comorbid insulin resistance and schizophrenia. Together, these findings suggest that shared genetic liability and inflammation may be putative biological mechanisms that underly the associations between cardiometabolic disorders and schizophrenia, over and above the common attributions of sociodemographic, lifestyle and clinical factors.

Fifth, turning to the clinical relevance of the cardiometabolic comorbidity of schizophrenia, I performed a systematic review of cardiometabolic risk prediction algorithms to examine whether any might be suitable for young people with psychosis. Despite identifying a large number of algorithms, most had significant methodological shortcomings, and none were developed for younger populations. Using ALSPAC data, I found that existing algorithms substantially underpredicted cardiometabolic risk in a younger psychosis-risk sample. Therefore, I concluded that no existing cardiometabolic risk prediction algorithm is likely suitable for young people with psychosis.

Finally, using patient data from three UK EIS, I developed and validated PsyMetRiC, the first cardiometabolic risk prediction algorithm developed especially for young people with psychosis. The predictive performance of PsyMetRiC was good in both the development and external validation samples, suggesting that PsyMetRiC is likely to be suitable for use in the UK EIS population.

8.2 Interpretation and Context of the Main Findings in This Thesis

8.2.1 Examining The Nature of Association between Cardiometabolic Disorders and Schizophrenia

The finding that disruption to glucose-insulin homeostasis may predate psychosis (Chapter 2) could represent a considerable advance in our understanding of the nature of association between cardiometabolic traits and schizophrenia. Given the known cardiometabolic risk-increasing associations between schizophrenia and a host of lifestyle and clinical factors (as detailed in Section 1.1.3), it was previously assumed that the cardiometabolic comorbidity so prevalent in schizophrenia was simply a consequence of the psychiatric disorder. This assumption was bolstered by the findings of early systematic reviews and meta-analyses, which reported an unremarkable prevalence of cardiometabolic disorders like T2D and metabolic syndrome in antipsychotic naïve FEP (Mitchell et al., 2013a). The authors concluded that the cardiometabolic comorbidity of schizophrenia most likely arises after the onset of the psychiatric disorder and so must be a consequence of it.

However, just as psychotic symptoms may be distributed over a spectrum in the general population (van Os et al., 2009), neither is cardiometabolic dysfunction a binary distinction. Rather, subtle forms of cardiometabolic dysfunction may be present in the absence of clinical T2D or metabolic syndrome. Early systematic reviews failed to consider that the absence of relatively mature cardiometabolic phenotypes such as T2D and metabolic syndrome does not necessarily equate to the absence of cardiometabolic dysfunction. Indeed, more recent meta-analyses have consistently shown that subtle forms of disrupted glucose-insulin homeostasis, such as insulin resistance and impaired glucose tolerance, are detectable in FEP (Perry et al., 2016, Pillinger et al., 2017a, Greenhalgh et al., 2017). Findings from these meta-analyses, therefore, called into question the traditional understanding of the direction of association between cardiometabolic dysfunction and schizophrenia. This is because participants included in the studies were antipsychotic naïve and relatively young, and so less affected by commonly attributed lifestyle and clinical factors. Nevertheless, since all studies included in meta-analyses were either cross-sectional or featured existing cases of FEP, further elucidation on the direction of association could not be ascertained.

In Chapter 2, I present evidence that disruption to glucose-insulin homeostasis predates FEP and may be detectable from childhood in at least some individuals. While the commonly attributed lifestyle and clinical factors are not to be devalued and remain crucial therapeutic targets for the cardiometabolic comorbidity in schizophrenia, my findings suggest that these factors are more likely to exacerbate rather than cause the comorbidity. This argument is strengthened in light of the findings in Chapter 2, where associations between persistently high fasting insulin and schizophrenia-

spectrum phenotypes endured even after adjusting for a detailed set of sociodemographic and lifestyle confounders. One previous study sought to examine longitudinal associations between a single point-measure of fasting insulin levels measured at age 9 years and risk of psychosis at age 18 years in the ALSPAC cohort and found no evidence for an association (Perry et al., 2018). This discrepancy in findings from the same cohort underscores the importance of taking into account dynamic temporal changes and fluctuations in cardiometabolic markers, which are captured more effectively across repeated measures. In addition, the incidence of psychosis at age 18 years is relatively low. Therefore, the previous study may have included too few cases of outcome to detect an association.

In Chapter 2, I also present evidence for specificity of association between cardiometabolic dysfunction and schizophrenia-spectrum phenotypes. In addition to disrupted glucose-insulin homeostasis, recent meta-analyses have also consistently reported evidence of dyslipidaemia detectable in antipsychotic naïve FEP (Pillinger et al., 2017b, Misiak et al., 2017). However, the pattern of dyslipidaemia in FEP could be further evidence for a primary disruption to glucose-insulin homeostasis in schizophrenia, rather than more wide-ranging primary cardiometabolic dysfunction, as I have explained in Section 1.1.4.2.2. Interestingly, I found that the ALSPAC participants grouped into the 'persistently-high' fasting insulin developmental trajectory also had mean levels of triglycerides and HDL outside of reference ranges at age 24 years, providing further consistency to the results.

I did not find that ALSPAC participants grouped into the 'persistently-high' fasting insulin developmental trajectory had mean BMI levels or other forms of cholesterol outside of reference ranges. Also, I did not find evidence for associations of BMI developmental trajectories with schizophrenia-spectrum phenotypes at age 24 years. In fact, I found striking differences between the longitudinal cardiometabolic associations of schizophrenia-spectrum outcomes compared with depression, where the latter did show strong associations with puberty-onset BMI increases but no associations with glucose-insulin homeostasis. This provides evidence of specificity for primary disruption to glucose-insulin homeostasis, but not adiposity, with increased risk of psychosis. These findings are in line with meta-analyses of both individuals with FEP (Perry et al., 2016) and younger individuals at risk of developing psychosis (Carney et al., 2016) which did not find differences in BMI between cases and controls. Conversely, longitudinal studies conducted in large samples have found associations of lower BMI in childhood and adolescence with increased risk of schizophrenia in adulthood (Zammit et al., 2007, Weiser et al., 2004, Sorensen et al., 2006).

While these longitudinal studies are limited by only including single point-measures for BMI, the large population-representative samples permit significant statistical power to detect a difference. Therefore, additional subtle trajectories of BMI may exist in the population that could not be

accurately delineated in my analysis. For example, I found that over 70% of included participants were grouped into the 'stable average' BMI trajectory, whose BMI remained close to the sample mean over time. Replication of my work in larger samples may uncover additional BMI trajectories subtly distinct from the 'stable average' trajectory I identified. Some of these may be associated with lower childhood BMI and may, in turn, be associated with an increased risk of schizophrenia.

At first glance, it may appear contradictory that abnormalities in glucose-insulin homeostasis could be detectable in the absence of adiposity and other cardiometabolic phenotypes. However, there is increasing recognition that in the early stages of cardiometabolic disruption in young adults, insulin resistance can occur in isolation and in advance of changes to adiposity (Wiebe et al., 2021). This has been coined the 'lean insulin resistant' phenotype (Penesova et al., 2011, Townsend et al., 2018, George et al., 2015, Gonzalez-Cantero et al., 2018). Interestingly, the 'lean insulin resistant' phenotype is also associated with higher levels of inflammation (Ding et al., 2016), and I will discuss the potential mechanistic involvement of inflammation below (Section 8.2.2).

8.2.2 Testing Potential Mechanisms of Association between Cardiometabolic Disorders and Schizophrenia

In Chapter 3, using ALSPAC data, I found that genetic predisposition to T2D was associated with increased risk of psychosis in early adulthood and *vice versa,* genetic predisposition to schizophrenia was associated with insulin resistance in early adulthood. These findings indicate the possibility of gene similarity between schizophrenia and disrupted glucose-insulin homeostasis. These findings align with other observational genetics studies (Chouinard et al., 2019, Tomasik et al., 2019, Hackinger et al., 2018) as described in Section 3.5, and extend upon them since existing studies are limited by relatively small sample sizes compared with the analysis I present in Chapter 3.

Furthermore, I can extend upon the findings of previous studies since I tested a potential mechanism for the genotype-phenotype associations. I found that the association between genetic predisposition for T2D and risk of psychosis was partly mediated by childhood inflammation. Whilst the effect size for mediation was small, and those analyses may have been limited by statistical power, the findings suggest that the genotype-phenotype associations of T2D and schizophrenia align at least in part due to genetic influences on inflammatory and immune pathways, which could increase the risk of both disorders simultaneously. Indeed, there is biological plausibility for this mechanism; longitudinal associations between childhood inflammatory markers and subsequent risk of psychosis have been reported in the same (Khandaker et al., 2014) and other cohorts (Kappelmann et al., 2019, Goldsmith et al., 2019, Metcalf et al., 2017, Osimo et al., 2021). Similarly, longitudinal associations of

inflammatory markers with disrupted glucose-insulin homeostasis have been reported in meta-analyses (Bowker et al., 2020).

In Chapter 4, I took a different approach to examine for shared genetic liability between schizophrenia, cardiometabolic and inflammatory traits, using genomic methods that leverage summary data from large-scale GWAS. The findings were consistent with those reported in Chapter 3, thus strengthening the argument that shared genetic liability may at least in part explain phenotypic associations between schizophrenia, cardiometabolic and inflammatory traits. Limited previous research has sought to examine for shared genetic architecture between schizophrenia and cardiometabolic traits. For example, one recent study reported a negative genetic correlation between schizophrenia and BMI (Bahrami et al., 2020). Another older study that featured smaller GWAS reported limited evidence for genetic correlation between schizophrenia and cardiometabolic traits (Bulik-Sullivan et al., 2015a).

However, as described in Section 4.1, the LDSC approach may have limitations that I was able to address with the use of novel complementary analytic methods. In line with previous research, I found evidence for a negative genetic correlation between schizophrenia, BMI, and T2D, confined to relatively common genetic variants. This cardiometabolic risk-decreasing pattern of correlation with schizophrenia differed from the pattern I identified in relatively less-common genetic variants, where I found consistent evidence of a cardiometabolic risk-increasing pattern of correlation with schizophrenia. The heterogeneity of these findings requires further investigation but may help to explain why lower BMI in childhood is longitudinally associated with schizophrenia in adulthood (Zammit et al., 2007, Weiser et al., 2004, Sorensen et al., 2006), and may help to explain how a 'lean insulin resistant' phenotype may be associated with schizophrenia, as discussed in Section 8.2.1. These findings together suggest that obesity, which is commonly observed in chronic schizophrenia (Mitchell et al., 2013b), may occur due to lifestyle and iatrogenic factors (see Section 1.1.3) or may occur over time secondary to intrinsically disrupted glucose-insulin homeostasis in schizophrenia.

In addition to addressing limitations of the LDSC approach, findings from Chapter 4 also provide further granularity around potential mechanisms that may link schizophrenia, inflammation and cardiometabolic traits. The colocalization analysis returned robust evidence for several genetic loci that may underly the shared genetic liability between schizophrenia, cardiometabolic and inflammatory traits. Several loci are related to pathways involving BDNF, which has biologically plausible roles in the development and maintenance of the immune and central nervous systems, and in the regulation of cardiometabolic function (See Section 4.5).

The research methods employed in Chapter 4, including genetic correlation and colocalization analyses, cannot test the direction of association. However, in Chapter 5, I used MR to first examine for evidence that insulin resistance and related cardiometabolic traits may be causally related to schizophrenia, and second, for evidence that inflammation may be a common mechanism for schizophrenia and insulin resistance. Crucially, MR can examine the direction of association and can address problems of residual confounding. My findings indicate that inflammation may be a novel therapeutic target for both schizophrenia and its cardiometabolic comorbidity. Existing MR studies have reported evidence that inflammation may be causally related to schizophrenia (Hartwig et al., 2017) and T2D (Yuan and Larsson, 2020, Bowker et al., 2020) separately. Therefore, my findings are consistent with previous research.

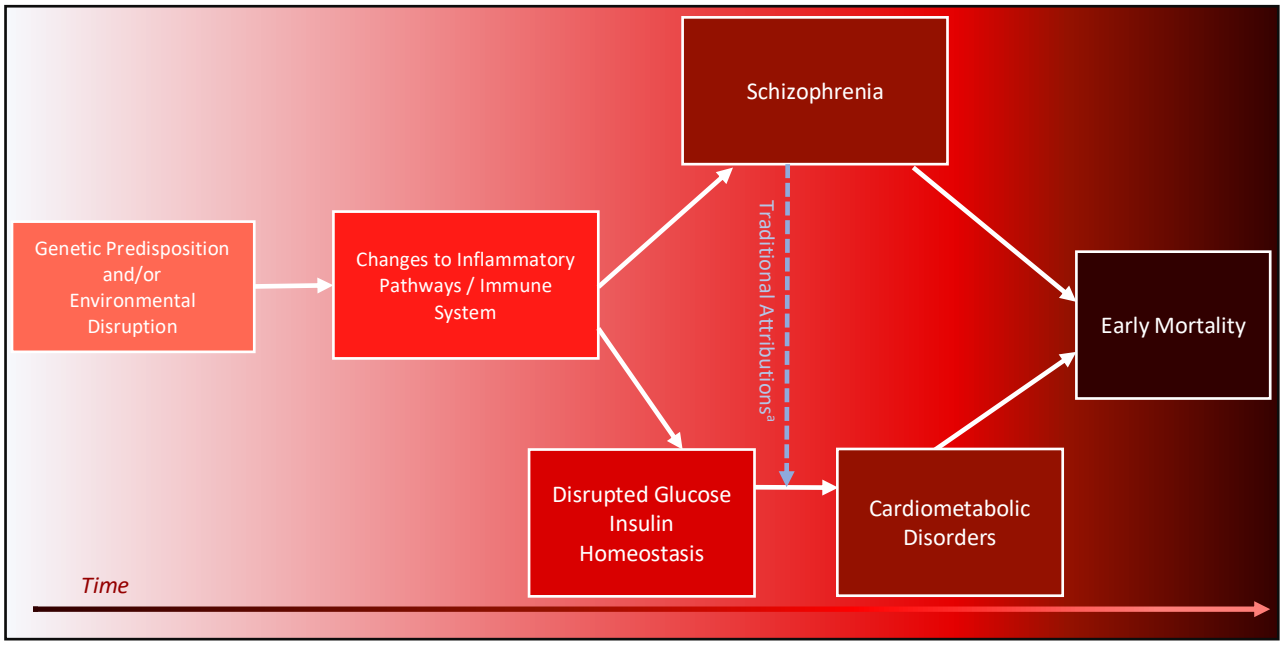
Where the findings from Chapters 3 & 4 showed how shared genetic liability might lead to inflammatory changes, disruption to glucose-insulin homeostasis, and increased risk of schizophrenia, findings from Chapter 5 show that genetic predisposition may not be the only mechanism for inflammation to exert a simultaneous influence on risk of schizophrenia and comorbid cardiometabolic disorders. In MR, while genetically predicted levels of the exposure are modelled, these are considered proxies for lifelong levels of environmental exposures free from measurement error or short-term environment-related fluctuations in the exposure (Davey Smith and Ebrahim, 2005). Therefore, findings from Chapter 5 indicate that increases in inflammation from any cause, whether genetic or environmental, could be potentially causally linked with schizophrenia and cardiometabolic disorders simultaneously.

Environmental adversity in early life through infection, stressful life events or malnutrition may permanently alter the immune system (Merlot et al., 2008, Harvey et al., 2010). This idea is consistent with Barker's developmental programming hypothesis (Barker et al., 1993), as described in Section 1.2.3. Indeed, in Chapter 2, I found that participants grouped into the 'persistently high' fasting insulin trajectory had significantly greater exposure to perinatal stressful life events and significantly lower birthweight compared with participants grouped into the 'stable average' trajectory. These associations may reflect an adverse early developmental environment, providing further evidence of the potential role of the developmental programming hypothesis in simultaneously increasing the risk of schizophrenia and cardiometabolic disorders.

The findings from Chapters 2-5, alongside existing research, can be framed together to delineate the most likely direction of association between genetically predisposed schizophrenia, cardiometabolic and inflammatory traits. For example, in Chapter 2, I found that disrupted glucose-insulin homeostasis may predate the onset of psychosis. In Chapter 3, I found that genetic predisposition may increase the risk of comorbid psychosis and disrupted glucose-insulin homeostasis, at least in part

due to a mediating influence of inflammation. Existing longitudinal research has shown that inflammation is likely to predate disrupted glucose-insulin homeostasis (Bowker et al., 2020) and psychosis (Khandaker et al., 2014). In Chapter 4, I found that the cardiometabolic traits colocalized with inflammation and schizophrenia were broader than those solely related to disrupted glucose-insulin homeostasis, for example, involving distal cardiometabolic endpoints such as CAD. Given that disruption to glucose-insulin homeostasis predisposes to CAD (Aronson and Edelman, 2014), CAD is likely to be a downstream colocalized trait from inflammation and disrupted glucose-insulin homeostasis. In Chapter 5, I found that lifelong levels of inflammation, either genetic or environmental, may be a common cause for comorbid insulin resistance and schizophrenia. Therefore, the most likely direction of association between schizophrenia, cardiometabolic and inflammatory traits is shown in Figure 29.

Figure 29: Schematic Outlining the Most Likely Direction of Association between Inflammation, Cardiometabolic Disorders and Schizophrenia Based Upon Findings of This Thesis



[a]Traditional attributions include sociodemographic, lifestyle and iatrogenic factors and are described in further detail in Section 1.1.3

8.2.3 Improving the Prediction of Cardiometabolic Risk in Young People with Psychosis

Chapters 2-6 show a consistent thread of evidence showing that schizophrenia is likely to carry inherent cardiometabolic risk, which may be first detectable before the onset of FEP. The findings suggest that the commonly attributed lifestyle and clinical factors (as described in Section 1.1.3) are likely to be exacerbating rather than causal factors for the comorbidity. Therefore, there is a clear and crucial need for tools that can accurately quantify this combined inherent and exacerbated

cardiometabolic risk at the soonest possible opportunity in the schizophrenia illness course. Such tools can assist healthcare professionals in preparing personalized treatment plans by accurately considering present and future cardiometabolic risk. This can help to attenuate the risk of more distal adverse cardiometabolic outcomes and close the substantial mortality gap faced by people with schizophrenia.

I identified an extensive array of existing cardiometabolic risk prediction algorithms in my systematic review (Chapter 6). However, the majority have not been externally validated, and most were rated as being at high risk of bias. Prediction algorithms that cannot evidence potential generalizability and have been reported so poorly cannot be clinically useful and so are arguably no more than contributors to research waste. This opinion is not new; an older systematic review of cardiometabolic risk prediction algorithms came to a similar conclusion (Damen et al., 2016). It is disheartening that despite the introduction of reporting standards (Collins et al., 2015) and risk of bias assessment tools (Wolff et al., 2019) for risk prediction algorithms, improvements in the literature are yet to materialize.

However, I did identify a few excellent examples of cardiometabolic risk prediction algorithms. QRISK3, QDiabetes and PRIMROSE were developed in extensive samples and were rated as relatively low risk of bias. Each was validated in large samples and so have evidenced generalizability; QRISK3 has been successfully integrated into clinical practice in the UK, a step most health outcome prediction algorithms fail to reach (Riley, 2019); and, PRIMROSE was the only identified algorithm to have a published economic analysis (Zomer et al., 2017).

Despite these positives, I identified several reasons why all three algorithms are unlikely to be suitable for young people with psychosis. These included the older populations the algorithms were developed for, the balance of predictor weightings, and the character and coding of included predictors. In the exploratory validation analysis, I found that all three algorithms substantially underpredicted cardiometabolic risk in young people who had or were at risk of developing psychosis.

Therefore, I developed and externally validated PsyMetRiC in real EIS samples of young adults (Chapter 7), following TRIPOD reporting guidelines (Collins et al., 2015). I found that PsyMetRiC showed good predictive performance in the development, external validation and sensitivity analysis samples, suggesting the algorithm is likely to be suitable for use in the UK EIS population. By involving a young person's advisory group in the design of PsyMetRiC, I ensured that the algorithm is likely to be acceptable for patients. By developing two PsyMetRiC versions, one with- and one without biochemical measures, I ensured that the algorithm is likely to be clinically useful.

While not yet bedside ready, I see PsyMetRiC as a useful starting point that can be taken forward towards routine clinical practice after further validation and fine-tuning. An important consideration that I purposely did not broach with PsyMetRiC is with prescribed risk score 'cut-offs' and associated treatment recommendations. Primarily, this is because I do not believe that algorithms could, or should, entirely dictate clinical decisions, which are complex assimilations of individual patient factors that are unlikely to be fully captured by any algorithm. Rather, I believe that prediction algorithms should more appropriately be placed in the context helping to inform the decision-making process. Nevertheless, while cut-offs are often helpful in healthcare settings, deciding upon them in the context of PsyMetRiC will require a separate body of multi-disciplinary and patient-centred work.

8.3 Strengths and Weaknesses of the Methodological Approaches Used in This Thesis

In this thesis, I have used a number of complementary approaches to examine the nature and mechanisms of the cardiometabolic comorbidity of schizophrenia and consider the prediction of cardiometabolic risk in young people with psychosis. Each method was selected for its potential strengths in being able to address the research questions posed and for its ability to address the limitations of previous research. However, each analysis I have conducted may have weaknesses that must be taken into consideration. I will now address the strengths and weaknesses of each analysis I have conducted in order of their presentation in this thesis.

8.3.1 Strengths and Weaknesses of the Methodological Approaches Used in Section B

The availability in ALSPAC of repeat measures of cardiometabolic indices to delineate trajectories of cardiometabolic development and test associations with schizophrenia-spectrum outcomes (Chapter 2) is a key strength. ALSPAC is a relatively large population-representative birth cohort and features a highly diverse range of collected data spanning biochemical, sociodemographic, anthropometric, genetic, psychiatric, and lifestyle data. Such a detailed set of available data permitted a thorough analysis of two important features of cardiometabolic development through childhood and adolescence. In my study, I included 12 measures of BMI between ages 1-24 years and four measures of fasting insulin between ages 9-24 years. Comparing my study with those included in a relatively recent systematic review of BMI developmental trajectories (Mattsson et al., 2019), my study features the longest temporal analysis period and the most extensive set of repeat measurements. To the best of my knowledge, my study also includes the first analysis of developmental trajectories of fasting

insulin through childhood and adolescence. In addition, given the richness of the ALSPAC dataset, I performed a detailed analysis of the identified trajectories. This included examining sociodemographic and lifestyle predictors of trajectory membership, clinical characteristics of the identified trajectories at age 24 years, and associations with schizophrenia-spectrum and depression outcomes at age 24 years. For the latter analysis, I adjusted for a detailed set of potential confounders, including sex, ethnicity, paternal social class, childhood emotional and behavioural problems, and cumulative scores of smoking, physical activity, alcohol use, substance use, sleep problems and average calorie intake. Such detailed confounding adjustment is rare for existing studies in the field and provides some confidence that the associations are unlikely to be fully explained by sociodemographic and lifestyle factors.

The data available in the ALSPAC cohort permitted a detailed consideration of several key aspects of the Bradford Hill criteria (Hill, 1965), which has thus far eluded existing studies in the field. First, I could examine the direction of association, where most existing research has been either cross-sectional or has included incident cases of psychosis. Evidence of longitudinal associations is key to unravelling pathophysiology and identifying genuine risk factors. Second, by including several related schizophrenia-spectrum outcomes, I could test for evidence of consistency and robustness of the results. Third, by including depression as an outcome, I could test for evidence of specificity of association. Fourth, I could examine a biological gradient related to both the exposure and outcomes. For example, my analysis delineated two adverse fasting insulin trajectories, but the strongest evidence of association was found for the most adverse fasting insulin trajectory. I also found the strongest evidence for an association with the most clinically relevant schizophrenia-spectrum outcomes, namely ARMS and psychotic disorder.

The use of repeat measure data also allowed me to overcome another key limitation of previous studies, which have typically included one-off measurements of cardiometabolic indices and so are blind to fluctuations over time. Cardiometabolic indices, including measures of glucose-insulin homeostasis (Moebus et al., 2011) and BMI (Turicchi et al., 2020), are subject to normal fluctuation. This variability cannot be appropriately considered with single-point measures. Repeated measures over time permitted a more granular and detailed examination of underlying biological mechanisms taking into account dynamic temporal changes in these indices.

The use of GMM as an analytical approach for the repeat measure data permitted the capture of information about interindividual differences in intraindividual change, taking into account unobserved heterogeneity within a larger population (Jung, 2007). On the other hand, regression-based modelling assumes that the growth trajectories of all individuals in a population could be adequately described using a single estimate of growth parameters; i.e., all individuals are drawn from

a homogenous population without discernible differences (Jung, 2007). This is most likely an oversimplification. Since GMM relaxes these assumptions and allows differences in growth parameters across unobserved subpopulations, I could delineate subgroups of the population likely to be plausibly different from one another in their trajectories of fasting insulin or BMI levels. Therefore, GMM is likely to be a more biologically plausible framework to consider cardiometabolic development than more standard regression-based methods.

Weaknesses of the ALSPAC data-based studies (Chapters 2 & 3) include missing data. As is common in most, if not all cohort studies, attrition has also affected ALSPAC, with males and those from more disadvantaged backgrounds more likely to have been lost to follow-up (Boyd et al., 2013a). Systematic differences in attrition can lead to selection bias in the analytic sample, which can bias results in either direction with a magnitude of impact that can be difficult to ascertain (Odgaard-Jensen et al., 2011). In addition, selective sampling can increase the risk of collider bias when the selected and missing samples differ on a variable which may be on the putative causal pathway analysed (Cole et al., 2010). Given that most psychiatric and cardiometabolic disorders are strongly associated with social class and that participants from lower social classes were disproportionately lost to follow-up in ALSPAC (Boyd et al., 2013a), collider bias may have affected the results.

While the GMM approach could address missing data in the delineation of cardiometabolic developmental trajectories using FIML, analyses involving psychiatric outcome data would have been most susceptible to missing data bias. For example, the analytic sample following confounding adjustment featured as low as 28% of the total sample for BMI-psychiatric outcome analyses and as low as 47% for fasting insulin-psychiatric outcome analyses.

Methods of addressing missing data do exist, such as multiple imputation, and when used carefully, are effective at reducing the impact of bias from missing data (White et al., 2011). However, I could not use multiple imputation in this study since I analysed trajectory-psychiatric outcome associations using the three-step GMM method (see Section 2.3.5.2). While the three-step method was methodologically appropriate since it considers classification uncertainty, adding multiple imputation to this process would have caused prohibitive computational burden. The field has yet to feasibly combine multiple imputation with the three-step method (Asparouhov, 2014). Therefore, given the risk of selection bias in my analyses, it cannot be concluded that the results I have obtained are generalizable to the whole population of young people with psychosis, and so replication of my findings is crucial.

Since my study was observational, I cannot confirm that residual confounding has not affected my findings. This issue was addressed in a separate study using MR (Chapter 5). Residual confounding

is a limitation of all observational research. Whilst detailed confounding adjustment is helpful to reduce the impact of confounding on the results, addressing the impact of any potential confounder relies on the assumption that the confounder has been accurately measured. This may not have been the case with ALSPAC data in some instances.

For example, socioeconomic inequality is a powerful predictor of health outcomes (Kivimaki et al., 2020), but measuring it is complex (Darin-Mattsson et al., 2017). Inequality is not static and temporal changes in socioeconomic factors can lead to measurable differences in health outcomes over time (McKenzie et al., 2014). I chose paternal social class at birth to represent socioeconomic status since it a good predictor of childhood outcomes (Erola, 2016). However, since my exposures were measured longitudinally through childhood and adolescence, I could not capture the potential impact of social mobility over time (Tiikkaja et al., 2013).

In addition, a number of the confounders I adjusted for were based on self-report data collected from questionnaires. Self-report health data is at notoriously high risk of measurement error (Butler, 1987). For example, I adjusted for calorie intake based upon data collected from food frequency questionnaires, but such data is at high risk of recall bias (Natarajan et al., 2010, Freedman et al., 2011). Furthermore, I adjusted for physical activity based upon self-reported questionnaire data, but past physical activity levels are commonly misreported in the general population (Lim et al., 2015) and even more so in people with psychosis (Firth et al., 2018). Whilst accelerometer data may be a more accurate objective measure of physical activity (Dyrstad et al., 2014), the sample size with available accelerometer data in ALSPAC was relatively small. Therefore, I used the self-report variable to maximise the available sample size.

Residual confounding may also have affected my results. This could have occurred due to known confounders I could not include or confounders that are unknown. For example, I could not adjust for psychological stress and the associated impacts upon the HPA axis (Smith and Vale, 2006) since cortisol data were available only for a small sub-section of the cohort at a single time-point. Unknown confounders by definition cannot be adjusted for but can still impact the results of observational analyses.

Another limitation is the possibility of reverse causality, which I also addressed by using MR in Chapter 5. Longitudinal research can only demonstrate the direction of association if it can be confirmed that the outcome did not occur before the exposure. Simply, it cannot be proven that the exposure occurred before the outcome just because the variables were measured in that order. In my study, the earliest point of assessment of fasting insulin was age 9 years, and there was no corresponding data on schizophrenia-spectrum outcomes at, or before this age. Therefore, the risk of

reverse causality cannot be discounted entirely. Despite that, psychosis in pre-pubertal children is rare (McClellan and Werry, 1997), and so the risk of reverse causality in this instance is small.

Other limitations relate to the statistical approaches I used in the study. First, the identified trajectories are a statistical phenomenon and not necessarily a biological one. Therefore, care must be taken in the interpretation of the results and extrapolation to external populations. Replication of my work in larger samples will increase confidence in the biological plausibility of the identified trajectories. Second, the statistical approach I used to examine associations of identified trajectories with psychiatric outcomes is grounded in examining average group-level differences. Therefore, the findings from my study cannot be extrapolated to the individual. For example, whether a raised fasting insulin level in childhood could predict adult psychosis cannot be determined from my results. Given that there is growing interest in developing prediction algorithms for transition to psychosis (Montemagni et al., 2020), my results suggest that fasting insulin levels in childhood could be a suitable candidate predictor. Nevertheless, this must be explored formally using appropriate prognosis research methods, such as those I employed in Chapter 7.

Finally, another important limitation that was unavoidable in my study was in the measurement of schizophrenia-spectrum outcomes. ALSPAC does not yet have health-record linkage and does not have data on whether participants met the criteria for an ICD or DSM diagnosis of schizophrenia or related psychoses. For example, PEs do not exclusively represent psychosis-risk and are associated with other mental disorders, including anxiety and depression (Varghese et al., 2011). Nevertheless, my chosen outcomes are likely to lie along the continuum of the schizophrenia spectrum, and the psychotic disorder outcome would likely meet a clinical threshold for the consideration of monitoring and treatment, and so is clinically relevant.

8.3.2 Strengths and Weaknesses of the Methodological Approaches Used in Section C

In Chapter 3, where I examined associations of genetic predisposition for T2D and risk of schizophrenia-spectrum outcomes in early adulthood, and *vice versa*, a primary strength of the study relates to the relatively large sample size and richness of the ALSPAC dataset. This is discussed above in Section 8.3.1.

In addition, in Chapter 3, I was also able to address several important features of the Bradford-Hill criteria (Hill, 1965). First, I used genotype as an exposure, and this is set at conception. Therefore, a clear direction of association is evidenced without the possibility of reverse causality. Second, by including several schizophrenia-spectrum outcomes, I was able to evidence consistency in findings both internally within the study and externally with previous research (Chouinard et al., 2019,

Hackinger et al., 2018, Ferentinos and Dikeos, 2012). Third, I found the strongest evidence for associations of genetic risk for T2D in the most clinically relevant schizophrenia-spectrum outcome, psychotic disorder. Therefore, my results provide evidence of a dose-response relationship toward a clinically relevant schizophrenia-spectrum outcome. Fourth, where previous research was unable to consider potential mechanisms for genotype-phenotype associations, I was able to test a hypothetical mechanism that genetic influences on inflammation may mediate associations between genetic predisposition for T2D and schizophrenia-spectrum outcomes and *vice versa*. I also performed a sensitivity analysis by removing a BMI-related T2D SNP and thus a potentially pleiotropic mechanism involving adiposity. In completing these steps, I was able to provide evidence of biological plausibility, which is an important criterion of the Bradford-Hill criteria.

In addition, findings presented in Chapter 3 are at low risk of residual confounding since confounding of genotype-phenotype is unlikely. However, there is evidence that GWAS can be affected by factors such as ethnicity (Huang et al., 2015), social class (Morris et al., 2020) and even voluntary study participation (Tyrrell et al., 2021). These factors relate to population stratification and can bias GWAS results (Hellwege et al., 2017). In ALSPAC, quality control measures included filtering participants of non-European ancestry to reduce the impact of confounding by ethnicity (see Section 3.3.4). In addition, I also adjusted all regression analyses for the first ten principal components, which reduces the risk of population stratification bias (see Section 3.3.4), and adjusted for ethnicity and social class (see Section 3.3.6). Therefore, while the risk of confounding by population stratification is possible in my analyses, I took appropriate steps to minimize it.

Regarding the weaknesses of the genotype-phenotype analysis (Chapter 3), as described above in Section 8.3.1, a primary unavoidable limitation of ALSPAC data is attrition. I performed a missing sample analysis and found that the missing sample were more likely to be male and from a lower social class. I also found that the missing sample had a higher mean score for PRS-schizophrenia but a lower mean score for PRS-T2D. Whilst attrition is common to cohort studies, it presents a challenge to analyses of GWAS data since it may introduce population stratification bias. Thus, my analyses may underestimate a true association of genetic predisposition for schizophrenia with insulin resistance. In contrast, the opposite might be the case for the association of genetic predisposition for T2D with schizophrenia-spectrum outcomes.

In addition, whilst I was able to address several key features of the Bradford-Hill criteria (Hill, 1965), some of the associations from my analyses were relatively weak and unlikely to meet the effect size criterion. For example, I only found weak evidence for an association of genetic risk for schizophrenia with insulin resistance. Also, I found partial mediation by inflammation for the association of genetic risk for T2D with PEs at age 18 years but not *vice versa.* Statistical power may be one potential

contributor. While my sample was larger than samples used in previous research (Chouinard et al., 2019, Hackinger et al., 2018, Ferentinos and Dikeos, 2012), mediation analyses typically require relatively large sample sizes, and studies employing tests of mediation are commonly underpowered (Fritz and Mackinnon, 2007). In addition, prospective analyses of genetic data are also commonly underpowered (Chanock et al., 2007), predisposing to an increased risk of type II statistical error (Hong and Park, 2012). In future, replication of my work in a larger sample will help to clarify the findings.

Further, I used psychotic experiences and operationally defined psychotic disorder rather than diagnoses of schizophrenia according to ICD or DSM as outcomes. I have described this limitation, including the relevance of these outcomes for schizophrenia in further detail in Section 8.3.1. This issue is particularly pertinent to analyses in Chapter 3. Since outcomes in this study were assessed at age 18 years, before the peak age of incidence of schizophrenia (Eranti et al., 2013, Castle et al., 1998), some genuine cases of psychosis may have been missed.

A limitation common to most analyses of existing GWAS data is that most GWAS have been conducted in either solely or mostly European samples. Therefore, it cannot be known whether the findings are relevant for non-European populations. This is a significant limitation of analyses of genetic data and their real-world relevance because non-Europeans contribute a substantial proportion of the global burden of schizophrenia and cardiometabolic disorders. While GWAS of non-European populations are growing in size and breadth, there remains much to do to achieve parity (Sirugo et al., 2019, Popejoy and Fullerton, 2016).

Finally, another limitation common to GWAS is that they typically only measure common genetic variation. For example, current GWAS efforts can explain only a fraction of the heritability of schizophrenia (Lee et al., 2013) and T2D (Billings and Florez, 2010). This suggests that a notable proportion of genetic liability remains undiscovered, possibly through large numbers of rare variants which individually contribute a small effect (Manolio et al., 2009). Therefore, this limits the power of analyses of GWAS data to detect genetic effects, and this 'selection bias' of more common genetic variants may increase the risk of both type I and type II error in secondary analyses. Encouragingly, studies such as the UK Biobank are soon to release data from the whole-genome sequencing of exceptionally large samples. This may lead to an improvement in the proportion of the heritability of human disease explained by genetic variation and may lead to improved PRS. Replication of my analysis in future, when more accurate and complete genetic data are available, will be helpful.

Regarding Chapter 4, where I used summary GWAS data to examine for potential genetic overlap between schizophrenia, cardiometabolic and inflammatory traits, a key strength of the study is in the

sample sizes available for analysis, which also helps to at least in part to address the potential power limitation of Chapter 3. While in Chapter 3 I could include less than 4,000 participants in total, the sample sizes of the GWAS included in analyses in Chapter 4 were between 42,854 and 898,130, providing a considerable increase in statistical power.

In addition, previous studies which have sought to examine for genetic overlap between schizophrenia and cardiometabolic traits from summary GWAS data have used the LDSC approach, yet the LDSC approach may have limitations, including a) downward bias of the effects of lower-frequency variants; b) opposing mechanisms; c) a lack of context with which one might consider biological plausibility, or indeed, distinguish potential causality from correlation. These limitations of previous research are described in further detail in Section 4.1. In Chapter 4, I sought to address these limitations by using a complementary set of independent methods which are better suited to examining genetic overlap between traits after considering the limitations of the LDSC approach. I was also able to include both cardiometabolic and inflammatory traits in the analysis to further test the hypothesis that inflammation may be mechanistically implicated in the associations of schizophrenia and cardiometabolic disorders. The convergent thread of evidence from the independent methods and consistency with the results from Chapter 3 helps to provide confidence in the study conclusions. Finally, I included several cardiometabolic traits at different ends of the spectrum of chronicity. For example, biochemical measures such as fasting insulin, HDL and triglycerides are likely to be adversely affected long before distal traits such as T2D or CAD are diagnosed. I found that clusters of cardiometabolic and inflammatory traits at varying levels of chronicity were correlated and colocalized with schizophrenia. Specifically for colocalization analysis, the greater the number of colocalized traits at a specific locus, the stronger the evidence for colocalization (Foley et al., 2021).

Weaknesses of the analysis presented in Chapter 4 can be divided into those arising from the GWAS samples analysed and those arising from the statistical methods and results. Regarding the weaknesses of the GWAS samples, there is a risk that as sample sizes increase, specificity toward the trait intended to be measured decreases. This phenomenon has been elegantly demonstrated in GWAS analyses of depression (Cai et al., 2020) and may also apply to schizophrenia. For example, I used the largest published GWAS for schizophrenia (Pardinas et al., 2018) to maximise statistical power. However, the predominant contributor to the larger sample of that GWAS, compared with previous schizophrenia GWAS, was clozapine treated patients. This is likely a result of convenience sampling since clozapine-treated patients receive regular blood tests, which might be readily analysed for genotype. Nevertheless, treatment resistance is a primary requirement for clozapine treatment. Some have hypothesized that non-response to antipsychotics may be a marker of a distinct subtype of

schizophrenia (Farooq et al., 2013, Gillespie et al., 2017), which might have distinct polygenicity compared with treatment-responsive schizophrenia (Vita et al., 2019, Pisanu and Squassina, 2019). Therefore, the larger GWAS sample size may have come at the cost of increased heterogeneity, which could impede research aiming to examine potential disease mechanisms (Cai et al., 2020).

Related to this, as GWAS samples increase, the granularity of the measured trait often decreases. For example, in the analysis of prospective ALSPAC data (Chapter 2), I examined associations of positive and negative psychotic symptoms separately. Yet, such granularity is not possible in secondary analyses of GWAS datasets. In addition, the GWAS of biochemical/anthropometric traits were mostly conducted based on a single point measurement of the trait. Yet, as described in more detail in Section 8.3.1, cardiometabolic and inflammatory traits are subject to normal fluctuations, which may not be adequately addressed from a single point measurement. Therefore, since the original GWAS methods of analysing the biochemical/anthropometric cardiometabolic and inflammatory phenotypes may be subject to measurement error, the GWAS summary estimates derived from these studies may also be affected.

Relatedly, all included GWAS featured adult participants. Yet, adverse cardiometabolic function is more common with increasing age due to the chronicity of most lifestyle factors (See Section 1.1.1.2). Therefore, measurement of cardiometabolic traits in adulthood may be subject to confounding. For example, it may not be possible to distinguish between the effect of a genetic variant on BMI directly from the effect of that genetic variant on smoking behaviour, which could in turn influence BMI. In addition, it is also not possible to determine whether the results apply outside of adult populations. This has been clearly demonstrated in the case of BMI, where genetic variants associated with childhood obesity show only partial overlap with those associated with obesity in adulthood (Vogelezang et al., 2020).

Further, summary data from GWAS is relatively inflexible, and prospective adjustment for factors such as social class, which I adjusted for in the prospective genetic analysis (Chapter 3), is not possible. Finally, limitations of GWAS discussed above in relation to Chapter 3 are also likely to apply to summary data from GWAS studies. For example, GWAS currently measure only common genetic variation, and the findings are only likely to be relevant for European populations.

Other limitations of Chapter 4 relate to the analytic methods of the summary data and the study results, which may limit firm conclusions from the analyses. First, the analytic methods used in Chapter 4 cannot elucidate the direction of association. This can only be inferred from research using other methods, both genetic and observational, that I have used elsewhere in this thesis.

Second, secondary analysis of GWAS data may be biased when there is sample overlap between analysed traits. It is commonly assumed that when data are derived from different GWAS consortia, the risk of sample overlap is small (Shi et al., 2017). However, a finite pool of individuals have consented to genotyping. As GWAS sample sizes increase, one might argue that even between different GWAS consortia, some level of sample overlap is possible. This may be particularly relevant for GWAS of binary traits such as T2D or schizophrenia, which require samples of healthy participants to act as controls.

Third, while LDSC provides estimates of shared genetic heritability between trait pairs with which one could infer the real-life relevance of potential genetic overlap, many of the complementary methods I used in Chapter 4 do not provide these estimates. One could argue that some level of genetic overlap may be expected by chance between complex phenotypes, and so the clinical relevance of my findings cannot be ascertained. Nevertheless, I found convergent and consistent evidence from a range of independent statistical methods in Chapters 3 and 4, fostering confidence that the results are unlikely to have occurred by chance and are therefore likely to have clinical relevance.

Fourth, despite the considerable increase in power when using summary data from GWAS, some of the results described in Chapter 4 were relatively weak. For example, a few findings in the MAF-stratified analysis did not reach the Bonferroni-corrected evidential threshold, although this could be explained by the presence of opposing mechanisms (Shi et al., 2017), which is described in more detail in Section 4.1. In future, further refinement of analytic methods coupled with better powered GWAS (and whole-genome sequenced datasets) will help to clarify the results I have presented.

Fifth, at present, there is a relative dearth of large-scale publicly available GWAS data for inflammatory markers beyond CRP. This meant CRP was the sole inflammatory trait that I could include in my analyses. Since CRP is a downstream and relatively generalized inflammatory marker, my analyses are limited in being able to elucidate a deeper mechanistic understanding of the broader constellation of inflammatory changes underlying the genetic correlation and colocalization findings involving CRP.

In Chapter 5, I used MR as a methodological approach, which uses as input large-scale GWAS datasets, imparting significant statistical power to the analysis. In addition, MR has several key methodological strengths. When the assumptions for valid instrumental variables are met, MR can evidence direction of association free of residual confounding. These are two key criteria for establishing genuine risk factors of a disease outcome, and the ability to address residual confounding is not possible with observational research. MR can achieve these aims because it analyses genetic variants inherited randomly at conception as unconfounded proxies of a modifiable exposure, to

examine whether that exposure may have a causal effect on a disease outcome (Smith, 2010). Whilst MR findings in isolation cannot prove causality, they can be assimilated with a coherent body of observational and experimental evidence, which can together imply likely causality.

As well as bi-directional two-sample MR, in Chapter 5 I employed a number of extensions to the method, and each has its own inherent strengths. Firstly, from a hypothesis that inflammation may be a shared mechanism for schizophrenia and its cardiometabolic comorbidity, I considered that genetic variants which influence levels of both cardiometabolic and inflammatory indices could represent a specific biological mechanism that could be associated with schizophrenia. While pleiotropy could invalidate MR results between an exposure and an outcome depending on whether it is horizontal (confounding) or vertical (mediating) (Verbanck et al., 2018), I tested associations of 'inflammation-related' cardiometabolic variants with schizophrenia to show that inflammation-related pleiotropy is likely to affect MR associations between cardiometabolic traits and schizophrenia. These findings implicate inflammation as a common causal mechanism for the comorbidity.

Second, I used a detailed set of sensitivity analyses that can help to test the assumptions of MR. These included the Cochran Q test for SNP heterogeneity, the MR Egger regression intercept test and MR-PRESSO for horizontal pleiotropy, and the I^2_{GX} statistic for measurement error. In conducting these sensitivity analyses, I found decreased levels of heterogeneity and horizontal pleiotropy in inflammation-related genetic variants, compared with all related cardiometabolic variants. This suggests that inflammation-related cardiometabolic variants are likely to be closer to a specific biological mechanism, further supporting my hypothesis. Third, I used MVMR, a methodological extension to MR that tests associations of genetically predicted levels of an exposure on an outcome, after conditioning on the genetic associations with another exposure(s). Put simply, MVMR can examine pleiotropic mechanisms that could explain a univariable association, much like including covariates in observational study regression models. I leveraged the MVMR approach to further evidence that inflammation could be a common mechanism for comorbid insulin resistance and schizophrenia.

However, there are a number of potential limitations of the MR approach. First and foremost, as mentioned above, MR can provide evidence of the direction of association free of residual confounding *if the assumptions for valid instrumental variables are met*. I used a varied set of MR methods to help probe the assumptions for MR. These included IVW (which assumes all genetic variants satisfy MR assumptions); weighted median (which can produce accurate results so long as 50% of the selected genetic variants satisfy MR assumptions), and MR Egger (which can produce accurate results even if all genetic variants are subject to pleiotropy, as long as the size of the pleiotropic effect is independent of the size of the genetic variants' effects on the exposure (Bowden

et al., 2015)). I also used a detailed set of sensitivity analyses as described above to help test the MR assumptions. Despite taking these steps, proving that the assumptions for valid MR analysis have been met is near-impossible.

For example, one fundamental assumption is that the selected genetic variants must be associated with the exposure. This is perhaps the assumption with which one can have the most confidence in my study since the genetic variants were selected as being strongly associated (at the genome-wide level) with the exposures from large-scale GWAS. Nevertheless, as I have described above in Section 8.3.2, the secondary use of GWAS data presents its own challenges, including confidence in gene-exposure associations due to the risk of population stratification; the risk of confounding by chronic lifestyle factors for cardiometabolic indices measured in adulthood; heterogeneity of the analysed trait which may not be captured in GWAS; and, the applicability of GWAS results to non-European samples.

Another key assumption is that the selected genetic variants must influence levels of the exposure *directly* and not via an alternate mechanism. In most instances, this is extremely difficult to prove. Previous MR studies have attempted to address this challenge by restricting selected genetic variants to those located close to known gene coding regions (for example, the *IL6R* and *CRP* genes for IL-6 and CRP, respectively (Hartwig et al., 2017)). This is because genetic variation in the coding region of the exposure is more likely to affect the exposure directly rather than through alternate mechanisms. Nevertheless, this methodological step is not a panacea and can only prove the assumption is met if a complete biological understanding of how the genetic variant influences the exposure is known. This is often not the case.

The final key MR assumption is that genetic variants must only influence the outcome through effects on the exposure of interest, and not via any other mechanism. This is perhaps the most challenging assumption to prove in MR studies, particularly when a complete biological understanding of how an exposure influences an outcome is not known. In future, evidence from MR studies will require detailed examination in experimental and animal model research to help add to our mechanistic understanding of how an exposure influences an outcome. This can in turn help to evidence whether the assumptions for valid MR analysis had been met. Nevertheless, it was the violation of this assumption that I aimed to leverage to test the hypothesis that inflammation may be a common mechanism for comorbid cardiometabolic disorders and schizophrenia. The evidence I present in Chapter 5 suggests that this assumption is likely to be violated in MR studies examining cardiometabolic traits and schizophrenia due to a common biological mechanism involving inflammation.

Another potential limitation of MR is that the exposures modelled in MR studies represent lifetime changes in the level of the exposure rather than, for example, short-term changes in the exposure during a critical developmental period, as per Barker's developmental programming hypothesis. While MR evidence could be consistent with the developmental programming hypothesis since early-life disruption could permanently alter biological mechanisms, it is also possible that changes in the levels of certain exposures may only be strong risk factors for a disease in a specific developmental period. MR may not be able to capture this 'critical period' effect. In addition, that MR approximates lifetime changes in the level of an exposure has been cited as a possible reason why MR studies of potential therapeutic targets often overestimate the experimental treatment effect observed in clinical trials. This is because treatments are generally not prescribed over a lifetime (Gill et al., 2021). Whilst this limitation does not directly affect the conclusions I have made in Chapter 5, it is important to consider that the interpretation of MR findings is not straightforward, and evidence from MR studies requires triangulation with experimental, observational and animal model evidence in order to be most impactful.

8.3.3 Strengths and Weaknesses of the Methodological Approaches Used in Section D

There are several strengths to the systematic review I performed in Chapter 6, where I sought to establish whether existing cardiometabolic risk prediction algorithms could be suitable for young people with psychosis. First, while only meta-analyses of RCTs feature at the top of the evidence hierarchy, any well-designed systematic review can provide a valuable summary of current research. An earlier systematic review of cardiometabolic risk prediction algorithms was conducted in 2016 but did not consider suitability for young people with psychosis (Damen et al., 2016). My study updates the findings of the previous review because many additional algorithms have been published since 2016, and also considers the suitability of existing algorithms for a different population.

Second, I designed my search strategy to be as inclusive as possible, incorporating algorithms developed for the general and psychiatric populations. I also limited the risk of publication bias in my review by including conference abstracts, theses, and pre-prints. Together, this meant that I could include over 100 algorithms in my review, allowing a rich examination of potential suitability for young people with psychosis.

Third, I followed the state-of-the-art for risk of bias and quality appraisal via the relatively new PROBAST tool (Wolff et al., 2019), and followed the PRISMA guidelines for the conduct and reporting of a systematic review (Moher et al., 2009). These guidelines are validated, expert-consensus driven and form the basis of a high-quality review.

Fourth, while I was unable to follow a meta-analytic approach in my review, I substituted this important quantitative component with an exploratory validation analysis of three favourable algorithms from the systematic review. In doing so, I was able to present a consistent message across both the findings of the narrative results synthesis and the quantitative validation analysis, showing that existing cardiometabolic risk prediction algorithms are unlikely to be suitable for young people with psychosis.

The systematic review also has some weaknesses. First, as mentioned above, the research question posed and the heterogeneity of included studies prevented a meta-analytic synthesis. Narrative synthesis can increase the risk of reporting bias and can limit the validity of conclusions (Campbell et al., 2020). I addressed this by performing a quantitative analysis using ALSPAC data, which meant that my conclusions were formed from both narrative and quantitative results. Nevertheless, my quantitative analysis, which consisted of three separate validation analyses, cannot be equated with a formal meta-analytic quantitative analysis.

Second, since I did not follow a meta-analytic approach, I could not quantitatively assess the risk of publication bias. While I did include conference abstracts, theses and pre-prints in my review, these were few in number compared with peer-reviewed research papers. Therefore, publication bias is likely to have affected the mix of studies included in my review. For example, relatively few externally validated algorithms were included in my review. This may be because external validation requires data from a second distinct population, which is often not possible. However, it could also be because external validation performance estimates are usually less favourable than internal validation performance estimates. Therefore, some externally validated algorithms may have been deemed a lower priority for publication by journal editors.

Third, there is growing interest in risk prediction algorithms for health outcomes (Riley, 2019). Indeed, a large proportion of studies included in my review were published in the last few years. Therefore, it is likely that further cardiometabolic risk prediction algorithms have been published since the date my search concluded. This is a limitation of most systematic reviews. Interestingly, The BMJ currently features a ‘living’ systematic review of coronavirus risk prediction algorithms, which is updated regularly with newly identified studies (Wynants et al., 2020). The review already includes over 200 studies, and this is highly likely to increase further. The ‘living’ nature of that review is a notable feat but is not feasible for a PhD conducted over a finite period.

In Chapter 7 I developed PsyMetRiC, the first cardiometabolic risk prediction algorithm tailored for young people with psychosis, and the work has several strengths. First, a significant strength relates to the robust external validation analysis, where I showed that PsyMetRiC performed well in a

geographically distinct UK EIS sample. External validation is a crucial step in demonstrating that a risk prediction algorithm is likely to be generalizable to the intended population (Riley, 2019), and is essential for demonstrating clinical usefulness. However, as I showed in my systematic review (Chapter 6), most existing cardiometabolic risk prediction algorithms have not been externally validated, and this problem extends to psychosis research. A recent systematic review of algorithms predicting risk of transition to psychosis found an alarming lack of studies that included an external validation step (Montemagni et al., 2020).

Second, I leveraged recent advances in prognosis research by formally conducting a sample size analysis before developing PsyMetRiC. In doing so, I was able to reduce the risk of overfit, which might lead to biased predictive performance estimates. I further reduced the risk of bias from overfit by shrinking PsyMetRiC regression coefficients for optimism. I believe that these steps are likely to have been fundamental to the favourable external validation performance of PsyMetRiC.

Third, I included a detailed set of predictive performance analyses, including measures of discrimination, calibration and a decision curve analysis, in line with recommendations (Steyerberg and Vergouwe, 2014, Collins et al., 2015). As I showed in my systematic review (Chapter 6), most existing cardiometabolic risk prediction algorithms have not reported measures of algorithm calibration. Poor reporting of algorithm calibration in published research is a problem that unfortunately applies to the entire field of prognosis research (Van Calster et al., 2019). Without an assessment of algorithm calibration, it cannot be concluded that risk estimates are reliable. Therefore, such studies may be misleading and could lead to potentially incorrect and even harmful clinical decisions (Van Calster et al., 2019).

Fourth, in the development of PsyMetRiC I considered two important barriers to potential future clinical use: patient acceptability and clinical practicality. I engaged actively with the McPin Foundation YPAG to help ensure that PsyMetRiC, and the information requested by it, was likely to be acceptable for young people. I also developed two versions of PsyMetRiC, one with and one without biochemical information, so that PsyMetRiC can still be useful in instances where blood tests have been refused or are not available.

Finally, with PsyMetRiC I aimed to develop an algorithm that balanced usefulness, acceptability, and generalizability with statistical methods carefully selected to suit the available data. Given that I had access to a relatively limited sample, I did not consider more complex modelling strategies such as interactions and non-linear terms and did not proceed with a variable selection method. Variable selection may have included traditional methods such as backward selection, or more complex automated machine-learning approaches. Whilst this meant PsyMetRiC was relatively 'simple'

compared with some risk prediction algorithms, I believe these considerations are likely to have been fundamental to the favourable external validation performance. For example, a substantial body of work, including meta-analyses, has shown no performance benefit of complex machine learning approaches over simple logistic regression for clinical prediction models (Christodoulou et al., 2019, van der Ploeg et al., 2016, Takada et al., 2021). Indeed, leading experts in prognosis research have recently called for more attention to be paid to sound methodology rather than an over-reliance on machine learning algorithmic complexity, arguing that the latter contributes to nothing more than "extensive research waste" (Wilkinson et al., 2020). Nevertheless, in future, when larger samples might be available, more complex modelling approaches could be carefully considered.

Despite the strengths of PsyMetRiC, there are some weaknesses and limitations which must be taken into consideration. First and foremost, the field of prognosis research must be contextualized alongside ideas first mooted by the British epidemiologist Geoffrey Rose. Rose considered two distinct strategies for disease prevention: the high-risk vs the population approach. Rose surmised that risk factors follow a normal distribution at a population level, and therefore proposed the "prevention paradox" (Khaw, 2008). He theorised, using cholesterol and CVD as an example, that a high-risk prevention strategy would target individuals at the extreme upper end of population cholesterol distribution to prevent cases of CVD. Rose posited that this would be less effective than targeting the whole population to shift the population distribution of cholesterol to the left. He argued that by focusing on only the small number of cases at the upper extreme of cholesterol distribution, a large number of CVD cases would be missed because the majority of cases arise from closer to the centre of the normal distribution curve, as a function of the sample size distribution (Khaw, 2008).

However, most risk prediction algorithms developed for health outcomes, including PsyMetRiC, are multivariable and aim to capture as much outcome variance as possible. Therefore, while a single risk factor may be a poor predictor for a health outcome because alone it may not capture sufficient outcome variance, a multivariable approach may be a preferable method with which to consider a high-risk strategy of disease prevention.

Nevertheless, Rose's principle is relevant, particularly given the work I have presented in Chapters 2-5, which show evidence for a potentially inherent cardiometabolic risk in schizophrenia. One could therefore posit that a population prevention approach would be preferable in this population. While I agree that there is more to be done to promote healthy lifestyle behaviours for all young patients with psychosis, I do not agree that certain PsyMetRiC-related interventions are likely to be suitable for a population prevention strategy. For example, such a strategy may conclude that metabolically-active antipsychotics should never be prescribed, yet metabolically-active antipsychotics can greatly improve the lives of people with psychosis, and may reduce the risk of other disabling side-effects

such as movement disorders (Leucht et al., 2013). Therefore, I believe that in future, the real art of cardiometabolic risk reduction in young people with psychosis will lie somewhere between the extremes of Rose's population and high-risk prevention strategies. Healthy lifestyle promotion is likely to be suitable for all young people with psychosis, but for certain interventions like antipsychotic selection, tools such as PsyMetRiC can be a helpful aid toward informed, personalized treatment decisions, balancing clinical knowledge with the preferences and beliefs of the young person.

Second, I could not include some predictors in PsyMetRiC that may be biologically relevant. Given the results of Chapter 2, fasting insulin is likely to be a relevant predictor of adverse cardiometabolic outcomes in young people with psychosis. I could not include the marker because it is not yet routinely measured in clinical practice. I addressed this by including triglycerides and HDL, whose ratio may be a suitable surrogate marker for insulin resistance (Murguia-Romero et al., 2013, McLaughlin et al., 2005). Also, I could not include an inflammatory marker in PsyMetRiC, for example, CRP. While CRP is frequently measured in clinical practice, predominantly it is measured when there is suspicion of infection. Therefore, the distribution of CRP in the available sample is likely to be skewed. In addition, there are significant discrepancies between laboratories in the reporting of CRP, with some reporting the exact result and others the exact result only after an arbitrary cut-off. This heterogeneity also prevented the inclusion of CRP as a predictor.

Third, the risk estimates generated from a prediction algorithm such as PsyMetRiC are never in reality static. PsyMetRiC was developed using retrospective data, as is common in modern risk prediction algorithms developed using electronic health records. Yet, the performance of algorithms when assessed prospectively may vary, precisely because they have been used and risk estimates observed. For example, either a very-high or very-low PsyMetRiC score may alter the behaviour of either the clinician or patient, which may affect the risk estimate in either direction over time in a manner that cannot be captured in a retrospective analysis. Therefore, a prospective assessment of PsyMetRiC in a sufficient sample is required. This is particularly pertinent given that the predicted outcome in PsyMetRiC, metabolic syndrome, is a cardiometabolic 'intermediate' with few immediate consequences. While a clinician should be sufficiently motivated to act in response to their patient returning a high PsyMetRiC score, it may be more challenging to persuade young people, who may be inherently more risk-tolerant than older adults (Albert and Duffy, 2012), to change their present behaviour to prevent outcomes which are more insidious and long-term.

Fourth, another limitation common to prognosis research is that algorithms can only be confirmed to be suited for the population they were validated in when the data was collected. PsyMetRiC was developed and validated in the UK, yet different global populations are likely to vary in population

health, social norms, culture, and legislation. These factors may impact the baseline risk, or the amount of risk apportioned by any individual risk factor. PsyMetRiC will therefore require international validation to assess transportability and may require recalibration to apply to global populations. In addition, period and cohort effects may impact the baseline risk or the amount of risk apportioned by any individual risk factor (Holford, 1991). This means even well-designed externally validated risk prediction algorithms require periodic updating over time. For example, the QRISK cardiovascular risk prediction algorithm is now on its third iteration (Hippisley-Cox et al., 2017), and the recent QCOVID risk prediction algorithm (Clift et al., 2020), which aims to predict risk of hospitalization and mortality from coronavirus disease, states that it will require updating over time to reflect changes in baseline risk due to the fast-moving global pandemic of 2020/2021. PsyMetRiC in future will require periodic recalibration of either the intercept, predictor weights, or both, to remain accurate and generalizable.

Fifth, most healthcare risk prediction algorithms are developed to predict a binary health outcome. This is likely to aid the interpretability of risk estimates in a clinical setting. However, the underlying biology of such outcomes often does not represent a binary distinction between 'health' and 'disease'. Rather, diagnostic criteria aim to capture individuals at the more extreme end of a health continuum. The binarization of health presents a challenge for prognosis research, and categorization of continuous data is generally discouraged in statistical modelling (Altman and Royston, 2006). For example, the absence of a metabolic syndrome diagnosis does not equate to the absence of cardiometabolic risk. In reality, there could be very little of substance to distinguish a case of metabolic syndrome from a non-case, even as little as 1mmHg of systolic blood pressure. A future iteration of PsyMetRiC may instead consider a continuous cardiometabolic risk score as an outcome since this may more accurately align with the underlying biology.

Sixth, missing data may also have affected PsyMetRiC. All the samples used for either development or validation featured varying amounts of missing data both for the predictors and outcome. I have discussed this limitation in more detail above in Section 8.3.1 and Section 8.3.2. I also found differences in the missing compared with the included samples, which might have affected PsyMetRiC coefficients. While I used multiple imputation to reduce the impact of bias from missing data, PsyMetRiC will require validation in larger samples and prospective assessment before it can be considered suitable for clinical use.

Finally, the clinical translation of risk prediction algorithms in healthcare presents a substantial stumbling block almost universally. For example, in my systematic review, less than 1% of the included algorithms are used regularly in the clinic in the general population. A substantial body of work stands in the way of regular clinical use of an externally validated risk prediction algorithm.

This involves health technology assessment, prospective validation, stakeholder involvement, regulatory approval, and even convincing clinicians to adopt the algorithm into their clinical practice. These are all barriers that must be overcome in future to prevent PsyMetRiC from contributing to little more than research waste.

8.4 Future Directions and Implications

In Section B, I found that disruption to glucose-insulin homeostasis is likely to predate the onset of psychosis; may not be fully explained by sociodemographic, lifestyle and iatrogenic effects; and may be psychosis specific. Studies examining longitudinal associations of trajectories of childhood cardiometabolic markers and psychiatric outcomes are scarce, and so this finding ideally requires replication in an external prospective sample. A replication analysis may also help to address the limitations listed in Section 8.3.1.

Nevertheless, if replicated, this finding could have significant implications for our understanding of the cardiometabolic comorbidity of schizophrenia. Rather than solely caused by the traditional attributions of sociodemographic, lifestyle and iatrogenic factors, the cardiometabolic comorbidity of schizophrenia may have early-life beginnings and may be inherent to it. Therefore, sociodemographic, lifestyle and iatrogenic factors may be *exacerbating* rather than *causal* features of the comorbidity (see Figure 29).

This finding renews and reinforces the critical impetus that all young people presenting with psychosis must receive a comprehensive physical health assessment at the soonest available opportunity. Subtle disruption to glucose-insulin homeostasis may not in its early forms present with abnormalities to FPG or HbA1C, so these broader and less-sensitive measures must not be relied upon to confirm normal glucose-insulin homeostasis. Since detailed measurement of glucose-insulin homeostasis is often not yet possible in most current psychiatric services in the UK (for example, use of the hyperinsulinaemic-euglycaemic clamp to assess insulin sensitivity, or even a blood measurement of fasting insulin which one could combine with FPG to calculate HOMA-IR), a suitable surrogate may be the triglyceride:HDL ratio (Murguia-Romero et al., 2013, McLaughlin et al., 2005). Improved education of healthcare professionals working in psychiatry to use this marker and recognise its implications for young people with psychosis is vital. In addition, improved funding for EIS may in future permit the introduction of more sensitive tests for disrupted glucose-insulin homeostasis for all young patients presenting with psychosis.

In Section C, I found that shared genetic liability and inflammation could be potential common mechanisms underlying the associations of disrupted glucose-insulin homeostasis and schizophrenia. My findings imply that genes, the environment, or likely both, could play a role in increasing systemic inflammation, which may in turn increase the risk of both disruptions to glucose-insulin homeostasis and schizophrenia simultaneously.

Findings from the genetic correlation and colocalization study in Chapter 4 highlight a number of potential biological pathways that could simultaneously increase systemic inflammation, the risk of

cardiometabolic disorders, and schizophrenia. These findings, and the pathways implicated, now require analysis using complementary research methods, which may first involve genomics and observational research, and could then extend to animal model and experimental research.

Regarding methods related to genomics, the implicated genes and pathways can be further interrogated using methods that harness expression quantitative trait loci (eQTL) data. eQTL analysis aims to identify allelic variants associated with gene expression on the basis that a proportion of transcripts are under genetic control. A transcript that is correlated with a risk variant in a relevant tissue or cell type represents a strong candidate susceptibility gene (Lawrenson et al., 2015). Therefore, eQTL analysis would be an important confirmatory step for my findings because several of the colocalized loci are located within non-coding regions of the genome, and so they may play a more indirect regulatory role in the expression of gene products. While there are a growing number of publicly available eQTL datasets which can be used for analysis, in the past, such analysis has been hampered by prohibitive heterogeneity between datasets (Kerimov et al., 2020). However, efforts are underway to harmonize these datasets, so in future, eQTL analysis with sufficiently powered sample sizes is likely to be possible (Kerimov et al., 2020).

Cohort studies that seek to examine the antecedents of schizophrenia should in future more frequently measure cardiometabolic and inflammatory indices, and in much greater granularity. Such studies may also seek to include biochemical measurements of BDNF and related pathways. Then, these studies could help to confirm the longitudinal findings from Chapter 2 and can triangulate evidence that BDNF-related pathways may be implicated as a common mechanism for schizophrenia, cardiometabolic and inflammatory traits.

Next, should observational and eQTL analysis provide further weight to the colocalization findings, knockouts for the genes implicated in the colocalization analysis could be studied in animal models. In such studies, gene knockout animals could be tested for levels of inflammation, cardiometabolic indices and behavioural outcomes simultaneously. Such findings could provide a richer pathophysiological understanding of both schizophrenia and its associated cardiometabolic comorbidity. A convergence of results may also provide compelling evidence of potential novel therapeutic or preventative targets for schizophrenia and its associated cardiometabolic comorbidity, which could be leveraged in experimental clinical trials.

Findings from my MR study in Chapter 5 suggest that targeting inflammation may be a putative therapeutic or preventative target for schizophrenia and its associated cardiometabolic comorbidity. Yet, the most recent meta-analyses of RCTs of anti-inflammatory agents for schizophrenia have shown relatively heterogenous evidence for their efficacy in schizophrenia (Jeppesen et al., 2020,

Cakici et al., 2019) and trials of anti-inflammatory agents for their efficacy on cardiometabolic indices of people with schizophrenia are scarce. There could be two explanations for the heterogeneous efficacy in schizophrenia. First, the heterogeneity may be due to trial inclusion criteria since participants with baseline evidence of inflammation may be better candidates for immunotherapy (Raison et al., 2013). Second, there may be a difference between a therapeutic and a preventative target; anti-inflammatory medications might be more effective when trialled before the onset of psychosis, or at FEP. Indeed, a recent meta-analysis found stronger evidence for efficacy in trials conducted on younger individuals with FEP (Cakici et al., 2019). Future RCTs conducted on young individuals at the onset of psychosis should consider examining cardiometabolic markers at baseline and follow-up in addition to psychiatric indices. For example, cardiometabolic markers might be used to help select participants for trial inclusion but may also be considered as outcome measures.

In Section D, in lieu of a suitable cardiometabolic risk prediction algorithm for young people with psychosis, I developed PsyMetRiC. I do not see PsyMetRiC, as it currently exists, as the final algorithm that should be used in clinical practice. Rather, it is a useful starting point and shows the potential that such a tool could have for young people with psychosis. In future, recalibration and updating of PsyMetRiC in larger retrospective samples will allow for the refinement of the algorithm, which might further improve the accuracy of risk estimates. Next, prospective validation of PsyMetRiC will be necessary to test the 'real-world' performance of PsyMetRiC and the clinical usefulness and acceptability of the algorithm. Concomitantly, international validations could assess the transportability of PsyMetRiC to different global populations, with local recalibrations conducted such that PsyMetRiC could be used across the world. Subsequently, a body of multidisciplinary work could be conducted, featuring clinicians, allied health professionals and young people with experience of psychosis, to determine the most appropriate PsyMetRiC score cut-offs and associated therapeutic options. Finally, regulatory approval must be sought. After meeting these considerable but necessary hurdles, PsyMetRiC might then be considered ready for implementation in clinical practice and could be included in guidelines for the management of FEP, both in the UK and internationally.

8.5 Concluding Remarks

In this thesis, I present evidence that furthers our understanding of the nature and mechanisms of association of cardiometabolic disorders and schizophrenia, and makes the first steps toward improving the prediction of cardiometabolic risk in young people with psychosis. On the nature of association between cardiometabolic disorders and schizophrenia, I found that disruption to glucose-insulin homeostasis may be inherent to schizophrenia and may be detectable from early life, long before the onset of psychosis. This finding is important, and some might argue that it relieves some of the blame frequently placed on the shoulders of people who have schizophrenia for their cardiometabolic comorbidity, since my findings suggest that factors such as an unhealthy diet and physical inactivity are likely to exacerbate rather than cause the comorbidity. On the mechanisms of association between cardiometabolic disorders and schizophrenia, I found a thread of consistent evidence across independent genetic and prospective studies suggesting that shared genetic liability and inflammation could be common biological mechanisms for schizophrenia and its cardiometabolic comorbidity. In addition, results from these studies implicate biologically plausible targets that could be further investigated for their therapeutic or preventative potential for schizophrenia and its associated cardiometabolic comorbidity, and for their potential insights into the pathophysiology of schizophrenia itself. On the prediction of cardiometabolic risk in young people with psychosis, PsyMetRiC is an encouraging first step on the journey toward a valuable future tool in the arsenal of EIS healthcare professionals, to factor physical health more appropriately into treatment decisions in a personalized and informed manner. The general population has benefited from clinic-ready cardiometabolic risk prediction algorithms for decades. It is surely time that such benefits can be extended to young people with psychosis, who are in crucial need of strategies to help close the mortality gap they may sadly be faced with.

SECTION E

BIBLIOGRAPHY

ABD EL-WAHAB, E. W., SHATAT, H. Z. & CHARL, F. 2019. Adapting a Prediction Rule for Metabolic Syndrome Risk Assessment Suitable for Developing Countries. *J Prim Care Community Health,* 10**,** 2150132719882760.

ABEL, K. M., WICKS, S., SUSSER, E. S., DALMAN, C., PEDERSEN, M. G., MORTENSEN, P. B. & WEBB, R. T. 2010. Birth weight, schizophrenia, and adult mental disorder: is risk confined to the smallest babies? *Arch Gen Psychiatry,* 67**,** 923-30.

AFSHIN, A., FOROUZANFAR, M. H., REITSMA, M. B., SUR, P., ESTEP, K., LEE, A., MARCZAK, L., MOKDAD, A. H., MORADI-LAKEH, M., NAGHAVI, M., SALAMA, J. S., VOS, T., ABATE, K. H., ABBAFATI, C., AHMED, M. B., AL-ALY, Z., ALKERWI, A., AL-RADDADI, R., AMARE, A. T., AMBERBIR, A., AMEGAH, A. K., AMINI, E., AMROCK, S. M., ANJANA, R. M., ARNLOV, J., ASAYESH, H., BANERJEE, A., BARAC, A., BAYE, E., BENNETT, D. A., BEYENE, A. S., BIADGILIGN, S., BIRYUKOV, S., BJERTNESS, E., BONEYA, D. J., CAMPOS-NONATO, I., CARRERO, J. J., CECILIO, P., CERCY, K., CIOBANU, L. G., CORNABY, L., DAMTEW, S. A., DANDONA, L., DANDONA, R., DHARMARATNE, S. D., DUNCAN, B. B., ESHRATI, B., ESTEGHAMATI, A., FEIGIN, V. L., FERNANDES, J. C., FURST, T., GEBREHIWOT, T. T., GOLD, A., GONA, P. N., GOTO, A., HABTEWOLD, T. D., HADUSH, K. T., HAFEZI-NEJAD, N., HAY, S. I., HORINO, M., ISLAMI, F., KAMAL, R., KASAEIAN, A., KATIKIREDDI, S. V., KENGNE, A. P., KESAVACHANDRAN, C. N., KHADER, Y. S., KHANG, Y. H., KHUBCHANDANI, J., KIM, D., KIM, Y. J., KINFU, Y., KOSEN, S., KU, T., DEFO, B. K., KUMAR, G. A., LARSON, H. J., LEINSALU, M., LIANG, X., LIM, S. S., LIU, P., LOPEZ, A. D., LOZANO, R., MAJEED, A., MALEKZADEH, R., MALTA, D. C., MAZIDI, M., MCALINDEN, C., MCGARVEY, S. T., MENGISTU, D. T., MENSAH, G. A., MENSINK, G. B. M., MEZGEBE, H. B., MIRRAKHIMOV, E. M., MUELLER, U. O., NOUBIAP, J. J., OBERMEYER, C. M., OGBO, F. A., OWOLABI, M. O., PATTON, G. C., et al. 2017. Health Effects of Overweight and Obesity in 195 Countries over 25 Years. *N Engl J Med,* 377**,** 13-27.

AGUIRRE, V., UCHIDA, T., YENUSH, L., DAVIS, R. & WHITE, M. F. 2000. The c-Jun NH(2)-terminal kinase promotes insulin resistance during association with insulin receptor substrate-1 and phosphorylation of Ser(307). *J Biol Chem,* 275**,** 9047-54.

AGUIRRE, V., WERNER, E. D., GIRAUD, J., LEE, Y. H., SHOELSON, S. E. & WHITE, M. F. 2002. Phosphorylation of Ser307 in insulin receptor substrate-1 blocks interactions with the insulin receptor and inhibits insulin action. *J Biol Chem,* 277**,** 1531-7.

ALAA, A. M., BOLTON, T., DI ANGELANTONIO, E., RUDD, J. H. F. & VAN DER SCHAAR, M. 2019. Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. *PLoS One,* 14**,** e0213653.

ALBA, A. C., AGORITSAS, T., WALSH, M., HANNA, S., IORIO, A., DEVEREAUX, P. J., MCGINN, T. & GUYATT, G. 2017. Discrimination and Calibration of Clinical Prediction Models: Users' Guides to the Medical Literature. *JAMA,* 318**,** 1377-1384.

ALBERT, S. M. & DUFFY, J. 2012. Differences in Risk Aversion between Young and Older Adults. *Neurosci Neuroecon,* 2012.

ALBERTI, K. G., ECKEL, R. H., GRUNDY, S. M., ZIMMET, P. Z., CLEEMAN, J. I., DONATO, K. A., FRUCHART, J. C., JAMES, W. P., LORIA, C. M., SMITH, S. C., JR., INTERNATIONAL DIABETES FEDERATION TASK FORCE ON, E., PREVENTION, HATIONAL HEART, L., BLOOD, I., AMERICAN HEART, A., WORLD HEART, F., INTERNATIONAL ATHEROSCLEROSIS, S. & INTERNATIONAL ASSOCIATION FOR THE STUDY OF, O. 2009. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation,* 120**,** 1640-5.

ALBERTI, K. G. & ZIMMET, P. Z. 1998. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet Med,* 15**,** 539-53.

ALEXANDER, B. T., DASINGER, J. H. & INTAPAD, S. 2015. Fetal programming and cardiovascular pathology. *Compr Physiol,* 5**,** 997-1025.

ALLENDER, S., FOSTER, C., SCARBOROUGH, P. & RAYNER, M. 2007. The burden of physical activity-related ill health in the UK. *J Epidemiol Community Health,* 61**,** 344-8.

ALONSO-PEDRERO, L., BES-RASTROLLO, M. & MARTI, A. 2019. Effects of antidepressant and antipsychotic use on weight gain: A systematic review. *Obes Rev,* 20**,** 1680-1690.

ALTMAN, D. G. & ROYSTON, P. 2006. The cost of dichotomising continuous variables. *BMJ,* 332**,** 1080.

ALTMAN, D. G., VERGOUWE, Y., ROYSTON, P. & MOONS, K. G. 2009. Prognosis and prognostic research: validating a prognostic model. *BMJ,* 338**,** b605.

ANAND, S. S., HAWKES, C., DE SOUZA, R. J., MENTE, A., DEHGHAN, M., NUGENT, R., ZULYNIAK, M. A., WEIS, T., BERNSTEIN, A. M., KRAUSS, R. M., KROMHOUT, D., JENKINS, D. J. A., MALIK, V., MARTINEZ-GONZALEZ, M. A., MOZAFFARIAN, D., YUSUF, S., WILLETT, W. C. & POPKIN, B. M. 2015. Food Consumption and its Impact

on Cardiovascular Disease: Importance of Solutions Focused on the Globalized Food System: A Report From the Workshop Convened by the World Heart Federation. *J Am Coll Cardiol,* 66**,** 1590-1614.

ANDERSON, J. P., PARIKH, J. R., SHENFELD, D. K., IVANOV, V., MARKS, C., CHURCH, B. W., LARAMIE, J. M., MARDEKIAN, J., PIPER, B. A., WILLKE, R. J. & RUBLEE, D. A. 2015. Reverse Engineering and Evaluation of Prediction Models for Progression to Type 2 Diabetes: An Application of Machine Learning Using Electronic Health Records. *J Diabetes Sci Technol,* 10**,** 6-18.

ANDREASSEN, O. A., DJUROVIC, S., THOMPSON, W. K., SCHORK, A. J., KENDLER, K. S., O'DONOVAN, M. C., RUJESCU, D., WERGE, T., VAN DE BUNT, M., MORRIS, A. P., MCCARTHY, M. I., INTERNATIONAL CONSORTIUM FOR BLOOD PRESSURE, G., DIABETES GENETICS, R., META-ANALYSIS, C., PSYCHIATRIC GENOMICS CONSORTIUM SCHIZOPHRENIA WORKING, G., RODDEY, J. C., MCEVOY, L. K., DESIKAN, R. S. & DALE, A. M. 2013. Improved detection of common variants associated with schizophrenia by leveraging pleiotropy with cardiovascular-disease risk factors. *Am J Hum Genet,* 92**,** 197-209.

ANTTILA, V., BULIK-SULLIVAN, B., FINUCANE, H. K., WALTERS, R. K., BRAS, J., DUNCAN, L., ESCOTT-PRICE, V., FALCONE, G. J., GORMLEY, P., MALIK, R., PATSOPOULOS, N. A., RIPKE, S., WEI, Z., YU, D., LEE, P. H., TURLEY, P., GRENIER-BOLEY, B., CHOURAKI, V., KAMATANI, Y., BERR, C., LETENNEUR, L., HANNEQUIN, D., AMOUYEL, P., BOLAND, A., DELEUZE, J. F., DURON, E., VARDARAJAN, B. N., REITZ, C., GOATE, A. M., HUENTELMAN, M. J., KAMBOH, M. I., LARSON, E. B., ROGAEVA, E., ST GEORGE-HYSLOP, P., HAKONARSON, H., KUKULL, W. A., FARRER, L. A., BARNES, L. L., BEACH, T. G., DEMIRCI, F. Y., HEAD, E., HULETTE, C. M., JICHA, G. A., KAUWE, J. S. K., KAYE, J. A., LEVERENZ, J. B., LEVEY, A. I., LIEBERMAN, A. P., PANKRATZ, V. S., POON, W. W., QUINN, J. F., SAYKIN, A. J., SCHNEIDER, L. S., SMITH, A. G., SONNEN, J. A., STERN, R. A., VAN DEERLIN, V. M., VAN ELDIK, L. J., HAROLD, D., RUSSO, G., RUBINSZTEIN, D. C., BAYER, A., TSOLAKI, M., PROITSI, P., FOX, N. C., HAMPEL, H., OWEN, M. J., MEAD, S., PASSMORE, P., MORGAN, K., NÖTHEN, M. M., ROSSOR, M., LUPTON, M. K., HOFFMANN, P., KORNHUBER, J., LAWLOR, B., MCQUILLIN, A., AL-CHALABI, A., BIS, J. C., RUIZ, A., BOADA, M., SESHADRI, S., BEISER, A., RICE, K., VAN DER LEE, S. J., DE JAGER, P. L., GESCHWIND, D. H., RIEMENSCHNEIDER, M., RIEDEL-HELLER, S., ROTTER, J. I., RANSMAYR, G., HYMAN, B. T., CRUCHAGA, C., ALEGRET, M., WINSVOLD, B., PALTA, P., FARH,

K. H., CUENCA-LEON, E., FURLOTTE, N., KURTH, T., et al. 2018. Analysis of shared heritability in common disorders of the brain. *Science,* 360.

ARCHIBALD, L., BRUNETTE, M. F., WALLIN, D. J. & GREEN, A. I. 2019. Alcohol Use Disorder and Schizophrenia or Schizoaffective Disorder. *Alcohol Res,* 40.

ARONSON, D. & EDELMAN, E. R. 2014. Coronary artery disease and diabetes mellitus. *Cardiol Clin,* 32**,** 439-55.

ARRUDA, A. P., MILANSKI, M., COOPE, A., TORSONI, A. S., ROPELLE, E., CARVALHO, D. P., CARVALHEIRA, J. B. & VELLOSO, L. A. 2011. Low-grade hypothalamic inflammation leads to defective thermogenesis, insulin resistance, and impaired insulin secretion. *Endocrinology,* 152**,** 1314-26.

ASKARI, M., HESHMATI, J., SHAHINFAR, H., TRIPATHI, N. & DANESHZAD, E. 2020. Ultra-processed food and the risk of overweight and obesity: a systematic review and meta-analysis of observational studies. *Int J Obes (Lond),* 44**,** 2080-2091.

ASPAROUHOV, T., MUTHEN, B. 2014. Auxiliary Variables in Mixture Modeling: Three-Step Approaches Using Mplus. *Structural Equation Modeling: A Multidisciplinary Journal,* 21**,** 329-341.

ASSMANN, G., CULLEN, P. & SCHULTE, H. 2002. Simple scoring scheme for calculating the risk of acute coronary events based on the 10-year follow-up of the prospective cardiovascular Munster (PROCAM) study. *Circulation,* 105**,** 310-5.

AUCOIN, M., LACHANCE, L., COOLEY, K. & KIDD, S. 2020. Diet and Psychosis: A Scoping Review. *Neuropsychobiology,* 79**,** 20-42.

AUSTIN, S. F., MORS, O., SECHER, R. G., HJORTHOJ, C. R., ALBERT, N., BERTELSEN, M., JENSEN, H., JEPPESEN, P., PETERSEN, L., RANDERS, L., THORUP, A. & NORDENTOFT, M. 2013. Predictors of recovery in first episode psychosis: the OPUS cohort at 10 year follow-up. *Schizophr Res,* 150**,** 163-8.

AYALA SOLARES, J. R., CANOY, D., RAIMONDI, F. E. D., ZHU, Y., HASSAINE, A., SALIMI-KHORSHIDI, G., TRAN, J., COPLAND, E., ZOTTOLI, M., PINHO-GOMES, A. C., NAZARZADEH, M. & RAHIMI, K. 2019. Long-Term Exposure to Elevated Systolic Blood Pressure in Predicting Incident Cardiovascular Disease: Evidence From Large-Scale Routine Electronic Health Records. *J Am Heart Assoc,* 8**,** e012129.

BAHRAMI, S., STEEN, N. E., SHADRIN, A., O'CONNELL, K., FREI, O., BETTELLA, F., WIRGENES, K. V., KRULL, F., FAN, C. C., DALE, A. M., SMELAND, O. B., DJUROVIC, S. & ANDREASSEN, O. A. 2020. Shared Genetic Loci Between Body Mass Index and Major Psychiatric Disorders: A Genome-wide Association Study. *JAMA Psychiatry*.

BAK, M., FRANSEN, A., JANSSEN, J., VAN OS, J. & DRUKKER, M. 2014. Almost all antipsychotics result in weight gain: a meta-analysis. *PLoS One,* 9**,** e94112.

BALKAU, B. & CHARLES, M. A. 1999. Comment on the provisional report from the WHO consultation. European Group for the Study of Insulin Resistance (EGIR). *Diabet Med,* 16**,** 442-3.

BAN, T. A. 2007. Fifty years chlorpromazine: a historical perspective. *Neuropsychiatr Dis Treat,* 3**,** 495-500.

BANKS, E., JOSHY, G., KORDA, R. J., STAVRESKI, B., SOGA, K., EGGER, S., DAY, C., CLARKE, N. E., LEWINGTON, S. & LOPEZ, A. D. 2019. Tobacco smoking and risk of 36 cardiovascular disease subtypes: fatal and non-fatal outcomes in a large prospective Australian study. *BMC Med,* 17**,** 128.

BARKER, D. J., GLUCKMAN, P. D., GODFREY, K. M., HARDING, J. E., OWENS, J. A. & ROBINSON, J. S. 1993. Fetal nutrition and cardiovascular disease in adult life. *Lancet,* 341**,** 938-41.

BARKER, D. J. & OSMOND, C. 1987. Death rates from stroke in England and Wales predicted from past maternal mortality. *Br Med J (Clin Res Ed),* 295**,** 83-6.

BARKER, D. J., OSMOND, C., FORSEN, T. J., KAJANTIE, E. & ERIKSSON, J. G. 2005. Trajectories of growth among children who have coronary events as adults. *N Engl J Med,* 353**,** 1802-9.

BARKER, D. J., OSMOND, C. & LAW, C. M. 1989. The intrauterine and early postnatal origins of cardiovascular disease and chronic bronchitis. *J Epidemiol Community Health,* 43**,** 237-40.

BARNES, T. R., DRAKE, R., PATON, C., COOPER, S. J., DEAKIN, B., FERRIER, I. N., GREGORY, C. J., HADDAD, P. M., HOWES, O. D., JONES, I., JOYCE, E. M., LEWIS, S., LINGFORD-HUGHES, A., MACCABE, J. H., OWENS, D. C., PATEL, M. X., SINCLAIR, J. M., STONE, J. M., TALBOT, P. S., UPTHEGROVE, R., WIECK, A. & YUNG, A. R. 2020. Evidence-based guidelines for the pharmacological treatment of schizophrenia: Updated recommendations from the British Association for Psychopharmacology. *J Psychopharmacol,* 34**,** 3-78.

BARREIRA, T. V., HARRINGTON, D. M., STAIANO, A. E., HEYMSFIELD, S. B. & KATZMARZYK, P. T. 2011. Body adiposity index, body mass index, and body fat in white and black adults. *JAMA,* 306**,** 828-30.

BATHINA, S. & DAS, U. N. 2019. PUFAs, BDNF and lipoxin A4 inhibit chemical-induced cytotoxicity of RIN5F cells in vitro and streptozotocin-induced type 2 diabetes mellitus in vivo. *Lipids Health Dis,* 18**,** 214.

BEALE, A. L., MEYER, P., MARWICK, T. H., LAM, C. S. P. & KAYE, D. M. 2018. Sex Differences in Cardiovascular Pathophysiology: Why Women Are Overrepresented in Heart Failure With Preserved Ejection Fraction. *Circulation,* 138**,** 198-205.

BELL, J. A., CARSLAKE, D., WADE, K. H., RICHMOND, R. C., LANGDON, R. J., VINCENT, E. E., HOLMES, M. V., TIMPSON, N. J. & DAVEY SMITH, G. 2018. Influence of puberty timing on adiposity and cardiometabolic traits: A Mendelian randomisation study. *PLoS Med,* 15**,** e1002641.

BERGMAN, R. N. & ADER, M. 2005. Atypical antipsychotics and glucose homeostasis. *J Clin Psychiatry,* 66**,** 504-14.

BERISA, T. & PICKRELL, J. K. 2016. Approximately independent linkage disequilibrium blocks in human populations. *Bioinformatics,* 32**,** 283-5.

BERMUDEZ, V., MARTINEZ, M. S., CHAVEZ-CASTILLO, M., OLIVAR, L. C., MORILLO, J., MEJIAS, J. C., ROJAS, M., SALAZAR, J., ROJAS, J., ANEZ, R. & CABRERA, M. 2015. Relationship between Alcohol Consumption and Components of the Metabolic Syndrome in Adult Population from Maracaibo City, Venezuela. *Adv Prev Med,* 2015**,** 352547.

BERNDT, S. I., GUSTAFSSON, S., MAGI, R., GANNA, A., WHEELER, E., FEITOSA, M. F., JUSTICE, A. E., MONDA, K. L., CROTEAU-CHONKA, D. C., DAY, F. R., ESKO, T., FALL, T., FERREIRA, T., GENTILINI, D., JACKSON, A. U., LUAN, J., RANDALL, J. C., VEDANTAM, S., WILLER, C. J., WINKLER, T. W., WOOD, A. R., WORKALEMAHU, T., HU, Y. J., LEE, S. H., LIANG, L., LIN, D. Y., MIN, J. L., NEALE, B. M., THORLEIFSSON, G., YANG, J., ALBRECHT, E., AMIN, N., BRAGG-GRESHAM, J. L., CADBY, G., DEN HEIJER, M., EKLUND, N., FISCHER, K., GOEL, A., HOTTENGA, J. J., HUFFMAN, J. E., JARICK, I., JOHANSSON, A., JOHNSON, T., KANONI, S., KLEBER, M. E., KONIG, I. R., KRISTIANSSON, K., KUTALIK, Z., LAMINA, C., LECOEUR, C., LI, G., MANGINO, M., MCARDLE, W. L., MEDINA-GOMEZ, C., MULLER-NURASYID, M., NGWA, J. S., NOLTE, I. M., PATERNOSTER, L., PECHLIVANIS, S., PEROLA, M., PETERS, M. J., PREUSS, M., ROSE, L. M., SHI, J., SHUNGIN, D., SMITH, A. V., STRAWBRIDGE, R. J., SURAKKA, I., TEUMER, A., TRIP, M. D., TYRER, J., VAN VLIET-OSTAPTCHOUK, J. V., VANDENPUT, L., WAITE, L. L., ZHAO, J. H., ABSHER, D., ASSELBERGS, F. W., ATALAY, M., ATTWOOD, A. P., BALMFORTH, A. J., BASART, H., BEILBY, J., BONNYCASTLE, L. L., BRAMBILLA, P., BRUINENBERG, M., CAMPBELL, H., CHASMAN, D. I., CHINES, P. S., COLLINS, F. S., CONNELL, J. M., COOKSON, W. O., DE FAIRE, U., DE VEGT, F., DEI, M., DIMITRIOU, M., EDKINS, S., ESTRADA, K., EVANS, D. M., FARRALL, M., FERRARIO, M. M., et al. 2013. Genome-wide meta-analysis identifies 11

new loci for anthropometric traits and provides insights into genetic architecture. *Nat Genet,* 45**,** 501-12.

BERRY, A., DRAKE, R. J., WEBB, R. T., ASHCROFT, D. M., CARR, M. J. & YUNG, A. R. 2018. Investigating the Agreement Between Cardiovascular Disease Risk Calculators Among People Diagnosed With Schizophrenia. *Front Psychiatry,* 9**,** 685.

BILLINGS, L. K. & FLOREZ, J. C. 2010. The genetics of type 2 diabetes: what have we learned from GWAS? *Ann N Y Acad Sci,* 1212**,** 59-77.

BLUM, B. P. & MANN, J. J. 2002. The GABAergic system in schizophrenia. *Int J Neuropsychopharmacol,* 5**,** 159-79.

BLYTHE, J. & WHITE, J. 2012. Role of the mental health nurse towards physical health care in serious mental illness: an integrative review of 10 years of UK literature. *Int J Ment Health Nurs,* 21**,** 193-201.

BODEN-ALBALA, B., SACCO, R. L., LEE, H. S., GRAHAME-CLARKE, C., RUNDEK, T., ELKIND, M. V., WRIGHT, C., GIARDINA, E. G., DITULLIO, M. R., HOMMA, S. & PAIK, M. C. 2008. Metabolic syndrome and ischemic stroke risk: Northern Manhattan Study. *Stroke,* 39**,** 30-5.

BOUCHER, H. 2019. *Exploring the Utility of MUAC in Classifying Adult Metabolic Syndrome Using NHANES 2015-16.*

BOWDEN, J., DAVEY SMITH, G. & BURGESS, S. 2015. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol,* 44**,** 512-25.

BOWDEN, J., DAVEY SMITH, G., HAYCOCK, P. C. & BURGESS, S. 2016a. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol,* 40**,** 304-14.

BOWDEN, J., DEL GRECO, M. F., MINELLI, C., DAVEY SMITH, G., SHEEHAN, N. & THOMPSON, J. 2017. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. *Stat Med,* 36**,** 1783-1802.

BOWDEN, J., DEL GRECO, M. F., MINELLI, C., DAVEY SMITH, G., SHEEHAN, N. A. & THOMPSON, J. R. 2016b. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I2 statistic. *Int J Epidemiol,* 45**,** 1961-1974.

BOWKER, N., SHAH, R. L., SHARP, S. J., LUAN, J., STEWART, I. D., WHEELER, E., FERREIRA, M. A. R., BARAS, A., WAREHAM, N. J., LANGENBERG, C. & LOTTA, L. A. 2020. Meta-analysis investigating the role of interleukin-6 mediated inflammation in type 2 diabetes. *EBioMedicine,* 61**,** 103062.

BOYD, A., GOLDING, J., MACLEOD, J., LAWLOR, D. A., FRASER, A., HENDERSON, J., MOLLOY, L., NESS, A., RING, S. & DAVEY SMITH, G. 2013a. Cohort Profile: the 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol,* 42**,** 111-27.

BOYD, A., GOLDING, J., MACLEOD, J., LAWLOR, D. A., FRASER, A., HENDERSON, J., MOLLOY, L., NESS, A., RING, S. & DAVEY SMITH, G. 2013b. Cohort Profile: The 'Children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol,* 42**,** 111-127.

BOYDELL, J., VAN OS, J., MCKENZIE, K., ALLARDYCE, J., GOEL, R., MCCREADIE, R. G. & MURRAY, R. M. 2001. Incidence of schizophrenia in ethnic minorities in London: ecological study into interactions with environment. *BMJ,* 323**,** 1336-8.

BRAND, R. J., ROSENMAN, R.H., SHOLTZ, R.I., FRIEDMAN, M. 1976. Multivariate Prediction of Coronary Heart Disease in the Western Collaborative Group Study Compared to the Findings of the Framingham Study. *Circulation,* 53**,** 348-355.

BREIER, A., SU, T. P., SAUNDERS, R., CARSON, R. E., KOLACHANA, B. S., DE BARTOLOMEIS, A., WEINBERGER, D. R., WEISENFELD, N., MALHOTRA, A. K., ECKELMAN, W. C. & PICKAR, D. 1997. Schizophrenia is associated with elevated amphetamine-induced synaptic dopamine concentrations: evidence from a novel positron emission tomography method. *Proc Natl Acad Sci U S A,* 94**,** 2569-74.

BREIMAN, L. 2001. Statistical Modeling: The Two Cultures. *Statist. Sci.,* 16**,** 199-231.

BROWN, S., INSKIP, H. & BARRACLOUGH, B. 2000. Causes of the excess mortality of schizophrenia. *Br J Psychiatry,* 177**,** 212-7.

BRUMPTON, B., SANDERSON, E., HEILBRON, K., HARTWIG, F. P., HARRISON, S., VIE, G. A., CHO, Y., HOWE, L. D., HUGHES, A., BOOMSMA, D. I., HAVDAHL, A., HOPPER, J., NEALE, M., NIVARD, M. G., PEDERSEN, N. L., REYNOLDS, C. A., TUCKER-DROB, E. M., GROTZINGER, A., HOWE, L., MORRIS, T., LI, S., WITHIN-FAMILY, C., ANDME RESEARCH, T., AUTON, A., WINDMEIJER, F., CHEN, W. M., BJORNGAARD, J. H., HVEEM, K., WILLER, C., EVANS, D. M., KAPRIO, J., DAVEY SMITH, G., ASVOLD, B. O., HEMANI, G. & DAVIES, N. M. 2020. Avoiding dynastic, assortative mating, and population stratification biases in Mendelian randomization through within-family analyses. *Nat Commun,* 11**,** 3519.

BRUNETTE, M. F., MUESER, K. T., BABBIN, S., MEYER-KALOS, P., ROSENHECK, R., CORRELL, C. U., CATHER, C., ROBINSON, D. G., SCHOOLER, N. R., PENN, D. L., ADDINGTON, J., ESTROFF, S. E., GOTTLIEB, J., GLYNN, S. M., MARCY, P.,

ROBINSON, J. & KANE, J. M. 2018. Demographic and clinical correlates of substance use disorders in first episode psychosis. *Schizophr Res,* 194**,** 4-12.

BUCKLEY, P. F., MILLER, B. J., LEHRER, D. S. & CASTLE, D. J. 2009. Psychiatric comorbidities and schizophrenia. *Schizophr Bull,* 35**,** 383-402.

BULIK-SULLIVAN, B., FINUCANE, H. K., ANTTILA, V., GUSEV, A., DAY, F. R., LOH, P. R., REPROGEN, C., PSYCHIATRIC GENOMICS, C., GENETIC CONSORTIUM FOR ANOREXIA NERVOSA OF THE WELLCOME TRUST CASE CONTROL, C., DUNCAN, L., PERRY, J. R., PATTERSON, N., ROBINSON, E. B., DALY, M. J., PRICE, A. L. & NEALE, B. M. 2015a. An atlas of genetic correlations across human diseases and traits. *Nat Genet,* 47**,** 1236-41.

BULIK-SULLIVAN, B. K., LOH, P. R., FINUCANE, H. K., RIPKE, S., YANG, J., SCHIZOPHRENIA WORKING GROUP OF THE PSYCHIATRIC GENOMICS, C., PATTERSON, N., DALY, M. J., PRICE, A. L. & NEALE, B. M. 2015b. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet,* 47**,** 291-5.

BURGESS, S., BUTTERWORTH, A. & THOMPSON, S. G. 2013. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol,* 37**,** 658-65.

BURGESS, S. & THOMPSON, S. G. 2015. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *Am J Epidemiol,* 181**,** 251-60.

BUSCOT, M. J., THOMSON, R. J., JUONALA, M., SABIN, M. A., BURGNER, D. P., LEHTIMAKI, T., HUTRI-KAHONEN, N., VIIKARI, J. S. A., RAITAKARI, O. T. & MAGNUSSEN, C. G. 2018. Distinct child-to-adult body mass index trajectories are associated with different levels of adult cardiometabolic risk. *Eur Heart J,* 39**,** 2263-2270.

BUTLER, J. S., BURKHAUSER, R.V., MITCHELL, J.M., PINCUS, T.P., 1987. Measurement Error in Self-Reported Health Variables. *The Review of Economics and Statistics,* 69**,** 644-650.

BUUREN, S. V., GROOTHUIS-OUDSHOORN, K. 2011. MICE: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software,* 45**,** 1-67.

BZDOK, D. & IOANNIDIS, J. P. A. 2019. Exploration, Inference, and Prediction in Neuroscience and Biomedicine. *Trends Neurosci,* 42**,** 251-262.

BZDOK, D., VAROQUAUX, G. & STEYERBERG, E. W. 2021. Prediction, Not Association, Paves the Road to Precision Medicine. *JAMA Psychiatry,* 78**,** 127-128.

CABO, J., ALONSO, R. & MATA, P. 2012. Omega-3 fatty acids and blood pressure. *Br J Nutr,* 107 Suppl 2**,** S195-200.

CADENHEAD, K. S., MINICHINO, A., KELSVEN, S., ADDINGTON, J., BEARDEN, C., CANNON, T. D., CORNBLATT, B. A., MATHALON, D., MCGLASHAN, T. H., PERKINS, D. O., SEIDMAN, L. J., TSUANG, M., WALKER, E. F., WOODS, S. W., YAO, J. & NORTH AMERICAN PRODROMAL LONGITUDINAL STUDIES, C. 2019. Metabolic abnormalities and low dietary Omega 3 are associated with symptom severity and worse functioning prior to the onset of psychosis: Findings from the North American Prodrome Longitudinal Studies Consortium. *Schizophr Res,* 204**,** 96-103.

CAI, N., REVEZ, J. A., ADAMS, M. J., ANDLAUER, T. F. M., BREEN, G., BYRNE, E. M., CLARKE, T. K., FORSTNER, A. J., GRABE, H. J., HAMILTON, S. P., LEVINSON, D. F., LEWIS, C. M., LEWIS, G., MARTIN, N. G., MILANESCHI, Y., MORS, O., MULLER-MYHSOK, B., PENNINX, B., PERLIS, R. H., PISTIS, G., POTASH, J. B., PREISIG, M., SHI, J., SMOLLER, J. W., STREIT, F., TIEMEIER, H., UHER, R., VAN DER AUWERA, S., VIKTORIN, A., WEISSMAN, M. M., CONSORTIUM, M. D. D. W. G. O. T. P. G., KENDLER, K. S. & FLINT, J. 2020. Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nat Genet,* 52**,** 437-447.

CAKICI, N., VAN BEVEREN, N. J. M., JUDGE-HUNDAL, G., KOOLA, M. M. & SOMMER, I. E. C. 2019. An update on the efficacy of anti-inflammatory agents for patients with schizophrenia: a meta-analysis. *Psychol Med***,** 1-13.

CALABRO, P., WILLERSON, J. T. & YEH, E. T. 2003. Inflammatory cytokines stimulated C-reactive protein production by human coronary artery smooth muscle cells. *Circulation,* 108**,** 1930-2.

CALABRO, P. & YEH, E. T. 2008. Intra-abdominal adiposity, inflammation, and cardiovascular risk: new insight into global cardiometabolic risk. *Curr Hypertens Rep,* 10**,** 32-8.

CAMBIEN, F. & TIRET, L. 2007. Genetics of cardiovascular diseases: from single mutations to the whole genome. *Circulation,* 116**,** 1714-24.

CAMPBELL, M., MCKENZIE, J. E., SOWDEN, A., KATIKIREDDI, S. V., BRENNAN, S. E., ELLIS, S., HARTMANN-BOYCE, J., RYAN, R., SHEPPERD, S., THOMAS, J., WELCH, V. & THOMSON, H. 2020. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *BMJ,* 368**,** l6890.

CANTOR-GRAAE, E. & SELTEN, J. P. 2005. Schizophrenia and migration: a meta-analysis and review. *Am J Psychiatry,* 162**,** 12-24.

CAPPUCCIO, F. P., COOPER, D., D'ELIA, L., STRAZZULLO, P. & MILLER, M. A. 2011. Sleep duration predicts cardiovascular outcomes: a systematic review and meta-analysis of prospective studies. *Eur Heart J,* 32**,** 1484-92.

CAPPUCCIO, F. P., D'ELIA, L., STRAZZULLO, P. & MILLER, M. A. 2010. Quantity and quality of sleep and incidence of type 2 diabetes: a systematic review and meta-analysis. *Diabetes Care,* 33**,** 414-20.

CAPPUCCIO, F. P. & MILLER, M. A. 2017. Sleep and Cardio-Metabolic Disease. *Curr Cardiol Rep,* 19**,** 110.

CARAVAGGIO, F., BORLIDO, C., HAHN, M., FENG, Z., FERVAHA, G., GERRETSEN, P., NAKAJIMA, S., PLITMAN, E., CHUNG, J. K., IWATA, Y., WILSON, A., REMINGTON, G. & GRAFF-GUERRERO, A. 2015. Reduced insulin sensitivity is related to less endogenous dopamine at D2/3 receptors in the ventral striatum of healthy nonobese humans. *Int J Neuropsychopharmacol,* 18**,** pyv014.

CARDINAL, R. N. 2017. Clinical records anonymisation and text extraction (CRATE): an open-source software system. *BMC Medical Informatics and Decision Making,* 17**,** 50.

CARLSSON, A. & LINDQVIST, M. 1963. Effect of Chlorpromazine or Haloperidol on Formation of 3methoxytyramine and Normetanephrine in Mouse Brain. *Acta Pharmacol Toxicol (Copenh),* 20**,** 140-4.

CARNEY, R., COTTER, J., BRADSHAW, T., FIRTH, J. & YUNG, A. R. 2016. Cardiometabolic risk factors in young people at ultra-high risk for psychosis: A systematic review and meta-analysis. *Schizophr Res,* 170**,** 290-300.

CARRERA, N., ARROJO, M., SANJUAN, J., RAMOS-RIOS, R., PAZ, E., SUAREZ-RAMA, J. J., PARAMO, M., AGRA, S., BRENLLA, J., MARTINEZ, S., RIVERO, O., COLLIER, D. A., PALOTIE, A., CICHON, S., NOTHEN, M. M., RIETSCHEL, M., RUJESCU, D., STEFANSSON, H., STEINBERG, S., SIGURDSSON, E., ST CLAIR, D., TOSATO, S., WERGE, T., STEFANSSON, K., GONZALEZ, J. C., VALERO, J., GUTIERREZ-ZOTES, A., LABAD, A., MARTORELL, L., VILELLA, E., CARRACEDO, A. & COSTAS, J. 2012. Association study of nonsynonymous single nucleotide polymorphisms in schizophrenia. *Biol Psychiatry,* 71**,** 169-77.

CARVALHO, E., JANSSON, P. A., AXELSEN, M., ERIKSSON, J. W., HUANG, X., GROOP, L., RONDINONE, C., SJÖSTRÖM, L. & SMITH, U. 1999. Low cellular IRS 1 gene and protein expression predict insulin resistance and NIDDM. *FASEB J,* 13**,** 2173-8.

CASELLA, G., BERGER, RL. 2002. *Statistical Inference,* California, USA, Duxbury Pacific Grove.

CASTLE, D., SHAM, P. & MURRAY, R. 1998. Differences in distribution of ages of onset in males and females with schizophrenia. *Schizophr Res,* 33**,** 179-83.

CEDIEL, G., CORVALAN, C., AGUIRRE, C., DE ROMANA, D. L. & UAUY, R. 2016. Serum 25-Hydroxyvitamin D associated with indicators of body fat and insulin resistance in prepubertal chilean children. *Int J Obes (Lond),* 40**,** 147-52.

CHAM, H., RESHETNYAK, E., ROSENFELD, B. & BREITBART, W. 2017. Full Information Maximum Likelihood Estimation for Latent Variable Interactions With Incomplete Indicators. *Multivariate Behav Res,* 52**,** 12-30.

CHANG, C. C., CHOW, C. C., TELLIER, L. C. A. M., VATTIKUTI, S., PURCELL, S. M. & LEE, J. J. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience,* 4.

CHANG, W., CHENG, W., ALLAIRE, J.J., XIE, Y., MCPHERSON, J., 2020. shiny: Web Application Framework for R.

CHANOCK, S. J., MANOLIO, T., BOEHNKE, M., BOERWINKLE, E., HUNTER, D. J., THOMAS, G., HIRSCHHORN, J. N., ABECASIS, G., ALTSHULER, D., BAILEY-WILSON, J. E., BROOKS, L. D., CARDON, L. R., DALY, M., DONNELLY, P., FRAUMENI, J. F., JR., FREIMER, N. B., GERHARD, D. S., GUNTER, C., GUTTMACHER, A. E., GUYER, M. S., HARRIS, E. L., HOH, J., HOOVER, R., KONG, C. A., MERIKANGAS, K. R., MORTON, C. C., PALMER, L. J., PHIMISTER, E. G., RICE, J. P., ROBERTS, J., ROTIMI, C., TUCKER, M. A., VOGAN, K. J., WACHOLDER, S., WIJSMAN, E. M., WINN, D. M. & COLLINS, F. S. 2007. Replicating genotype-phenotype associations. *Nature,* 447**,** 655-60.

CHEAH, S. Y., LAWFORD, B. R., YOUNG, R. M., CONNOR, J. P., PHILLIP MORRIS, C. & VOISEY, J. 2014. BDNF SNPs are implicated in comorbid alcohol dependence in schizophrenia but not in alcohol-dependent patients without schizophrenia. *Alcohol Alcohol,* 49**,** 491-7.

CHEN, L., TONKIN, A. M., MOON, L., MITCHELL, P., DOBSON, A., GILES, G., HOBBS, M., PHILLIPS, P. J., SHAW, J. E., SIMMONS, D., SIMONS, L. A., FITZGERALD, A. P., DE BACKER, G. & DE BACQUER, D. 2009. Recalibration and validation of the SCORE risk chart in the Australian population: the AusSCORE chart. *Eur J Cardiovasc Prev Rehabil,* 16**,** 562-70.

CHOE, E. K., RHEE, H., LEE, S., SHIN, E., OH, S. W., LEE, J. E. & CHOI, S. H. 2018. Metabolic Syndrome Prediction Using Machine Learning Models with Genetic and Clinical Information from a Nonobese Healthy Population. *Genomics Inform,* 16**,** e31.

CHOUINARD, S., POULIN, J., STIP, E. & GODBOUT, R. 2004. Sleep in untreated patients with schizophrenia: a meta-analysis. *Schizophr Bull,* 30**,** 957-67.

CHOUINARD, V. A., HENDERSON, D. C., DALLA MAN, C., VALERI, L., GRAY, B. E., RYAN, K. P., CYPESS, A. M., COBELLI, C., COHEN, B. M. & ONGUR, D. 2019. Impaired insulin signaling in unaffected siblings and patients with first-episode psychosis. *Mol Psychiatry,* 24**,** 1513-1522.

CHRISTODOULOU, E., MA, J., COLLINS, G. S., STEYERBERG, E. W., VERBAKEL, J. Y. & VAN CALSTER, B. 2019. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *J Clin Epidemiol,* 110**,** 12-22.

CHRYSANT, S. G. 2011. A new paradigm in the treatment of the cardiovascular disease continuum: focus on prevention. *Hippokratia,* 15**,** 7-11.

CHURCH, T. S., THOMAS, D. M., TUDOR-LOCKE, C., KATZMARZYK, P. T., EARNEST, C. P., RODARTE, R. Q., MARTIN, C. K., BLAIR, S. N. & BOUCHARD, C. 2011. Trends over 5 decades in U.S. occupation-related physical activity and their associations with obesity. *PLoS One,* 6**,** e19657.

CIOMPI, L. 1980. The natural history of schizophrenia in the long term. *Br J Psychiatry,* 136**,** 413-20.

CLIFT, A. K., COUPLAND, C. A. C., KEOGH, R. H., DIAZ-ORDAZ, K., WILLIAMSON, E., HARRISON, E. M., HAYWARD, A., HEMINGWAY, H., HORBY, P., MEHTA, N., BENGER, J., KHUNTI, K., SPIEGELHALTER, D., SHEIKH, A., VALABHJI, J., LYONS, R. A., ROBSON, J., SEMPLE, M. G., KEE, F., JOHNSON, P., JEBB, S., WILLIAMS, T. & HIPPISLEY-COX, J. 2020. Living risk prediction algorithm (QCOVID) for risk of hospital admission and mortality from coronavirus 19 in adults: national derivation and validation cohort study. *BMJ,* 371**,** m3731.

COHN, G., VALDES, G. & CAPUZZI, D. M. 2001. Pathophysiology and treatment of the dyslipidemia of insulin resistance. *Curr Cardiol Rep,* 3**,** 416-23.

COHRS, S. 2008. Sleep disturbances in patients with schizophrenia : impact and effect of antipsychotics. *CNS Drugs,* 22**,** 939-62.

COLE, S. R., PLATT, R. W., SCHISTERMAN, E. F., CHU, H., WESTREICH, D., RICHARDSON, D. & POOLE, C. 2010. Illustrating bias due to conditioning on a collider. *Int J Epidemiol,* 39**,** 417-20.

COLLINS, G. S., OGUNDIMU, E. O. & ALTMAN, D. G. 2016. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. *Stat Med,* 35**,** 214-26.

COLLINS, G. S., REITSMA, J. B., ALTMAN, D. G. & MOONS, K. G. 2015. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med,* 162**,** 55-63.

COMPTON, M. T. 2004. Duration of untreated psychosis significantly associated with positive symptoms one year after treatment. *Evid Based Ment Health,* 7**,** 101.

CORRELL, C. U., SIKICH, L., REEVES, G., JOHNSON, J., KEETON, C., SPANOS, M., KAPOOR, S., BUSSELL, K., MILLER, L., CHANDRASEKHAR, T., SHERIDAN, E. M., PIRMOHAMED, S., REINBLATT, S. P., ALDERMAN, C., SCHEER, A., BORNER, I., BETHEA, T. C., EDWARDS, S., HAMER, R. M. & RIDDLE, M. A. 2020. Metformin add-on vs. antipsychotic switch vs. continued antipsychotic treatment plus healthy lifestyle education in overweight or obese youth with severe mental illness: results from the IMPACT trial. *World Psychiatry,* 19**,** 69-80.

CORRELL, C. U., SOLMI, M., VERONESE, N., BORTOLATO, B., ROSSON, S., SANTONASTASO, P., THAPA-CHHETRI, N., FORNARO, M., GALLICCHIO, D., COLLANTONI, E., PIGATO, G., FAVARO, A., MONACO, F., KOHLER, C., VANCAMPFORT, D., WARD, P. B., GAUGHRAN, F., CARVALHO, A. F. & STUBBS, B. 2017. Prevalence, incidence and mortality from cardiovascular disease in patients with pooled and specific severe mental illness: a large-scale meta-analysis of 3,211,768 patients and 113,383,368 controls. *World Psychiatry,* 16**,** 163-180.

CRAWFORD, M. J., JAYAKUMAR, S., LEMMEY, S. J., ZALEWSKA, K., PATEL, M. X., COOPER, S. J. & SHIERS, D. 2014. Assessment and treatment of physical health problems among people with schizophrenia: national cross-sectional study. *Br J Psychiatry,* 205**,** 473-7.

CUI, H., JIN, Y., WANG, J., WENG, X. & LI, C. 2012. Serum brain-derived neurotrophic factor (BDNF) levels in schizophrenia: A systematic review. *Shanghai Arch Psychiatry,* 24**,** 250-61.

CURTIS, J., WATKINS, A., ROSENBAUM, S., TEASDALE, S., KALUCY, M., SAMARAS, K. & WARD, P. B. 2016. Evaluating an individualized lifestyle and life skills intervention to prevent antipsychotic-induced weight gain in first-episode psychosis. *Early Interv Psychiatry,* 10**,** 267-76.

DALAL, P. K. & AGARWAL, M. 2015. Postmenopausal syndrome. *Indian J Psychiatry,* 57**,** S222-32.

DAMEN, J. A., HOOFT, L., SCHUIT, E., DEBRAY, T. P., COLLINS, G. S., TZOULAKI, I., LASSALE, C. M., SIONTIS, G. C., CHIOCCHIA, V., ROBERTS, C., SCHLUSSEL, M. M., GERRY, S., BLACK, J. A., HEUS, P., VAN DER SCHOUW, Y. T., PEELEN, L. M. &

MOONS, K. G. 2016. Prediction models for cardiovascular disease risk in the general population: systematic review. *BMJ,* 353**,** i2416.

DANDONA, P., ALJADA, A. & BANDYOPADHYAY, A. 2004. Inflammation: the link between insulin resistance, obesity and diabetes. *Trends Immunol,* 25**,** 4-7.

DARIN-MATTSSON, A., FORS, S. & KAREHOLT, I. 2017. Different indicators of socioeconomic status and their relative importance as determinants of health in old age. *Int J Equity Health,* 16**,** 173.

DAUMIT, G. L., PRONOVOST, P. J., ANTHONY, C. B., GUALLAR, E., STEINWACHS, D. M. & FORD, D. E. 2006. Adverse events during medical and surgical hospitalizations for persons with schizophrenia. *Arch Gen Psychiatry,* 63**,** 267-72.

DAVEY SMITH, G. & EBRAHIM, S. 2005. What can mendelian randomisation tell us about modifiable behavioural and environmental exposures? *BMJ,* 330**,** 1076-9.

DAVIES, N. M., VON HINKE KESSLER SCHOLDER, S., FARBMACHER, H., BURGESS, S., WINDMEIJER, F. & SMITH, G. D. 2015. The many weak instruments problem and Mendelian randomization. *Stat Med,* 34**,** 454-68.

DAVIES, R. W., DANDONA, S., STEWART, A. F., CHEN, L., ELLIS, S. G., TANG, W. H., HAZEN, S. L., ROBERTS, R., MCPHERSON, R. & WELLS, G. A. 2010. Improved prediction of cardiovascular disease based on a panel of single nucleotide polymorphisms identified through genome-wide association studies. *Circ Cardiovasc Genet,* 3**,** 468-74.

DE HERT, M., CORRELL, C. U., BOBES, J., CETKOVICH-BAKMAS, M., COHEN, D., ASAI, I., DETRAUX, J., GAUTAM, S., MOLLER, H. J., NDETEI, D. M., NEWCOMER, J. W., UWAKWE, R. & LEUCHT, S. 2011. Physical illness in patients with severe mental disorders. I. Prevalence, impact of medications and disparities in health care. *World Psychiatry,* 10**,** 52-77.

DE LEON, J. & DIAZ, F. J. 2005. A meta-analysis of worldwide studies demonstrates an association between schizophrenia and tobacco smoking behaviors. *Schizophr Res,* 76**,** 135-57.

DE LUCA, C. & OLEFSKY, J. M. 2008. Inflammation and insulin resistance. *FEBS Lett,* 582**,** 97-105.

DE SILVA, V. A., SURAWEERA, C., RATNATUNGA, S. S., DAYABANDARA, M., WANNIARACHCHI, N. & HANWELLA, R. 2016. Metformin in prevention and treatment of antipsychotic induced weight gain: a systematic review and meta-analysis. *BMC Psychiatry,* 16**,** 341.

DEBOER, M. D. 2011. Ethnicity, obesity and the metabolic syndrome: implications on assessing risk and targeting intervention. *Expert Rev Endocrinol Metab,* 6**,** 279-289.

DEFRONZO, R. A. & FERRANNINI, E. 1991. Insulin resistance. A multifaceted syndrome responsible for NIDDM, obesity, hypertension, dyslipidemia, and atherosclerotic cardiovascular disease. *Diabetes Care,* 14**,** 173-94.

DELAVARI, A., FOROUZANFAR, M. H., ALIKHANI, S., SHARIFIAN, A. & KELISHADI, R. 2009. First nationwide study of the prevalence of the metabolic syndrome and optimal cutoff points of waist circumference in the Middle East: the national survey of risk factors for noncommunicable diseases of Iran. *Diabetes Care,* 32**,** 1092-7.

DELAY, J., DENIKER, P. & HARL, J. M. 1952. [Therapeutic use in psychiatry of phenothiazine of central elective action (4560 RP)]. *Ann Med Psychol (Paris),* 110**,** 112-7.

DEMJAHA, A., MURRAY, R. M., MCGUIRE, P. K., KAPUR, S. & HOWES, O. D. 2012. Dopamine synthesis capacity in patients with treatment-resistant schizophrenia. *Am J Psychiatry,* 169**,** 1203-10.

DENG, M. Y., LAM, S., MEYER, U., FELDON, J., LI, Q., WEI, R., LUK, L., CHUA, S. E., SHAM, P., WANG, Y. & MCALONAN, G. M. 2011. Frontal-subcortical protein expression following prenatal exposure to maternal inflammation. *PLoS One,* 6**,** e16638.

DENG, Y., XIAO, Y., YUAN, F., LIU, Y., JIANG, X., DENG, J., FEJES-TOTH, G., NARAY-FEJES-TOTH, A., CHEN, S., CHEN, Y., YING, H., ZHAI, Q., SHU, Y. & GUO, F. 2018. SGK1/FOXO3 Signaling in Hypothalamic POMC Neurons Mediates Glucocorticoid-Increased Adiposity. *Diabetes,* 67**,** 569-580.

DEPARTMENT OF HEALTH AND SOCIAL CARE. 2019. UK Chief Medical Officers' Physical Activity Guidelines.

DEPARTMENT OF HEALTH. 2014. Annual Report of the Chief Medical Officer 2013: Public Mental Health Priorities: Investing In The Evidence. London, UK.

DHINGRA, R. & VASAN, R. S. 2012. Age as a risk factor. *Med Clin North Am,* 96**,** 87-91.

DICKENS, A. M., SEN, P., KEMPTON, M. J., BARRANTES-VIDAL, N., IYEGBE, C., NORDENTOFT, M., POLLAK, T., RIECHER-ROSSLER, A., RUHRMANN, S., SACHS, G., BRESSAN, R., KREBS, M. O., AMMINGER, G. P., DE HAAN, L., VAN DER GAAG, M., VALMAGGIA, L., HYOTYLAINEN, T., GROUP, E.-G. H. R. S., ORESIC, M. & MCGUIRE, P. 2021. Dysregulated Lipid Metabolism Precedes Onset of Psychosis. *Biol Psychiatry,* 89**,** 288-297.

DICKERSON, F. 2019. Smoking and Schizophrenia: Still a Burning Problem. *Schizophrenia Bulletin,* 45**,** S95.

DING, C., CHAN, Z. & MAGKOS, F. 2016. Lean, but not healthy: the 'metabolically obese, normal-weight' phenotype. *Curr Opin Clin Nutr Metab Care,* 19**,** 408-417.

DODD, G. T. & TIGANIS, T. 2017. Insulin action in the brain: Roles in energy and glucose homeostasis. *J Neuroendocrinol,* 29.

DONG, Y. & PENG, C. Y. 2013. Principled missing data methods for researchers. *Springerplus,* 2**,** 222.

DRURY, K. K., FARRON-RIDGE, C. 1921. Some observations of the types of blood-sugar curves found in different forms of insanity. *Journal of Mental Science,* 71**,** 8-29.

DUNCAN, M. S., FREIBERG, M. S., GREEVY, R. A., JR., KUNDU, S., VASAN, R. S. & TINDLE, H. A. 2019. Association of Smoking Cessation With Subsequent Risk of Cardiovascular Disease. *JAMA,* 322**,** 642-650.

DUNDER, K., LIND, L., ZETHELIUS, B., BERGLUND, L. & LITHELL, H. 2004. Evaluation of a scoring scheme, including proinsulin and the apolipoprotein B/apolipoprotein A1 ratio, for the risk of acute coronary events in middle-aged men: Uppsala Longitudinal Study of Adult Men (ULSAM). *Am Heart J,* 148**,** 596-601.

DUPUIS, J., LANGENBERG, C., PROKOPENKO, I., SAXENA, R., SORANZO, N., JACKSON, A. U., WHEELER, E., GLAZER, N. L., BOUATIA-NAJI, N., GLOYN, A. L., LINDGREN, C. M., MÄGI, R., MORRIS, A. P., RANDALL, J., JOHNSON, T., ELLIOTT, P., RYBIN, D., THORLEIFSSON, G., STEINTHORSDOTTIR, V., HENNEMAN, P., GRALLERT, H., DEHGHAN, A., HOTTENGA, J. J., FRANKLIN, C. S., NAVARRO, P., SONG, K., GOEL, A., PERRY, J. R., EGAN, J. M., LAJUNEN, T., GRARUP, N., SPARSØ, T., DONEY, A., VOIGHT, B. F., STRINGHAM, H. M., LI, M., KANONI, S., SHRADER, P., CAVALCANTI-PROENÇA, C., KUMARI, M., QI, L., TIMPSON, N. J., GIEGER, C., ZABENA, C., ROCHELEAU, G., INGELSSON, E., AN, P., O'CONNELL, J., LUAN, J., ELLIOTT, A., MCCARROLL, S. A., PAYNE, F., ROCCASECCA, R. M., PATTOU, F., SETHUPATHY, P., ARDLIE, K., ARIYUREK, Y., BALKAU, B., BARTER, P., BEILBY, J. P., BEN-SHLOMO, Y., BENEDIKTSSON, R., BENNETT, A. J., BERGMANN, S., BOCHUD, M., BOERWINKLE, E., BONNEFOND, A., BONNYCASTLE, L. L., BORCH-JOHNSEN, K., BÖTTCHER, Y., BRUNNER, E., BUMPSTEAD, S. J., CHARPENTIER, G., CHEN, Y. D., CHINES, P., CLARKE, R., COIN, L. J., COOPER, M. N., CORNELIS, M., CRAWFORD, G., CRISPONI, L., DAY, I. N., DE GEUS, E. J., DELPLANQUE, J., DINA, C., ERDOS, M. R., FEDSON, A. C., FISCHER-ROSINSKY, A., FOROUHI, N. G., FOX, C. S., FRANTS, R., FRANZOSI, M. G., GALAN, P., GOODARZI, M. O., GRAESSLER, J., GROVES, C. J., GRUNDY, S., GWILLIAM, R., GYLLENSTEN, U., HADJADJ, S., et al. 2010. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet,* 42**,** 105-16.

DUVIVIER, B., BOLIJN, J. E., KOSTER, A., SCHALKWIJK, C. G., SAVELBERG, H. & SCHAPER, N. C. 2018. Reducing sitting time versus adding exercise: differential effects on biomarkers of endothelial dysfunction and metabolic risk. *Sci Rep,* 8**,** 8657.

DYRSTAD, S. M., HANSEN, B. H., HOLME, I. M. & ANDERSSEN, S. A. 2014. Comparison of self-reported versus accelerometer-measured physical activity. *Med Sci Sports Exerc,* 46**,** 99-106.

EGAN, M. F., KOJIMA, M., CALLICOTT, J. H., GOLDBERG, T. E., KOLACHANA, B. S., BERTOLINO, A., ZAITSEV, E., GOLD, B., GOLDMAN, D., DEAN, M., LU, B. & WEINBERGER, D. R. 2003. The BDNF val66met polymorphism affects activity-dependent secretion of BDNF and human memory and hippocampal function. *Cell,* 112**,** 257-69.

EIJKELENBOOM, A. & BURGERING, B. M. 2013. FOXOs: signalling integrators for homeostasis maintenance. *Nat Rev Mol Cell Biol,* 14**,** 83-97.

EKELUND, U., STEENE-JOHANNESSEN, J., BROWN, W. J., FAGERLAND, M. W., OWEN, N., POWELL, K. E., BAUMAN, A., LEE, I. M., LANCET PHYSICAL ACTIVITY SERIES 2 EXECUTIVE, C. & LANCET SEDENTARY BEHAVIOUR WORKING, G. 2016. Does physical activity attenuate, or even eliminate, the detrimental association of sitting time with mortality? A harmonised meta-analysis of data from more than 1 million men and women. *Lancet,* 388**,** 1302-10.

ELKS, C. E., DEN HOED, M., ZHAO, J. H., SHARP, S. J., WAREHAM, N. J., LOOS, R. J. & ONG, K. K. 2012. Variability in the heritability of body mass index: a systematic review and meta-regression. *Front Endocrinol (Lausanne),* 3**,** 29.

ENDERS, C. K. 2001a. A Primer on Maximum Likelihood Algorithms Available for Use With Missing Data. *Structural Equation Modeling: A Multidisciplinary Journal,* 8**,** 128-141.

ENDERS, C. K., BANDALOS, D.L. 2001b. The Relative Performance of Full Information Maximum Likelihood Estimation for Missing Data in Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal,* 8**,** 430-457.

ENSOR, J., MARTIN, E.C., RILEY, R.D., 2019. pmsampsize: Calculates the Minimum Sample Size Required for Developing a Multivariable Prediction Model. https://CRAN.R-project.org/package=pmsampsize.

ENTRINGER, S., BUSS, C., SWANSON, J. M., COOPER, D. M., WING, D. A., WAFFARN, F. & WADHWA, P. D. 2012. Fetal programming of body composition, obesity, and metabolic function: the role of intrauterine stress and stress biology. *J Nutr Metab,* 2012**,** 632548.

ERANTI, S. V., MACCABE, J. H., BUNDY, H. & MURRAY, R. M. 2013. Gender difference in age at onset of schizophrenia: a meta-analysis. *Psychol Med,* 43**,** 155-67.

EROLA, J., JALONEN, S., LEHTI, H., 2016. Parental education, class and income over early life course and children's achievement. *Research in Social Stratification and Mobility,* 44**,** 33-43.

EYLES, D. W., TRZASKOWSKI, M., VINKHUYZEN, A. A. E., MATTHEISEN, M., MEIER, S., GOOCH, H., ANGGONO, V., CUI, X., TAN, M. C., BURNE, T. H. J., JANG, S. E., KVASKOFF, D., HOUGAARD, D. M., NORGAARD-PEDERSEN, B., COHEN, A., AGERBO, E., PEDERSEN, C. B., BORGLUM, A. D., MORS, O., SAH, P., WRAY, N. R., MORTENSEN, P. B. & MCGRATH, J. J. 2018. The association between neonatal vitamin D status and risk of schizophrenia. *Sci Rep,* 8**,** 17692.

FAERCH, K., BORCH-JOHNSEN, K., HOLST, J. J. & VAAG, A. 2009. Pathophysiology and aetiology of impaired fasting glycaemia and impaired glucose tolerance: does it matter for prevention and treatment of type 2 diabetes? *Diabetologia,* 52**,** 1714-23.

FAROOQ, S., AGID, O., FOUSSIAS, G. & REMINGTON, G. 2013. Using treatment response to subtype schizophrenia: proposal for a new paradigm in classification. *Schizophr Bull,* 39**,** 1169-72.

FARRELLY, L., FOCKING, M., PIONTKEWITZ, Y., DICKER, P., ENGLISH, J., WYNNE, K., CANNON, M., CAGNEY, G. & COTTER, D. R. 2015. Maternal immune activation induces changes in myelin and metabolic proteins, some of which can be prevented with risperidone in adolescence. *Dev Neurosci,* 37**,** 43-55.

FATEMI, S. H. & FOLSOM, T. D. 2009. The neurodevelopmental hypothesis of schizophrenia, revisited. *Schizophr Bull,* 35**,** 528-48.

FEATHERSTONE, R. E., KAPUR, S. & FLETCHER, P. J. 2007. The amphetamine-induced sensitized state as a model of schizophrenia. *Prog Neuropsychopharmacol Biol Psychiatry,* 31**,** 1556-71.

FERENCE, B. A., GINSBERG, H. N., GRAHAM, I., RAY, K. K., PACKARD, C. J., BRUCKERT, E., HEGELE, R. A., KRAUSS, R. M., RAAL, F. J., SCHUNKERT, H., WATTS, G. F., BOREN, J., FAZIO, S., HORTON, J. D., MASANA, L., NICHOLLS, S. J., NORDESTGAARD, B. G., VAN DE SLUIS, B., TASKINEN, M. R., TOKGOZOGLU, L., LANDMESSER, U., LAUFS, U., WIKLUND, O., STOCK, J. K., CHAPMAN, M. J. & CATAPANO, A. L. 2017. Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. *Eur Heart J,* 38**,** 2459-2472.

FERENTINOS, P. & DIKEOS, D. 2012. Genetic correlates of medical comorbidity associated with schizophrenia and treatment with antipsychotics. *Curr Opin Psychiatry,* 25**,** 381-90.

FERNANDES, B. S., STEINER, J., BERNSTEIN, H. G., DODD, S., PASCO, J. A., DEAN, O. M., NARDIN, P., GONCALVES, C. A. & BERK, M. 2016. C-reactive protein is increased in schizophrenia but is not altered by antipsychotics: meta-analysis and implications. *Mol Psychiatry,* 21**,** 554-64.

FERNANDEZ-ABASCAL, B., SUAREZ-PINILLA, P., COBO-CORRALES, C., CRESPO-FACORRO, B. & SUAREZ-PINILLA, M. 2021. In- and outpatient lifestyle interventions on diet and exercise and their effect on physical and psychological health: a systematic review and meta-analysis of randomised controlled trials in patients with schizophrenia spectrum disorders and first episode of psychosis. *Neurosci Biobehav Rev*.

FERNANDEZ-SOLA, J., LLUIS, M., SACANELLA, E., ESTRUCH, R., ANTUNEZ, E. & URBANO-MARQUEZ, A. 2011. Increased myostatin activity and decreased myocyte proliferation in chronic alcoholic cardiomyopathy. *Alcohol Clin Exp Res,* 35**,** 1220-9.

FERRARIO, C. R. & REAGAN, L. P. 2018. Insulin-mediated synaptic plasticity in the CNS: Anatomical, functional and temporal contexts. *Neuropharmacology,* 136**,** 182-191.

FERRARIO, M., CHIODINI, P., CHAMBLESS, L. E., CESANA, G., VANUZZO, D., PANICO, S., SEGA, R., PILOTTO, L., PALMIERI, L., GIAMPAOLI, S. & GROUP, C. P. R. 2005. Prediction of coronary events in a low incidence population. Assessing accuracy of the CUORE Cohort Study prediction equation. *Int J Epidemiol,* 34**,** 413-21.

FIGLEWICZ, D. P., EVANS, S. B., MURPHY, J., HOEN, M. & BASKIN, D. G. 2003. Expression of receptors for insulin and leptin in the ventral tegmental area/substantia nigra (VTA/SN) of the rat. *Brain Res,* 964**,** 107-15.

FIRTH, J., CARNEY, R., JEROME, L., ELLIOTT, R., FRENCH, P. & YUNG, A. R. 2016. The effects and determinants of exercise participation in first-episode psychosis: a qualitative study. *BMC Psychiatry,* 16**,** 36.

FIRTH, J., SIDDIQI, N., KOYANAGI, A., SISKIND, D., ROSENBAUM, S., GALLETLY, C., ALLAN, S., CANEO, C., CARNEY, R., CARVALHO, A. F., CHATTERTON, M. L., CORRELL, C. U., CURTIS, J., GAUGHRAN, F., HEALD, A., HOARE, E., JACKSON, S. E., KISELY, S., LOVELL, K., MAJ, M., MCGORRY, P. D., MIHALOPOULOS, C., MYLES, H., O'DONOGHUE, B., PILLINGER, T., SARRIS, J., SCHUCH, F. B., SHIERS, D., SMITH, L., SOLMI, M., SUETANI, S., TAYLOR, J., TEASDALE, S. B., THORNICROFT, G., TOROUS, J., USHERWOOD, T., VANCAMPFORT, D., VERONESE, N., WARD, P. B., YUNG, A. R., KILLACKEY, E. & STUBBS, B. 2019. The Lancet Psychiatry Commission: a blueprint for protecting physical health in people with mental illness. *Lancet Psychiatry,* 6**,** 675-712.

FIRTH, J., STUBBS, B., VANCAMPFORT, D., SCHUCH, F. B., ROSENBAUM, S., WARD, P. B., FIRTH, J. A., SARRIS, J. & YUNG, A. R. 2018. The Validity and Value of Self-reported Physical Activity and Accelerometry in People With Schizophrenia: A Population-Scale Study of the UK Biobank. *Schizophr Bull,* 44**,** 1293-1300.

FOLEY, C. N., STALEY, J. R., BREEN, P. G., SUN, B. B., KIRK, P. D. W., BURGESS, S. & HOWSON, J. M. M. 2021. A fast and efficient colocalization algorithm for identifying shared genetic risk factors across multiple traits. *Nat Commun,* 12**,** 764.

FOLEY, C. N., STALEY, J.R., BREEN, P.G., SUN, B.B., KIRK, P.D.W., BURGESS, S., HOWSON, J.M.M. 2019. A fast and efficient colocalization algorithm for identifying shared genetic risk factors across multiple traits. *bioRxiv*.

FRASER, A., MACDONALD-WALLIS, C., TILLING, K., BOYD, A., GOLDING, J., DAVEY SMITH, G., HENDERSON, J., MACLEOD, J., MOLLOY, L., NESS, A., RING, S., NELSON, S. M. & LAWLOR, D. A. 2013. Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol,* 42**,** 97-110.

FRAYLING, T. M., TIMPSON, N. J., WEEDON, M. N., ZEGGINI, E., FREATHY, R. M., LINDGREN, C. M., PERRY, J. R., ELLIOTT, K. S., LANGO, H., RAYNER, N. W., SHIELDS, B., HARRIES, L. W., BARRETT, J. C., ELLARD, S., GROVES, C. J., KNIGHT, B., PATCH, A. M., NESS, A. R., EBRAHIM, S., LAWLOR, D. A., RING, S. M., BEN-SHLOMO, Y., JARVELIN, M. R., SOVIO, U., BENNETT, A. J., MELZER, D., FERRUCCI, L., LOOS, R. J., BARROSO, I., WAREHAM, N. J., KARPE, F., OWEN, K. R., CARDON, L. R., WALKER, M., HITMAN, G. A., PALMER, C. N., DONEY, A. S., MORRIS, A. D., SMITH, G. D., HATTERSLEY, A. T. & MCCARTHY, M. I. 2007. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science,* 316**,** 889-94.

FREEDMAN, L. S., SCHATZKIN, A., MIDTHUNE, D. & KIPNIS, V. 2011. Dealing with dietary measurement error in nutritional cohort studies. *J Natl Cancer Inst,* 103**,** 1086-92.

FREEDMAN, R. 2014. alpha7-nicotinic acetylcholine receptor agonists for cognitive enhancement in schizophrenia. *Annu Rev Med,* 65**,** 245-61.

FRIEDLAND, D. R., CEDERBERG, C. & TARIMA, S. 2009. Audiometric pattern as a predictor of cardiovascular status: development of a model for assessment of risk. *Laryngoscope,* 119**,** 473-86.

FRITZ, M. S. & MACKINNON, D. P. 2007. Required sample size to detect the mediated effect. *Psychol Sci,* 18**,** 233-9.

FUJISHIRO, M., GOTOH, Y., KATAGIRI, H., SAKODA, H., OGIHARA, T., ANAI, M., ONISHI, Y., ONO, H., ABE, M., SHOJIMA, N., FUKUSHIMA, Y., KIKUCHI, M., OKA,

Y. & ASANO, T. 2003. Three mitogen-activated protein kinases inhibit insulin signaling by different mechanisms in 3T3-L1 adipocytes. *Mol Endocrinol,* 17**,** 487-97.

FUKUMA, S., SHIMIZU, S., SHINTANI, A., KAMITANI, T., AKIZAWA, T. & FUKUHARA, S. 2018. Development and validation of a prediction model for loss of physical function in elderly hemodialysis patients. *Nephrol Dial Transplant,* 33**,** 1452-1458.

FUNG, T. T., WILLETT, W. C., STAMPFER, M. J., MANSON, J. E. & HU, F. B. 2001. Dietary patterns and the risk of coronary heart disease in women. *Arch Intern Med,* 161**,** 1857-62.

FUSAR-POLI, P., BONOLDI, I., YUNG, A. R., BORGWARDT, S., KEMPTON, M. J., VALMAGGIA, L., BARALE, F., CAVERZASI, E. & MCGUIRE, P. 2012. Predicting psychosis: meta-analysis of transition outcomes in individuals at high clinical risk. *Arch Gen Psychiatry,* 69**,** 220-9.

GAO, W. G., DONG, Y. H., PANG, Z. C., NAN, H. R., WANG, S. J., REN, J., ZHANG, L., TUOMILEHTO, J. & QIAO, Q. 2010. A simple Chinese risk score for undiagnosed diabetes. *Diabet Med,* 27**,** 274-81.

GAO, W. G., QIAO, Q., PITKANIEMI, J., WILD, S., MAGLIANO, D., SHAW, J., SODERBERG, S., ZIMMET, P., CHITSON, P., KNOWLESSUR, S., ALBERTI, G. & TUOMILEHTO, J. 2009. Risk prediction models for the development of diabetes in Mauritian Indians. *Diabet Med,* 26**,** 996-1002.

GARIEPY, G., NITKA, D. & SCHMITZ, N. 2010. The association between obesity and anxiety disorders in the population: a systematic review and meta-analysis. *Int J Obes (Lond),* 34**,** 407-19.

GEHART, H., KUMPF, S., ITTNER, A. & RICCI, R. 2010. MAPK signalling in cellular metabolism: stress or wellness? *EMBO Rep,* 11**,** 834-40.

GELONEZE, B., REPETTO, E. M., GELONEZE, S. R., TAMBASCIA, M. A. & ERMETICE, M. N. 2006. The threshold value for insulin resistance (HOMA-IR) in an admixtured population IR in the Brazilian Metabolic Syndrome Study. *Diabetes Res Clin Pract,* 72**,** 219-20.

GEORGAKIS, M. K., MALIK, R., GILL, D., FRANCESCHINI, N., SUDLOW, C. L. M., DICHGANS, M. & INVENT CONSORTIUM, C. I. W. G. 2020. Interleukin-6 Signaling Effects on Ischemic Stroke and Other Cardiovascular Outcomes: A Mendelian Randomization Study. *Circ Genom Precis Med,* 13**,** e002872.

GEORGE, A. M., JACOB, A. G. & FOGELFELD, L. 2015. Lean diabetes mellitus: An emerging entity in the era of obesity. *World J Diabetes,* 6**,** 613-20.

GIAMBARTOLOMEI, C., ZHENLI LIU, J., ZHANG, W., HAUBERG, M., SHI, H., BOOCOCK, J., PICKRELL, J., JAFFE, A. E., COMMONMIND, C., PASANIUC, B. & ROUSSOS, P.

2018. A Bayesian framework for multiple trait colocalization from summary association statistics. *Bioinformatics,* 34**,** 2538-2545.

GIBSON-SMITH, D., HALLDORSSON, T. I., BOT, M., BROUWER, I. A., VISSER, M., THORSDOTTIR, I., BIRGISDOTTIR, B. E., GUDNASON, V., EIRIKSDOTTIR, G., LAUNER, L. J., HARRIS, T. B. & GUNNARSDOTTIR, I. 2020. Childhood overweight and obesity and the risk of depression across the lifespan. *BMC Pediatr,* 20**,** 25.

GILL, D., GEORGAKIS, M. K., WALKER, V. M., SCHMIDT, A. F., GKATZIONIS, A., FREITAG, D. F., FINAN, C., HINGORANI, A. D., HOWSON, J. M. M., BURGESS, S., SWERDLOW, D. I., DAVEY SMITH, G., HOLMES, M. V., DICHGANS, M., SCOTT, R. A., ZHENG, J., PSATY, B. M. & DAVIES, N. M. 2021. Mendelian randomization for studying the effects of perturbing drug targets. *Wellcome Open Res,* 6**,** 16.

GILLESPIE, A. L., SAMANAITE, R., MILL, J., EGERTON, A. & MACCABE, J. H. 2017. Is treatment-resistant schizophrenia categorically distinct from treatment-responsive schizophrenia? a systematic review. *BMC Psychiatry,* 17**,** 12.

GOLDMAN, L. S. 1999. Medical illness in patients with schizophrenia. *J Clin Psychiatry,* 60 Suppl 21**,** 10-5.

GOLDSMITH, D. R., HAROON, E., MILLER, A. H., ADDINGTON, J., BEARDEN, C., CADENHEAD, K., CANNON, T., CORNBLATT, B., MATHALON, D., MCGLASHAN, T., SEIDMAN, L., TSUANG, M., WOODS, S. W., WALKER, E. F. & PERKINS, D. O. 2019. Association of baseline inflammatory markers and the development of negative symptoms in individuals at clinical high risk for psychosis. *Brain Behav Immun,* 76**,** 268-274.

GOLDSTEIN, J. L. & BROWN, M. S. 1974. Binding and degradation of low density lipoproteins by cultured human fibroblasts. Comparison of cells from a normal subject and from a patient with homozygous familial hypercholesterolemia. *J Biol Chem,* 249**,** 5153-62.

GONCALVES, A., JHUND, P. S., CLAGGETT, B., SHAH, A. M., KONETY, S., BUTLER, K., KITZMAN, D. W., ROSAMOND, W., FUCHS, F. D. & SOLOMON, S. D. 2015. Relationship between alcohol consumption and cardiac structure and function in the elderly: the Atherosclerosis Risk In Communities Study. *Circ Cardiovasc Imaging,* 8.

GONZALEZ-CANTERO, J., MARTIN-RODRIGUEZ, J. L., GONZALEZ-CANTERO, A., ARREBOLA, J. P. & GONZALEZ-CALVIN, J. L. 2018. Insulin resistance in lean and overweight non-diabetic Caucasian adults: Study of its relationship with liver triglyceride content, waist circumference and BMI. *PLoS One,* 13**,** e0192663.

GOODMAN, R. 2001. Psychometric properties of the strengths and difficulties questionnaire. *J Am Acad Child Adolesc Psychiatry,* 40**,** 1337-45.

GORCZYNSKI, P. & FAULKNER, G. 2010. Exercise therapy for schizophrenia. *Cochrane Database Syst Rev***,** CD004412.

GOUGH, S. C. & O'DONOVAN, M. C. 2005. Clustering of metabolic comorbidity in schizophrenia: a genetic contribution? *J Psychopharmacol,* 19**,** 47-55.

GREENHALGH, A. M., GONZALEZ-BLANCO, L., GARCIA-RIZO, C., FERNANDEZ-EGEA, E., MILLER, B., ARROYO, M. B. & KIRKPATRICK, B. 2017. Meta-analysis of glucose tolerance, insulin, and insulin resistance in antipsychotic-naive patients with nonaffective psychosis. *Schizophr Res,* 179**,** 57-63.

GRIFFIN, S. J., LITTLE, P. S., HALES, C. N., KINMONTH, A. L. & WAREHAM, N. J. 2000. Diabetes risk score: towards earlier detection of type 2 diabetes in general practice. *Diabetes Metab Res Rev,* 16**,** 164-71.

GRILLO, C. A., PIROLI, G. G., LAWRENCE, R. C., WRIGHTEN, S. A., GREEN, A. J., WILSON, S. P., SAKAI, R. R., KELLY, S. J., WILSON, M. A., MOTT, D. D. & REAGAN, L. P. 2015. Hippocampal Insulin Resistance Impairs Spatial Learning and Synaptic Plasticity. *Diabetes,* 64**,** 3927-36.

GRUNDY, S. M., CLEEMAN, J. I., DANIELS, S. R., DONATO, K. A., ECKEL, R. H., FRANKLIN, B. A., GORDON, D. J., KRAUSS, R. M., SAVAGE, P. J., SMITH, S. C., JR., SPERTUS, J. A., COSTA, F., AMERICAN HEART, A., NATIONAL HEART, L. & BLOOD, I. 2005. Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement. *Circulation,* 112**,** 2735-52.

GUAN, F., ZHANG, T., LI, L., FU, D., LIN, H., CHEN, G. & CHEN, T. 2016. Two-stage replication of previous genome-wide association studies of AS3MT-CNNM2-NT5C2 gene cluster region in a large schizophrenia case-control sample from Han Chinese population. *Schizophr Res,* 176**,** 125-130.

GUPTA, P., PRIETO-MERINO, D., AJAY, V.S., SINGH, K., ROY, A., KRISHNAN, A., NARAYAN, V.N., ALI, M.K., TANDON, N., DORAIRAJ, P., PEREL, P. 2019. Cardiovascular Risk Prediction in India: Comparison of the Original and Recalibrated Framingham Prognostic Models in Urban Populations. [Version 1; peer review : 2 approved with reservations]. *Wellcome Open Research,* 4**,** 1-14.

GUPTE, A. A., POWNALL, H. J. & HAMILTON, D. J. 2015. Estrogen: an emerging regulator of insulin action and mitochondrial function. *J Diabetes Res,* 2015**,** 916585.

GURILLO, P., JAUHAR, S., MURRAY, R. M. & MACCABE, J. H. 2015. Does tobacco use cause psychosis? Systematic review and meta-analysis. *Lancet Psychiatry,* 2**,** 718-725.

HA, K. H., LEE, Y. H., SONG, S. O., LEE, J. W., KIM, D. W., CHO, K. H. & KIM, D. J. 2018. Development and Validation of the Korean Diabetes Risk Score: A 10-Year National Cohort Study. *Diabetes Metab J,* 42**,** 402-414.

HACKINGER, S., PRINS, B., MAMAKOU, V., ZENGINI, E., MAROULI, E., BRCIC, L., SERAFETINIDIS, I., LAMNISSOU, K., KONTAXAKIS, V., DEDOUSSIS, G., GONIDAKIS, F., THANOPOULOU, A., TENTOLOURIS, N., TSEZOU, A. & ZEGGINI, E. 2018. Evidence for genetic contribution to the increased risk of type 2 diabetes in schizophrenia. *Transl Psychiatry,* 8**,** 252.

HADDAWAY, N. R., COLLINS, A. M., COUGHLIN, D. & KIRK, S. 2015. The Role of Google Scholar in Evidence Reviews and Its Applicability to Grey Literature Searching. *PLoS One,* 10**,** e0138237.

HAPPELL, B., SCOTT, D. & PLATANIA-PHUNG, C. 2012. Perceptions of barriers to physical health care for people with serious mental illness: a review of the international literature. *Issues Ment Health Nurs,* 33**,** 752-61.

HARRELL, F. E., JR. 2001. *Regressing Model Strategies: With Applications to Linearc Models, Logistic Regression, and Survival Analysis,* New York, Springer-Verlag.

HARRIS, P. A., TAYLOR, R., MINOR, B. L., ELLIOTT, V., FERNANDEZ, M., O'NEAL, L., MCLEOD, L., DELACQUA, G., DELACQUA, F., KIRBY, J., DUDA, S. N. & CONSORTIUM, R. E. 2019. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform,* 95**,** 103208.

HARRIS, P. A., TAYLOR, R., THIELKE, R., PAYNE, J., GONZALEZ, N. & CONDE, J. G. 2009. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform,* 42**,** 377-81.

HARTWIG, F. P., BORGES, M. C., HORTA, B. L., BOWDEN, J. & DAVEY SMITH, G. 2017. Inflammatory Biomarkers and Risk of Schizophrenia: A 2-Sample Mendelian Randomization Study. *JAMA Psychiatry,* 74**,** 1226-1233.

HARTZ, S. M., PATO, C. N., MEDEIROS, H., CAVAZOS-REHG, P., SOBELL, J. L., KNOWLES, J. A., BIERUT, L. J., PATO, M. T. & GENOMIC PSYCHIATRY COHORT, C. 2014. Comorbidity of severe psychotic disorders with measures of substance use. *JAMA Psychiatry,* 71**,** 248-54.

HARVEY, L., BURNE, T. H., MCGRATH, J. J. & EYLES, D. W. 2010. Developmental vitamin D3 deficiency induces alterations in immune organ morphology and function in adult offspring. *J Steroid Biochem Mol Biol,* 121**,** 239-42.

HARVEY, R. C., JAMES, A. C. & SHIELDS, G. E. 2016. A Systematic Review and Network Meta-Analysis to Assess the Relative Efficacy of Antipsychotics for the Treatment of Positive and Negative Symptoms in Early-Onset Schizophrenia. *CNS Drugs,* 30**,** 27-39.

HAYES, J. F., LUNDIN, A., WICKS, S., LEWIS, G., WONG, I. C. K., OSBORN, D. P. J. & DALMAN, C. 2019. Association of Hydroxylmethyl Glutaryl Coenzyme A Reductase Inhibitors, L-Type Calcium Channel Antagonists, and Biguanides With Rates of Psychiatric Hospitalization and Self-Harm in Individuals With Serious Mental Illness. *JAMA Psychiatry*.

HAYES, J. F., MARSTON, L., WALTERS, K., KING, M. B. & OSBORN, D. P. J. 2017. Mortality gap for people with bipolar disorder and schizophrenia: UK-based cohort study 2000-2014. *Br J Psychiatry,* 211**,** 175-181.

HEALD, A., PENDLEBURY, J., ANDERSON, S., NARAYAN, V., GUY, M., GIBSON, M., HADDAD, P. & LIVINGSTON, M. 2017. Lifestyle factors and the metabolic syndrome in Schizophrenia: a cross-sectional study. *Ann Gen Psychiatry,* 16**,** 12.

HEDBLAD, B., NILSSON, P., JANZON, L. & BERGLUND, G. 2000. Relation between insulin resistance and carotid intima-media thickness and stenosis in non-diabetic subjects. Results from a cross-sectional study in Malmo, Sweden. *Diabet Med,* 17**,** 299-307.

HEDELIN, M., LOF, M., OLSSON, M., LEWANDER, T., NILSSON, B., HULTMAN, C. M. & WEIDERPASS, E. 2010. Dietary intake of fish, omega-3, omega-6 polyunsaturated fatty acids and vitamin D and the prevalence of psychotic-like symptoms in a cohort of 33,000 women from the general population. *BMC Psychiatry,* 10**,** 38.

HELLWEGE, J. N., KEATON, J. M., GIRI, A., GAO, X., VELEZ EDWARDS, D. R. & EDWARDS, T. L. 2017. Population Stratification in Genetic Association Studies. *Curr Protoc Hum Genet,* 95**,** 1 22 1-1 22 23.

HEMANI, G., BOWDEN, J. & DAVEY SMITH, G. 2018. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum Mol Genet,* 27**,** R195-R208.

HERMSDORFF, H. H., ZULET, M. A., PUCHAU, B. & MARTINEZ, J. A. 2011. Central adiposity rather than total adiposity measurements are specifically involved in the inflammatory status from healthy young adults. *Inflammation,* 34**,** 161-70.

HERRERA, B. M., KEILDSON, S. & LINDGREN, C. M. 2011. Genetics and epigenetics of obesity. *Maturitas,* 69**,** 41-9.

HERS, I., VINCENT, E. E. & TAVARE, J. M. 2011. Akt signalling in health and disease. *Cell Signal,* 23**,** 1515-27.

HESTON, L. L. 1966. Psychiatric disorders in foster home reared children of schizophrenic mothers. *Br J Psychiatry,* 112**,** 819-25.

HILDRUM, B., MYKLETUN, A., DAHL, A. A. & MIDTHJELL, K. 2009. Metabolic syndrome and risk of mortality in middle-aged versus elderly individuals: the Nord-Trondelag Health Study (HUNT). *Diabetologia,* 52**,** 583-90.

HILKER, R., HELENIUS, D., FAGERLUND, B., SKYTTHE, A., CHRISTENSEN, K., WERGE, T. M., NORDENTOFT, M. & GLENTHOJ, B. 2018. Heritability of Schizophrenia and Schizophrenia Spectrum Based on the Nationwide Danish Twin Register. *Biol Psychiatry,* 83**,** 492-498.

HILL, A. B. 1965. THE ENVIRONMENT AND DISEASE: ASSOCIATION OR CAUSATION? *Proc R Soc Med,* 58**,** 295-300.

HIPPISLEY-COX, J. & COUPLAND, C. 2017. Development and validation of QDiabetes-2018 risk prediction algorithm to estimate future risk of type 2 diabetes: cohort study. *BMJ,* 359**,** j5019.

HIPPISLEY-COX, J., COUPLAND, C. & BRINDLE, P. 2017. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *BMJ,* 357**,** j2099.

HIPPISLEY-COX, J., COUPLAND, C., VINOGRADOVA, Y., ROBSON, J., MAY, M. & BRINDLE, P. 2007. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. *BMJ,* 335**,** 136.

HJORTHOJ, C., OSTERGAARD, M. L., BENROS, M. E., TOFTDAHL, N. G., ERLANGSEN, A., ANDERSEN, J. T. & NORDENTOFT, M. 2015. Association between alcohol and substance use disorders and all-cause and cause-specific mortality in schizophrenia, bipolar disorder, and unipolar depression: a nationwide, prospective, register-based study. *Lancet Psychiatry,* 2**,** 801-8.

HO, P. A., DAHLE, D. N. & NOORDSY, D. L. 2018. Why Do People With Schizophrenia Exercise? A Mixed Methods Analysis Among Community Dwelling Regular Exercisers. *Front Psychiatry,* 9**,** 596.

HOCHLEHNERT, A., NIEHOFF, D., WILD, B., JUNGER, J., HERZOG, W. & LOWE, B. 2011. Psychiatric comorbidity in cardiovascular inpatients: costs, net gain, and length of hospitalization. *J Psychosom Res,* 70**,** 135-9.

HOLFORD, T. R. 1991. Understanding the effects of age, period, and cohort on incidence and mortality rates. *Annu Rev Public Health,* 12**,** 425-57.

HOLLAND, B. S., COPENHAVER, M. 1987. An improved sequentially rejective Bonferroni test procedure. *Biometrics,* 43**,** 417-423.

HOLM, S. 1979. A Simple Sequentially Rejective Multiple Test Procedure. 6**,** 65-70.

HOLZMANN, J., HENNCHEN, M. & ROHRER, H. 2015. Prox1 identifies proliferating neuroblasts and nascent neurons during neurogenesis in sympathetic ganglia. *Dev Neurobiol,* 75**,** 1352-67.

HONG, E. P. & PARK, J. W. 2012. Sample size and statistical power calculation in genetic association studies. *Genomics Inform,* 10**,** 117-22.

HONG, K. W., LIM, J. E., GO, M. J., SHIN CHO, Y., AHN, Y., HAN, B. G. & OH, B. 2012. Recapitulation of the association of the Val66Met polymorphism of BDNF gene with BMI in Koreans. *Obesity (Silver Spring),* 20**,** 1871-5.

HOSSAIN, R., MAHMUS, S.M.H., HOSSIN, M,A., NOORI, S.R.H., JAHAN, H. 2018. PRMT: Predicting Risk Factor of Obesity among Middle-Aged People Using Data Mining Techniques. *Procedia Computer Science,* 132**,** 1068-1076.

HOU, Y., LIANG, W., ZHANG, J., LI, Q., OU, H., WANG, Z., LI, S., HUANG, X. & ZHAO, C. 2018. Schizophrenia-associated rs4702 G allele-specific downregulation of FURIN expression by miR-338-3p reduces BDNF production. *Schizophr Res,* 199**,** 176-180.

HOWES, O. D. & KAPUR, S. 2009. The dopamine hypothesis of schizophrenia: version III--the final common pathway. *Schizophr Bull,* 35**,** 549-62.

HOWES, O. D. & KAPUR, S. 2014. A neurobiological hypothesis for the classification of schizophrenia: type A (hyperdopaminergic) and type B (normodopaminergic). *Br J Psychiatry,* 205**,** 1-3.

HOWES, O. D., WILLIAMS, M., IBRAHIM, K., LEUNG, G., EGERTON, A., MCGUIRE, P. K. & TURKHEIMER, F. 2013. Midbrain dopamine function in schizophrenia and depression: a post-mortem and positron emission tomographic imaging study. *Brain,* 136**,** 3242-51.

HSIAO, E. Y. & PATTERSON, P. H. 2011. Activation of the maternal immune system induces endocrine changes in the placenta via IL-6. *Brain Behav Immun,* 25**,** 604-15.

HUANG, T., SHU, Y. & CAI, Y. D. 2015. Genetic differences among ethnic groups. *BMC Genomics,* 16**,** 1093.

HUHN, M., NIKOLAKOPOULOU, A., SCHNEIDER-THOMA, J., KRAUSE, M., SAMARA, M., PETER, N., ARNDT, T., BACKERS, L., ROTHE, P., CIPRIANI, A., DAVIS, J., SALANTI, G. & LEUCHT, S. 2019. Comparative efficacy and tolerability of 32 oral antipsychotics for the acute treatment of adults with multi-episode schizophrenia: a systematic review and network meta-analysis. *Lancet,* 394**,** 939-951.

HUI, C., WARDWELL, B. & TSAI, G. E. 2009. Novel therapies for schizophrenia: understanding the glutamatergic synapse and potential targets for altering N-methyl-D-aspartate neurotransmission. *Recent Pat CNS Drug Discov,* 4**,** 220-38.

HUNT, G. E., LARGE, M. M., CLEARY, M., LAI, H. M. X. & SAUNDERS, J. B. 2018. Prevalence of comorbid substance use in schizophrenia spectrum disorders in community and clinical settings, 1990-2017: Systematic review and meta-analysis. *Drug Alcohol Depend,* 191**,** 234-258.

HUTTER, N., SCHNURR, A. & BAUMEISTER, H. 2010. Healthcare costs in patients with diabetes mellitus and comorbid mental disorders--a systematic review. *Diabetologia,* 53**,** 2470-9.

ISOMAA, B., ALMGREN, P., TUOMI, T., FORSEN, B., LAHTI, K., NISSEN, M., TASKINEN, M. R. & GROOP, L. 2001. Cardiovascular morbidity and mortality associated with the metabolic syndrome. *Diabetes Care,* 24**,** 683-9.

JABLENSKY, A. 2010. The diagnostic concept of schizophrenia: its history, evolution, and future prospects. *Dialogues Clin Neurosci,* 12**,** 271-87.

JACKSON, H. C., NEEDHAM, A. M., HUTCHINS, L. J., MAZURKIEWICZ, S. E. & HEAL, D. J. 1997. Comparison of the effects of sibutramine and other monoamine reuptake inhibitors on food intake in the rat. *Br J Pharmacol,* 121**,** 1758-62.

JEPPESEN, R., CHRISTENSEN, R. H. B., PEDERSEN, E. M. J., NORDENTOFT, M., HJORTHOJ, C., KOHLER-FORSBERG, O. & BENROS, M. E. 2020. Efficacy and safety of anti-inflammatory agents in treatment of psychotic disorders - A comprehensive systematic review and meta-analysis. *Brain Behav Immun,* 90**,** 364-380.

JESTE, D. V., GLADSJO, J. A., LINDAMER, L. A. & LACRO, J. P. 1996. Medical comorbidity in schizophrenia. *Schizophr Bull,* 22**,** 413-30.

JONAS, K. G., FOCHTMANN, L. J., PERLMAN, G., TIAN, Y., KANE, J. M., BROMET, E. J. & KOTOV, R. 2020. Lead-Time Bias Confounds Association Between Duration of Untreated Psychosis and Illness Course in Schizophrenia. *Am J Psychiatry,* 177**,** 327-334.

JONES, H. J., STERGIAKOULI, E., TANSEY, K. E., HUBBARD, L., HERON, J., CANNON, M., HOLMANS, P., LEWIS, G., LINDEN, D. E., JONES, P. B., DAVEY SMITH, G., O'DONOVAN, M. C., OWEN, M. J., WALTERS, J. T. & ZAMMIT, S. 2016. Phenotypic Manifestation of Genetic Risk for Schizophrenia During Adolescence in the General Population. *JAMA Psychiatry,* 73**,** 221-8.

JONES, S., HOWARD, L. & THORNICROFT, G. 2008. 'Diagnostic overshadowing': worse physical health care for people with mental illness. *Acta Psychiatr Scand,* 118**,** 169-71.

JUNG, T., WICKRAMA, K.A.S. 2007. An Introduction to Latent Class Growth Analysis and Growth Mixture Modeling. *Social and Personality Psychology Compass,* 2**,** 302-317.

JURCOVICOVA, J. 2014. Glucose transport in brain - effect of inflammation. *Endocr Regul,* 48**,** 35-48.

KAHN, R. S., SOMMER, I. E., MURRAY, R. M., MEYER-LINDENBERG, A., WEINBERGER, D. R., CANNON, T. D., O'DONOVAN, M., CORRELL, C. U., KANE, J. M., VAN OS, J. & INSEL, T. R. 2015. Schizophrenia. *Nat Rev Dis Primers,* 1**,** 15067.

KAN, C., PEDERSEN, N. L., CHRISTENSEN, K., BORNSTEIN, S. R., LICINIO, J., MACCABE, J. H., ISMAIL, K. & RIJSDIJK, F. 2016. Genetic overlap between type 2 diabetes and depression in Swedish and Danish twin registries. *Mol Psychiatry,* 21**,** 903-9.

KANE, J. M. & SHARIF, Z. A. 2008. Atypical antipsychotics: sedation versus efficacy. *J Clin Psychiatry,* 69 Suppl 1**,** 18-31.

KAPPELMANN, N., KHANDAKER, G. M., DAL, H., STOCHL, J., KOSIDOU, K., JONES, P. B., DALMAN, C. & KARLSSON, H. 2019. Systemic inflammation and intelligence in early adulthood and subsequent risk of schizophrenia and other non-affective psychoses: a longitudinal cohort and co-relative study. *Psychol Med,* 49**,** 295-302.

KARAGEORGIOU, V., MILAS, G. P. & MICHOPOULOS, I. 2018. Neutrophil-to-lymphocyte ratio in schizophrenia: A systematic review and meta-analysis. *Schizophr Res*.

KARSON, C., DUFFY, R. A., ERAMO, A., NYLANDER, A. G. & OFFORD, S. J. 2016. Long-term outcomes of antipsychotic treatment in patients with first-episode schizophrenia: a systematic review. *Neuropsychiatr Dis Treat,* 12**,** 57-67.

KATO, M., KAJIMURA, N., OKUMA, T., SEKIMOTO, M., WATANABE, T., YAMADERA, H. & TAKAHASHI, K. 1999. Association between delta waves during sleep and negative symptoms in schizophrenia. Pharmaco-eeg studies by using structurally different hypnotics. *Neuropsychobiology,* 39**,** 165-72.

KAYMAZ, N., DRUKKER, M., LIEB, R., WITTCHEN, H. U., WERBELOFF, N., WEISER, M., LATASTER, T. & VAN OS, J. 2012. Do subthreshold psychotic experiences predict clinical outcomes in unselected non-help-seeking population-based samples? A systematic review and meta-analysis, enriched with new results. *Psychol Med,* 42**,** 2239-53.

KEINANEN, J., MANTERE, O., KIESEPPA, T., MANTYLA, T., TORNIAINEN, M., LINDGREN, M., SUNDVALL, J. & SUVISAARI, J. 2015. Early insulin resistance predicts weight gain and waist circumference increase in first-episode psychosis--A one year follow-up study. *Schizophr Res,* 169**,** 458-463.

KEMP, V., FISHER, C., LAWN, S., BATTERSBY, M., ISAAC, M. 2014. Small Steps: Barriers and Facilitators to Physical Health Self-Management by People Living with Mental Illness. *Int J Ment Health Promot,* 16**,** 216-230.

KERIMOV, N., HAYHURST, J. D., MANNING, J. R., WALTER, P., KOLBERG, L., PEIKOVA, K., SAMOVICA, M., BURDETT, T., JUPP, S., PARKINSON, H., PAPATHEODOROU,

I., ZERBINO, D. & ALASOO, K. 2020. eQTL Catalogue: A compendium of uniformly processed human gene expression and splicing QTLs. *bioRxiv*.

KERNER, B. 2015. Comorbid substance use disorders in schizophrenia: a latent class approach. *Psychiatry Res,* 225**,** 395-401.

KERSHAW, K. N., BRENES, G. A., CHARLES, L. E., CODAY, M., DAVIGLUS, M. L., DENBURG, N. L., KROENKE, C. H., SAFFORD, M. M., SAVLA, T., TINDLE, H. A., TINKER, L. F. & VAN HORN, L. 2014. Associations of stressful life events and social strain with incident cardiovascular disease in the Women's Health Initiative. *J Am Heart Assoc,* 3**,** e000687.

KHANDAKER, G. M., PEARSON, R. M., ZAMMIT, S., LEWIS, G. & JONES, P. B. 2014. Association of serum interleukin 6 and C-reactive protein in childhood with depression and psychosis in young adult life: a population-based longitudinal study.

KHANDAKER, G. M., ZAMMIT, S., BURGESS, S., LEWIS, G. & JONES, P. B. 2017. Association between a functional interleukin 6 receptor genetic variant and risk of depression and psychosis in a population-based birth cohort. *Brain Behav Immun*.

KHANDAKER, G. M., ZIMBRON, J., DALMAN, C., LEWIS, G. & JONES, P. B. 2012. Childhood infection and adult schizophrenia: a meta-analysis of population-based studies. *Schizophr Res,* 139**,** 161-8.

KHANDAKER, G. M., ZIMBRON, J., LEWIS, G. & JONES, P. B. 2013. Prenatal maternal infection, neurodevelopment and adult schizophrenia: a systematic review of population-based studies. *Psychol Med,* 43**,** 239-57.

KHANDAKER, G. M., ZUBER, V., REES, J. M. B., CARVALHO, L., MASON, A. M., FOLEY, C. N., GKATZIONIS, A., JONES, P. B. & BURGESS, S. 2019. Shared mechanisms between coronary heart disease and depression: findings from a large UK general population-based cohort. *Mol Psychiatry*.

KHANTZIAN, E. J. 1997. The self-medication hypothesis of substance use disorders: a reconsideration and recent applications. *Harv Rev Psychiatry,* 4**,** 231-44.

KHAW, K. T. 2008. *Rose's Strategy of Preventative Medicine,* USA, Oxford University Press.

KILPELÄINEN, T. O., CARLI, J. F., SKOWRONSKI, A. A., SUN, Q., KRIEBEL, J., FEITOSA, M. F., HEDMAN, Å., DRONG, A. W., HAYES, J. E., ZHAO, J., PERS, T. H., SCHICK, U., GRARUP, N., KUTALIK, Z., TROMPET, S., MANGINO, M., KRISTIANSSON, K., BEEKMAN, M., LYYTIKÄINEN, L. P., ERIKSSON, J., HENNEMAN, P., LAHTI, J., TANAKA, T., LUAN, J., DEL GRECO M, F., PASKO, D., RENSTRÖM, F., WILLEMS, S. M., MAHAJAN, A., ROSE, L. M., GUO, X., LIU, Y., KLEBER, M. E., PÉRUSSE, L., GAUNT, T., AHLUWALIA, T. S., JU SUNG, Y., RAMOS, Y. F., AMIN, N., AMUZU, A.,

BARROSO, I., BELLIS, C., BLANGERO, J., BUCKLEY, B. M., BÖHRINGER, S., I CHEN, Y. D., DE CRAEN, A. J., CROSSLIN, D. R., DALE, C. E., DASTANI, Z., DAY, F. R., DEELEN, J., DELGADO, G. E., DEMIRKAN, A., FINUCANE, F. M., FORD, I., GARCIA, M. E., GIEGER, C., GUSTAFSSON, S., HALLMANS, G., HANKINSON, S. E., HAVULINNA, A. S., HERDER, C., HERNANDEZ, D., HICKS, A. A., HUNTER, D. J., ILLIG, T., INGELSSON, E., IOAN-FACSINAY, A., JANSSON, J. O., JENNY, N. S., JØRGENSEN, M. E., JØRGENSEN, T., KARLSSON, M., KOENIG, W., KRAFT, P., KWEKKEBOOM, J., LAATIKAINEN, T., LADWIG, K. H., LEDUC, C. A., LOWE, G., LU, Y., MARQUES-VIDAL, P., MEISINGER, C., MENNI, C., MORRIS, A. P., MYERS, R. H., MÄNNISTÖ, S., NALLS, M. A., PATERNOSTER, L., PETERS, A., PRADHAN, A. D., RANKINEN, T., RASMUSSEN-TORVIK, L. J., RATHMANN, W., RICE, T. K., BRENT RICHARDS, J., RIDKER, P. M., SATTAR, N., SAVAGE, D. B., et al. 2016. Genome-wide meta-analysis uncovers novel loci influencing circulating leptin levels. *Nat Commun,* 7**,** 10494.

KIM, J. H., BACHMANN, R. A. & CHEN, J. 2009. Interleukin-6 and insulin resistance. *Vitam Horm,* 80**,** 613-33.

KIM, S. H. & REAVEN, G. 2013. Sex differences in insulin resistance and cardiovascular disease risk. *J Clin Endocrinol Metab,* 98**,** E1716-21.

KINGSBURY, M., WEEKS, M., MACKINNON, N., EVANS, J., MAHEDY, L., DYKXHOORN, J. & COLMAN, I. 2016. Stressful Life Events During Pregnancy and Offspring Depression: Evidence From a Prospective Cohort Study. *J Am Acad Child Adolesc Psychiatry,* 55**,** 709-716 e2.

KIRKBRIDE, J. B., FEARON, P., MORGAN, C., DAZZAN, P., MORGAN, K., MURRAY, R. M. & JONES, P. B. 2007. Neighbourhood variation in the incidence of psychotic disorders in Southeast London. *Soc Psychiatry Psychiatr Epidemiol,* 42**,** 438-45.

KIRKPATRICK, B., BUCHANAN, R. W., ROSS, D. E. & CARPENTER, W. T., JR. 2001. A separate disease within the syndrome of schizophrenia. *Arch Gen Psychiatry,* 58**,** 165-71.

KIVIMAKI, M., BATTY, G. D., PENTTI, J., SHIPLEY, M. J., SIPILA, P. N., NYBERG, S. T., SUOMINEN, S. B., OKSANEN, T., STENHOLM, S., VIRTANEN, M., MARMOT, M. G., SINGH-MANOUX, A., BRUNNER, E. J., LINDBOHM, J. V., FERRIE, J. E. & VAHTERA, J. 2020. Association between socioeconomic status and the development of mental and physical health conditions in adulthood: a multi-cohort study. *Lancet Public Health,* 5**,** e140-e149.

KNUESEL, I., CHICHA, L., BRITSCHGI, M., SCHOBEL, S. A., BODMER, M., HELLINGS, J. A., TOOVEY, S. & PRINSSEN, E. P. 2014. Maternal immune activation and abnormal brain development across CNS disorders. *Nat Rev Neurol,* 10**,** 643-60.

KNUTSON, K. L., VAN CAUTER, E., RATHOUZ, P. J., YAN, L. L., HULLEY, S. B., LIU, K. & LAUDERDALE, D. S. 2009. Association between sleep and blood pressure in midlife: the CARDIA sleep study. *Arch Intern Med,* 169**,** 1055-61.

KOHEN, D. 2004. Diabetes mellitus and schizophrenia: historical perspective. *Br J Psychiatry Suppl,* 47**,** S64-6.

KOOY, F. H. 1919. Hyperglycemia in mental disorders. *Brain,* 42**,** 214-289.

KOWARIK, A., TEMPL, M. 2016. Imputation with the R Package VIM. *Journal of Statistical Software,* 74.

KRINGLEN, E. & CRAMER, G. 1989. Offspring of monozygotic twins discordant for schizophrenia. *Arch Gen Psychiatry,* 46**,** 873-7.

KROEZE, W. K., HUFEISEN, S. J., POPADAK, B. A., RENOCK, S. M., STEINBERG, S., ERNSBERGER, P., JAYATHILAKE, K., MELTZER, H. Y. & ROTH, B. L. 2003. H1-histamine receptor affinity predicts short-term weight gain for typical and atypical antipsychotic drugs. *Neuropsychopharmacology,* 28**,** 519-26.

KRYSTAL, A. D., GOFORTH, H. W. & ROTH, T. 2008. Effects of antipsychotic medications on sleep in schizophrenia. *Int Clin Psychopharmacol,* 23**,** 150-60.

KUK, J. L. & ARDERN, C. I. 2010. Age and sex differences in the clustering of metabolic syndrome factors: association with mortality risk. *Diabetes Care,* 33**,** 2457-61.

KUMPHUNE, S., CHATTIPAKORN, S. & CHATTIPAKORN, N. 2013. Roles of p38-MAPK in insulin resistant heart: evidence from bench to future bedside application. *Curr Pharm Des,* 19**,** 5742-54.

KYOSSEVA, S. V., ELBEIN, A. D., GRIFFIN, W. S., MRAK, R. E., LYON, M. & KARSON, C. N. 1999. Mitogen-activated protein kinases in schizophrenia. *Biol Psychiatry,* 46**,** 689-96.

L'ITALIEN, G., FORD, I., NORRIE, J., LAPUERTA, P., EHRETH, J., JACKSON, J. & SHEPHERD, J. 2000. The cardiovascular event reduction tool (CERT)--a simplified cardiac risk prediction model developed from the West of Scotland Coronary Prevention Study (WOSCOPS). *Am J Cardiol,* 85**,** 720-4.

LABORIT, H., HUGUENARD, P. & ALLUAUME, R. 1952. [A new vegetative stabilizer; 4560 R.P..]. *Presse Med,* 60**,** 206-8.

LAGOU, V., MAGI, R., HOTTENGA, JJ., 2019. Sex-dimorphic genetic effects and novel loci for fasting glucose and insulin variability. *accepted in principle*.

LAMERS, F., MILANESCHI, Y., DE JONGE, P., GILTAY, E. J. & PENNINX, B. 2018. Metabolic and inflammatory markers: associations with individual depressive symptoms. *Psychol Med,* 48**,** 1102-1110.

LAMICHHANE, S., DICKENS, A. M., SEN, P., LAURIKAINEN, H., BORGAN, F., SUVISAARI, J., HYOTYLAINEN, T., HOWES, O., HIETALA, J. & ORESIC, M. 2021. Association Between Circulating Lipids and Future Weight Gain in Individuals With an At-Risk Mental State and in First-Episode Psychosis. *Schizophr Bull,* 47**,** 160-169.

LANDER, E. S., LINTON, L. M., BIRREN, B., NUSBAUM, C., ZODY, M. C., BALDWIN, J., DEVON, K., DEWAR, K., DOYLE, M., FITZHUGH, W., FUNKE, R., GAGE, D., HARRIS, K., HEAFORD, A., HOWLAND, J., KANN, L., LEHOCZKY, J., LEVINE, R., MCEWAN, P., MCKERNAN, K., MELDRIM, J., MESIROV, J. P., MIRANDA, C., MORRIS, W., NAYLOR, J., RAYMOND, C., ROSETTI, M., SANTOS, R., SHERIDAN, A., SOUGNEZ, C., STANGE-THOMANN, Y., STOJANOVIC, N., SUBRAMANIAN, A., WYMAN, D., ROGERS, J., SULSTON, J., AINSCOUGH, R., BECK, S., BENTLEY, D., BURTON, J., CLEE, C., CARTER, N., COULSON, A., DEADMAN, R., DELOUKAS, P., DUNHAM, A., DUNHAM, I., DURBIN, R., FRENCH, L., GRAFHAM, D., GREGORY, S., HUBBARD, T., HUMPHRAY, S., HUNT, A., JONES, M., LLOYD, C., MCMURRAY, A., MATTHEWS, L., MERCER, S., MILNE, S., MULLIKIN, J. C., MUNGALL, A., PLUMB, R., ROSS, M., SHOWNKEEN, R., SIMS, S., WATERSTON, R. H., WILSON, R. K., HILLIER, L. W., MCPHERSON, J. D., MARRA, M. A., MARDIS, E. R., FULTON, L. A., CHINWALLA, A. T., PEPIN, K. H., GISH, W. R., CHISSOE, S. L., WENDL, M. C., DELEHAUNTY, K. D., MINER, T. L., DELEHAUNTY, A., KRAMER, J. B., COOK, L. L., FULTON, R. S., JOHNSON, D. L., MINX, P. J., CLIFTON, S. W., HAWKINS, T., BRANSCOMB, E., PREDKI, P., RICHARDSON, P., WENNING, S., SLEZAK, T., DOGGETT, N., CHENG, J. F., OLSEN, A., LUCAS, S., ELKIN, C., UBERBACHER, E., FRAZIER, M., et al. 2001. Initial sequencing and analysis of the human genome. *Nature,* 409**,** 860-921.

LAPIZ, M. D., FULFORD, A., MUCHIMAPURA, S., MASON, R., PARKER, T. & MARSDEN, C. A. 2003. Influence of postweaning social isolation in the rat on brain development, conditioned behavior, and neurotransmission. *Neurosci Behav Physiol,* 33**,** 13-29.

LARISCY, J. T. 2019. Smoking-attributable mortality by cause of death in the United States: An indirect approach. *SSM Popul Health,* 7**,** 100349.

LAURSEN, T. M., PLANA-RIPOLL, O., ANDERSEN, P. K., MCGRATH, J. J., TOENDER, A., NORDENTOFT, M., CANUDAS-ROMO, V. & ERLANGSEN, A. 2019. Cause-specific

life years lost among persons diagnosed with schizophrenia: Is it getting better or worse? *Schizophr Res,* 206**,** 284-290.

LAWRENCE, D. & KISELY, S. 2010. Inequalities in healthcare provision for people with severe mental illness. *J Psychopharmacol,* 24**,** 61-8.

LAWRENSON, K., LI, Q., KAR, S., SEO, J. H., TYRER, J., SPINDLER, T. J., LEE, J., CHEN, Y., KARST, A., DRAPKIN, R., ABEN, K. K., ANTON-CULVER, H., ANTONENKOVA, N., AUSTRALIAN OVARIAN CANCER STUDY, G., BAKER, H., BANDERA, E. V., BEAN, Y., BECKMANN, M. W., BERCHUCK, A., BISOGNA, M., BJORGE, L., BOGDANOVA, N., BRINTON, L. A., BROOKS-WILSON, A., BRUINSMA, F., BUTZOW, R., CAMPBELL, I. G., CARTY, K., CHANG-CLAUDE, J., CHENEVIX-TRENCH, G., CHEN, A., CHEN, Z., COOK, L. S., CRAMER, D. W., CUNNINGHAM, J. M., CYBULSKI, C., DANSONKA-MIESZKOWSKA, A., DENNIS, J., DICKS, E., DOHERTY, J. A., DORK, T., DU BOIS, A., DURST, M., ECCLES, D., EASTON, D. T., EDWARDS, R. P., EILBER, U., EKICI, A. B., FASCHING, P. A., FRIDLEY, B. L., GAO, Y. T., GENTRY-MAHARAJ, A., GILES, G. G., GLASSPOOL, R., GOODE, E. L., GOODMAN, M. T., GROWNWALD, J., HARRINGTON, P., HARTER, P., HASMAD, H. N., HEIN, A., HEITZ, F., HILDEBRANDT, M. A., HILLEMANNS, P., HOGDALL, E., HOGDALL, C., HOSONO, S., IVERSEN, E. S., JAKUBOWSKA, A., JAMES, P., JENSEN, A., JI, B. T., KARLAN, B. Y., KRUGER KJAER, S., KELEMEN, L. E., KELLAR, M., KELLEY, J. L., KIEMENEY, L. A., KRAKSTAD, C., KUPRYJANCZYK, J., LAMBRECHTS, D., LAMBRECHTS, S., LE, N. D., LEE, A. W., LELE, S., LEMINEN, A., LESTER, J., LEVINE, D. A., LIANG, D., LISSOWSKA, J., LU, K., LUBINSKI, J., LUNDVALL, L., MASSUGER, L. F., MATSUO, K., MCGUIRE, V., MCLAUGHLIN, J. R., NEVANLINNA, H., MCNEISH, I., MENON, U., et al. 2015. Cis-eQTL analysis and functional validation of candidate susceptibility genes for high-grade serous ovarian cancer. *Nat Commun,* 6**,** 8234.

LAWS, A. & REAVEN, G. M. 1992. Evidence for an independent relationship between insulin resistance and fasting plasma HDL-cholesterol, triglyceride and insulin concentrations. *J Intern Med,* 231**,** 25-30.

LEAR, S. A. & GASEVIC, D. 2019. Ethnicity and Metabolic Syndrome: Implications for Assessment, Management and Prevention. *Nutrients,* 12.

LEE, H. J., HUBER, J.,. Multiple imputation with large proportions of missing data: How much is too much? United Kingdom Stata Users' Group Meetings 2011, 2011. Stata Users Group.

LEE, I. M., SHIROMA, E. J., LOBELO, F., PUSKA, P., BLAIR, S. N., KATZMARZYK, P. T. & LANCET PHYSICAL ACTIVITY SERIES WORKING, G. 2012. Effect of physical

inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet,* 380**,** 219-29.

LEE, K. S. & VAILLANCOURT, T. 2018. Longitudinal Associations Among Bullying by Peers, Disordered Eating Behavior, and Symptoms of Depression During Adolescence. *JAMA Psychiatry,* 75**,** 605-612.

LEE, S. & DONG, H. H. 2017. FoxO integration of insulin signaling with glucose and lipid metabolism. *J Endocrinol,* 233**,** R67-R79.

LEE, S. H., RIPKE, S., NEALE, B. M., FARAONE, S. V., PURCELL, S. M., PERLIS, R. H., MOWRY, B. J., THAPAR, A., GODDARD, M. E., WITTE, J. S., ABSHER, D., AGARTZ, I., AKIL, H., AMIN, F., ANDREASSEN, O. A., ANJORIN, A., ANNEY, R., ANTTILA, V., ARKING, D. E., ASHERSON, P., AZEVEDO, M. H., BACKLUND, L., BADNER, J. A., BAILEY, A. J., BANASCHEWSKI, T., BARCHAS, J. D., BARNES, M. R., BARRETT, T. B., BASS, N., BATTAGLIA, A., BAUER, M., BAYES, M., BELLIVIER, F., BERGEN, S. E., BERRETTINI, W., BETANCUR, C., BETTECKEN, T., BIEDERMAN, J., BINDER, E. B., BLACK, D. W., BLACKWOOD, D. H., BLOSS, C. S., BOEHNKE, M., BOOMSMA, D. I., BREEN, G., BREUER, R., BRUGGEMAN, R., CORMICAN, P., BUCCOLA, N. G., BUITELAAR, J. K., BUNNEY, W. E., BUXBAUM, J. D., BYERLEY, W. F., BYRNE, E. M., CAESAR, S., CAHN, W., CANTOR, R. M., CASAS, M., CHAKRAVARTI, A., CHAMBERT, K., CHOUDHURY, K., CICHON, S., CLONINGER, C. R., COLLIER, D. A., COOK, E. H., COON, H., CORMAND, B., CORVIN, A., CORYELL, W. H., CRAIG, D. W., CRAIG, I. W., CROSBIE, J., CUCCARO, M. L., CURTIS, D., CZAMARA, D., DATTA, S., DAWSON, G., DAY, R., DE GEUS, E. J., DEGENHARDT, F., DJUROVIC, S., DONOHOE, G. J., DOYLE, A. E., DUAN, J., DUDBRIDGE, F., DUKETIS, E., EBSTEIN, R. P., EDENBERG, H. J., ELIA, J., ENNIS, S., ETAIN, B., FANOUS, A., FARMER, A. E., FERRIER, I. N., FLICKINGER, M., FOMBONNE, E., FOROUD, T., FRANK, J., FRANKE, B., FRASER, C., et al. 2013. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet,* 45**,** 984-94.

LEIBOWITZ, S. F., ROOSSIN, P., ROSENN, M. 1984. Chronic norepinephrine injection into the hypothalamic paraventricular nucleus produces hyperphagia and increased body weight in the rat. *Pharmacology Biochemistry and Behavior,* 21**,** 801-808.

LENCZ, T., MORGAN, T. V., ATHANASIOU, M., DAIN, B., REED, C. R., KANE, J. M., KUCHERLAPATI, R. & MALHOTRA, A. K. 2007. Converging evidence for a pseudoautosomal cytokine receptor gene locus in schizophrenia. *Mol Psychiatry,* 12**,** 572-80.

LENG, Y., CAPPUCCIO, F. P., WAINWRIGHT, N. W., SURTEES, P. G., LUBEN, R., BRAYNE, C. & KHAW, K. T. 2015. Sleep duration and risk of fatal and nonfatal stroke: a prospective study and meta-analysis. *Neurology,* 84**,** 1072-9.

LEUCHT, S., CIPRIANI, A., SPINELI, L., MAVRIDIS, D., OREY, D., RICHTER, F., SAMARA, M., BARBUI, C., ENGEL, R. R., GEDDES, J. R., KISSLING, W., STAPF, M. P., LASSIG, B., SALANTI, G. & DAVIS, J. M. 2013. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet,* 382**,** 951-62.

LEVER, J., KRZYWINSKI, M., ALTMAN, N. 2016. Model Selection and Overfitting. *Nature Methods,* 13**,** 703-704.

LEVY, J. C., MATTHEWS, D. R. & HERMANS, M. P. 1998. Correct homeostasis model assessment (HOMA) evaluation uses the computer program. *Diabetes Care,* 21**,** 2191-2.

LEWIS, G., IOANNIDIS, K., VAN HARMELEN, A. L., NEUFELD, S., STOCHL, J., LEWIS, G., JONES, P. B. & GOODYER, I. 2018. The association between pubertal status and depressive symptoms and diagnoses in adolescent females: A population-based cohort study. *PLoS One,* 13**,** e0198804.

LEWIS, G., PELOSI, A. J., ARAYA, R. & DUNN, G. 1992. Measuring psychiatric disorder in the community: a standardized assessment for use by lay interviewers. *Psychol Med,* 22**,** 465-86.

LI, J. & MELTZER, H. Y. 2014. A genetic locus in 7p12.2 associated with treatment resistant schizophrenia. *Schizophr Res,* 159**,** 333-9.

LI, W., LIU, Q., DENG, X., CHEN, Y., LIU, S. & STORY, M. 2017a. Association between Obesity and Puberty Timing: A Systematic Review and Meta-Analysis. *Int J Environ Res Public Health,* 14.

LI, Z., CHEN, J., YU, H., HE, L., XU, Y., ZHANG, D., YI, Q., LI, C., LI, X., SHEN, J., SONG, Z., JI, W., WANG, M., ZHOU, J., CHEN, B., LIU, Y., WANG, J., WANG, P., YANG, P., WANG, Q., FENG, G., LIU, B., SUN, W., LI, B., HE, G., LI, W., WAN, C., XU, Q., LI, W., WEN, Z., LIU, K., HUANG, F., JI, J., RIPKE, S., YUE, W., SULLIVAN, P. F., O'DONOVAN, M. C. & SHI, Y. 2017b. Genome-wide association analysis identifies 30 new susceptibility loci for schizophrenia. *Nat Genet,* 49**,** 1576-1583.

LI, Z., CHEN, P., CHEN, J., XU, Y., WANG, Q., LI, X., LI, C., HE, L. & SHI, Y. 2018. Glucose and Insulin-Related Traits, Type 2 Diabetes and Risk of Schizophrenia: A Mendelian Randomization Study. *EBioMedicine,* 34**,** 182-188.

LIAO, S. G., LIN, Y., KANG, D.D., CHANDRA, D., BON, J., KAMINSKI, N., SCIURBA, F.C., TSENG, G.C. 2014. Missing value imputation in high-dimensional phenomic data: Imputable or not, and how? *BMC Bioinformatics,* 15**,** 346.

LIEBERMAN, J. A. & FIRST, M. B. 2018. Psychotic Disorders. *N Engl J Med,* 379**,** 270-280.

LIEBERMAN, J. A., KANE, J. M. & ALVIR, J. 1987. Provocative tests with psychostimulant drugs in schizophrenia. *Psychopharmacology (Berl),* 91**,** 415-33.

LIGTHART, S., DE VRIES, P. S., UITTERLINDEN, A. G., HOFMAN, A., GROUP, C. I. W., FRANCO, O. H., CHASMAN, D. I. & DEHGHAN, A. 2015. Pleiotropy among common genetic loci identified for cardiometabolic disorders and C-reactive protein. *PLoS One,* 10**,** e0118859.

LIGTHART, S., VAEZ, A., VÕSA, U., STATHOPOULOU, M. G., DE VRIES, P. S., PRINS, B. P., VAN DER MOST, P. J., TANAKA, T., NADERI, E., ROSE, L. M., WU, Y., KARLSSON, R., BARBALIC, M., LIN, H., POOL, R., ZHU, G., MACÉ, A., SIDORE, C., TROMPET, S., MANGINO, M., SABATER-LLEAL, M., KEMP, J. P., ABBASI, A., KACPROWSKI, T., VERWEIJ, N., SMITH, A. V., HUANG, T., MARZI, C., FEITOSA, M. F., LOHMAN, K. K., KLEBER, M. E., MILANESCHI, Y., MUELLER, C., HUQ, M., VLACHOPOULOU, E., LYYTIKÄINEN, L. P., OLDMEADOW, C., DEELEN, J., PEROLA, M., ZHAO, J. H., FEENSTRA, B., AMINI, M., LAHTI, J., SCHRAUT, K. E., FORNAGE, M., SUKTITIPAT, B., CHEN, W. M., LI, X., NUTILE, T., MALERBA, G., LUAN, J., BAK, T., SCHORK, N., DEL GRECO M, F., THIERING, E., MAHAJAN, A., MARIONI, R. E., MIHAILOV, E., ERIKSSON, J., OZEL, A. B., ZHANG, W., NETHANDER, M., CHENG, Y. C., ASLIBEKYAN, S., ANG, W., GANDIN, I., YENGO, L., PORTAS, L., KOOPERBERG, C., HOFER, E., RAJAN, K. B., SCHURMANN, C., DEN HOLLANDER, W., AHLUWALIA, T. S., ZHAO, J., DRAISMA, H. H. M., FORD, I., TIMPSON, N., TEUMER, A., HUANG, H., WAHL, S., LIU, Y., HUANG, J., UH, H. W., GELLER, F., JOSHI, P. K., YANEK, L. R., TRABETTI, E., LEHNE, B., VOZZI, D., VERBANCK, M., BIINO, G., SABA, Y., MEULENBELT, I., O'CONNELL, J. R., LAAKSO, M., GIULIANINI, F., MAGNUSSON, P. K. E., BALLANTYNE, C. M., HOTTENGA, J. J., et al. 2018. Genome Analyses of >200,000 Individuals Identify 58 Loci for Chronic Inflammation and Highlight Pathways that Link Inflammation and Complex Disorders. *Am J Hum Genet,* 103**,** 691-706.

LIM, S., WYKER, B., BARTLEY, K. & EISENHOWER, D. 2015. Measurement error of self-reported physical activity levels in New York City: assessment and correction. *Am J Epidemiol,* 181**,** 648-55.

LIN, B. D., ALKEMA, A., PETERS, T., ZINKSTOK, J., LIBUDA, L., HEBEBRAND, J., ANTEL, J., HINNEY, A., CAHN, W., ADAN, R. & LUYKX, J. J. 2019. Assessing causal links between metabolic traits, inflammation and schizophrenia: a univariable and multivariable, bidirectional Mendelian-randomization study. *Int J Epidemiol,* 48**,** 1505-1514.

LIN, P. I. & SHULDINER, A. R. 2010. Rethinking the genetic basis for comorbidity of schizophrenia and type 2 diabetes. *Schizophr Res,* 123**,** 234-43.

LITTLE, T. D., JORGENSEN, T. D., LANG, K. M. & MOORE, E. W. 2014. On the joys of missing data. *J Pediatr Psychol,* 39**,** 151-62.

LIU, D. J., PELOSO, G. M., YU, H., BUTTERWORTH, A. S., WANG, X., MAHAJAN, A., SALEHEEN, D., EMDIN, C., ALAM, D., ALVES, A. C., AMOUYEL, P., DI ANGELANTONIO, E., ARVEILER, D., ASSIMES, T. L., AUER, P. L., BABER, U., BALLANTYNE, C. M., BANG, L. E., BENN, M., BIS, J. C., BOEHNKE, M., BOERWINKLE, E., BORK-JENSEN, J., BOTTINGER, E. P., BRANDSLUND, I., BROWN, M., BUSONERO, F., CAULFIELD, M. J., CHAMBERS, J. C., CHASMAN, D. I., CHEN, Y. E., CHEN, Y. I., CHOWDHURY, R., CHRISTENSEN, C., CHU, A. Y., CONNELL, J. M., CUCCA, F., CUPPLES, L. A., DAMRAUER, S. M., DAVIES, G., DEARY, I. J., DEDOUSSIS, G., DENNY, J. C., DOMINICZAK, A., DUBE, M. P., EBELING, T., EIRIKSDOTTIR, G., ESKO, T., FARMAKI, A. E., FEITOSA, M. F., FERRARIO, M., FERRIERES, J., FORD, I., FORNAGE, M., FRANKS, P. W., FRAYLING, T. M., FRIKKE-SCHMIDT, R., FRITSCHE, L. G., FROSSARD, P., FUSTER, V., GANESH, S. K., GAO, W., GARCIA, M. E., GIEGER, C., GIULIANINI, F., GOODARZI, M. O., GRALLERT, H., GRARUP, N., GROOP, L., GROVE, M. L., GUDNASON, V., HANSEN, T., HARRIS, T. B., HAYWARD, C., HIRSCHHORN, J. N., HOLMEN, O. L., HUFFMAN, J., HUO, Y., HVEEM, K., JABEEN, S., JACKSON, A. U., JAKOBSDOTTIR, J., JARVELIN, M. R., JENSEN, G. B., JORGENSEN, M. E., JUKEMA, J. W., JUSTESEN, J. M., KAMSTRUP, P. R., KANONI, S., KARPE, F., KEE, F., KHERA, A. V., KLARIN, D., KOISTINEN, H. A., KOONER, J. S., KOOPERBERG, C., KUULASMAA, K., KUUSISTO, J., LAAKSO, M., LAKKA, T., et al. 2017. Exome-wide association study of plasma lipids in >300,000 individuals. *Nat Genet,* 49**,** 1758-1766.

LIU, Y., LI, Z., ZHANG, M., DENG, Y., YI, Z. & SHI, T. 2013. Exploring the pathogenetic association between schizophrenia and type 2 diabetes mellitus diseases based on pathway analysis. *BMC Med Genomics,* 6 Suppl 1**,** S17.

LLOYD-JONES, D. M., HUFFMAN, M. D., KARMALI, K. N., SANGHAVI, D. M., WRIGHT, J. S., PELSER, C., GULATI, M., MASOUDI, F. A. & GOFF, D. C., JR. 2017. Estimating Longitudinal Risks and Benefits From Cardiovascular Preventive Therapies Among

Medicare Patients: The Million Hearts Longitudinal ASCVD Risk Assessment Tool: A Special Report From the American Heart Association and American College of Cardiology. *Circulation,* 135**,** e793-e813.

LOCKE, A. E., KAHALI, B., BERNDT, S. I., JUSTICE, A. E., PERS, T. H., DAY, F. R., POWELL, C., VEDANTAM, S., BUCHKOVICH, M. L., YANG, J., CROTEAU-CHONKA, D. C., ESKO, T., FALL, T., FERREIRA, T., GUSTAFSSON, S., KUTALIK, Z., LUAN, J., MÄGI, R., RANDALL, J. C., WINKLER, T. W., WOOD, A. R., WORKALEMAHU, T., FAUL, J. D., SMITH, J. A., ZHAO, J. H., ZHAO, W., CHEN, J., FEHRMANN, R., HEDMAN, Å., KARJALAINEN, J., SCHMIDT, E. M., ABSHER, D., AMIN, N., ANDERSON, D., BEEKMAN, M., BOLTON, J. L., BRAGG-GRESHAM, J. L., BUYSKE, S., DEMIRKAN, A., DENG, G., EHRET, G. B., FEENSTRA, B., FEITOSA, M. F., FISCHER, K., GOEL, A., GONG, J., JACKSON, A. U., KANONI, S., KLEBER, M. E., KRISTIANSSON, K., LIM, U., LOTAY, V., MANGINO, M., LEACH, I. M., MEDINA-GOMEZ, C., MEDLAND, S. E., NALLS, M. A., PALMER, C. D., PASKO, D., PECHLIVANIS, S., PETERS, M. J., PROKOPENKO, I., SHUNGIN, D., STANČÁKOVÁ, A., STRAWBRIDGE, R. J., SUNG, Y. J., TANAKA, T., TEUMER, A., TROMPET, S., VAN DER LAAN, S. W., VAN SETTEN, J., VAN VLIET-OSTAPTCHOUK, J. V., WANG, Z., YENGO, L., ZHANG, W., ISAACS, A., ALBRECHT, E., ÄRNLÖV, J., ARSCOTT, G. M., ATTWOOD, A. P., BANDINELLI, S., BARRETT, A., BAS, I. N., BELLIS, C., BENNETT, A. J., BERNE, C., BLAGIEVA, R., BLÜHER, M., BÖHRINGER, S., BONNYCASTLE, L. L., BÖTTCHER, Y., BOYD, H. A., BRUINENBERG, M., CASPERSEN, I. H., CHEN, Y. I., CLARKE, R., DAW, E. W., DE CRAEN, A. J. M., DELGADO, G., DIMITRIOU, M., et al. 2015. Genetic studies of body mass index yield new insights for obesity biology. *Nature,* 518**,** 197-206.

LOEBEL, A. D., LIEBERMAN, J. A., ALVIR, J. M., MAYERHOFF, D. I., GEISLER, S. H. & SZYMANSKI, S. R. 1992. Duration of psychosis and outcome in first-episode schizophrenia. *Am J Psychiatry,* 149**,** 1183-8.

LORENZ, W. F. 1922. Sugar tolerance in dementia praecoxand other mental states. *Archives of Neurology and Psychiatry,* 8**,** 184-196.

LOTTA, L. A., GULATI, P., DAY, F. R., PAYNE, F., ONGEN, H., VAN DE BUNT, M., GAULTON, K. J., EICHER, J. D., SHARP, S. J., LUAN, J., DE LUCIA ROLFE, E., STEWART, I. D., WHEELER, E., WILLEMS, S. M., ADAMS, C., YAGHOOTKAR, H., CONSORTIUM, E. P.-I., CAMBRIDGE, F. C., FOROUHI, N. G., KHAW, K. T., JOHNSON, A. D., SEMPLE, R. K., FRAYLING, T., PERRY, J. R., DERMITZAKIS, E., MCCARTHY, M. I., BARROSO, I., WAREHAM, N. J., SAVAGE, D. B.,

LANGENBERG, C., O'RAHILLY, S. & SCOTT, R. A. 2017a. Integrative genomic analysis implicates limited peripheral adipose storage capacity in the pathogenesis of human insulin resistance. *Nat Genet,* 49**,** 17-26.

LOTTA, L. A., GULATI, P., DAY, F. R., PAYNE, F., ONGEN, H., VAN DE BUNT, M., GAULTON, K. J., EICHER, J. D., SHARP, S. J., LUAN, J., DE LUCIA ROLFE, E., STEWART, I. D., WHEELER, E., WILLEMS, S. M., ADAMS, C., YAGHOOTKAR, H., FOROUHI, N. G., KHAW, K. T., JOHNSON, A. D., SEMPLE, R. K., FRAYLING, T., PERRY, J. R., DERMITZAKIS, E., MCCARTHY, M. I., BARROSO, I., WAREHAM, N. J., SAVAGE, D. B., LANGENBERG, C., O'RAHILLY, S., SCOTT, R. A., CONSORTIUM, E.-I. & CONSORTIUM, C. F. 2017b. Integrative genomic analysis implicates limited peripheral adipose storage capacity in the pathogenesis of human insulin resistance. *Nat Genet,* 49**,** 17-26.

LOU, M., LUO, P., TANG, R., PENG, Y., YU, S., HUANG, W. & HE, L. 2015. Relationship between neutrophil-lymphocyte ratio and insulin resistance in newly diagnosed type 2 diabetes mellitus patients. *BMC Endocr Disord,* 15**,** 9.

LU, Q., LI, B., OU, D., ERLENDSDOTTIR, M., POWLES, R. L., JIANG, T., HU, Y., CHANG, D., JIN, C., DAI, W., HE, Q., LIU, Z., MUKHERJEE, S., CRANE, P. K. & ZHAO, H. 2017. A Powerful Approach to Estimating Annotation-Stratified Genetic Covariance via GWAS Summary Statistics. *Am J Hum Genet,* 101**,** 939-964.

LUCATCH, A. M., LOWE, D. J. E., CLARK, R. C., KOZAK, K. & GEORGE, T. P. 2018. Neurobiological Determinants of Tobacco Smoking in Schizophrenia. *Front Psychiatry,* 9**,** 672.

LUM, A., SKELTON, E., WYNNE, O. & BONEVSKI, B. 2018. A Systematic Review of Psychosocial Barriers and Facilitators to Smoking Cessation in People Living With Schizophrenia. *Front Psychiatry,* 9**,** 565.

LUNDELL, L. S., MASSART, J., ALTINTAS, A., KROOK, A. & ZIERATH, J. R. 2019. Regulation of glucose uptake and inflammation markers by FOXO1 and FOXO3 in skeletal muscle. *Mol Metab,* 20**,** 79-88.

LUO, Q., CHEN, Q., WANG, W., DESRIVIERES, S., QUINLAN, E. B., JIA, T., MACARE, C., ROBERT, G. H., CUI, J., GUEDJ, M., PALANIYAPPAN, L., KHERIF, F., BANASCHEWSKI, T., BOKDE, A. L. W., BUCHEL, C., FLOR, H., FROUIN, V., GARAVAN, H., GOWLAND, P., HEINZ, A., ITTERMANN, B., MARTINOT, J. L., ARTIGES, E., PAILLERE-MARTINOT, M. L., NEES, F., ORFANOS, D. P., POUSTKA, L., FROHNER, J. H., SMOLKA, M. N., WALTER, H., WHELAN, R., CALLICOTT, J. H., MATTAY, V. S., PAUSOVA, Z., DARTIGUES, J. F., TZOURIO, C., CRIVELLO, F.,

BERMAN, K. F., LI, F., PAUS, T., WEINBERGER, D. R., MURRAY, R. M., SCHUMANN, G., FENG, J. & CONSORTIUM, I. 2019. Association of a Schizophrenia-Risk Nonsynonymous Variant With Putamen Volume in Adolescents: A Voxelwise and Genome-Wide Association Study. *JAMA Psychiatry,* 76**,** 435-445.

LUPPINO, F. S., DE WIT, L. M., BOUVY, P. F., STIJNEN, T., CUIJPERS, P., PENNINX, B. W. & ZITMAN, F. G. 2010. Overweight, obesity, and depression: a systematic review and meta-analysis of longitudinal studies. *Arch Gen Psychiatry,* 67**,** 220-9.

LV, W. Q., ZHANG, X., ZHANG, Q., HE, J. Y., LIU, H. M., XIA, X., FAN, K., ZHAO, Q., SHI, X. Z., ZHANG, W. D., SUN, C. Q. & DENG, H. W. 2017. Novel common variants associated with body mass index and coronary artery disease detected using a pleiotropic cFDR method. *J Mol Cell Cardiol,* 112**,** 1-7.

MACHIELA, M. J. & CHANOCK, S. J. 2015. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics,* 31**,** 3555-7.

MACKAY, A. V., IVERSEN, L. L., ROSSOR, M., SPOKES, E., BIRD, E., ARREGUI, A., CREESE, I. & SYNDER, S. H. 1982. Increased brain dopamine and dopamine receptors in schizophrenia. *Arch Gen Psychiatry,* 39**,** 991-7.

MADLEY-DOWD, P., HUGHES, R., TILLING, K. & HERON, J. 2019. The proportion of missing data should not be used to guide decisions on multiple imputation. *J Clin Epidemiol,* 110**,** 63-73.

MADRID-GAMBIN, F., FOCKING, M., SABHERWAL, S., HEURICH, M., ENGLISH, J. A., O'GORMAN, A., SUVITAIVAL, T., AHONEN, L., CANNON, M., LEWIS, G., MATTILA, I., SCAIFE, C., MADDEN, S., HYOTYLAINEN, T., ORESIC, M., ZAMMIT, S., CAGNEY, G., COTTER, D. R. & BRENNAN, L. 2019. Integrated Lipidomics and Proteomics Point to Early Blood-Based Changes in Childhood Preceding Later Development of Psychotic Experiences: Evidence From the Avon Longitudinal Study of Parents and Children. *Biol Psychiatry,* 86**,** 25-34.

MAEKAWA, F., FUJIWARA, K., TORIYA, M., MAEJIMA, Y., NISHIO, T., TOYODA, Y., NOHARA, K., YASHIRO, T. & YADA, T. 2013. Brain-derived neurotrophic factor in VMH as the causal factor for and therapeutic tool to treat visceral adiposity and hyperleptinemia in type 2 diabetic Goto-Kakizaki rats. *Front Synaptic Neurosci,* 5**,** 7.

MAHAJAN, A., GO, M. J., ZHANG, W., BELOW, J. E., GAULTON, K. J., FERREIRA, T., HORIKOSHI, M., JOHNSON, A. D., NG, M. C., PROKOPENKO, I., SALEHEEN, D., WANG, X., ZEGGINI, E., ABECASIS, G. R., ADAIR, L. S., ALMGREN, P., ATALAY, M., AUNG, T., BALDASSARRE, D., BALKAU, B., BAO, Y., BARNETT, A. H.,

BARROSO, I., BASIT, A., BEEN, L. F., BEILBY, J., BELL, G. I., BENEDIKTSSON, R., BERGMAN, R. N., BOEHM, B. O., BOERWINKLE, E., BONNYCASTLE, L. L., BURTT, N., CAI, Q., CAMPBELL, H., CAREY, J., CAUCHI, S., CAULFIELD, M., CHAN, J. C., CHANG, L. C., CHANG, T. J., CHANG, Y. C., CHARPENTIER, G., CHEN, C. H., CHEN, H., CHEN, Y. T., CHIA, K. S., CHIDAMBARAM, M., CHINES, P. S., CHO, N. H., CHO, Y. M., CHUANG, L. M., COLLINS, F. S., CORNELIS, M. C., COUPER, D. J., CRENSHAW, A. T., VAN DAM, R. M., DANESH, J., DAS, D., DE FAIRE, U., DEDOUSSIS, G., DELOUKAS, P., DIMAS, A. S., DINA, C., DONEY, A. S., DONNELLY, P. J., DORKHAN, M., VAN DUIJN, C., DUPUIS, J., EDKINS, S., ELLIOTT, P., EMILSSON, V., ERBEL, R., ERIKSSON, J. G., ESCOBEDO, J., ESKO, T., EURY, E., FLOREZ, J. C., FONTANILLAS, P., FOROUHI, N. G., FORSEN, T., FOX, C., FRASER, R. M., FRAYLING, T. M., FROGUEL, P., FROSSARD, P., GAO, Y., GERTOW, K., GIEGER, C., GIGANTE, B., GRALLERT, H., GRANT, G. B., GRROP, L. C., GROVES, C. J., GRUNDBERG, E., GUIDUCCI, C., HAMSTEN, A., HAN, B. G., HARA, K., HASSANALI, N., et al. 2014. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat Genet,* 46**,** 234-44.

MAHAJAN, A., WESSEL, J., WILLEMS, S. M., ZHAO, W., ROBERTSON, N. R., CHU, A. Y., GAN, W., KITAJIMA, H., TALIUN, D., RAYNER, N. W., GUO, X., LU, Y., LI, M., JENSEN, R. A., HU, Y., HUO, S., LOHMAN, K. K., ZHANG, W., COOK, J. P., PRINS, B. P., FLANNICK, J., GRARUP, N., TRUBETSKOY, V. V., KRAVIC, J., KIM, Y. J., RYBIN, D. V., YAGHOOTKAR, H., MULLER-NURASYID, M., MEIDTNER, K., LI-GAO, R., VARGA, T. V., MARTEN, J., LI, J., SMITH, A. V., AN, P., LIGTHART, S., GUSTAFSSON, S., MALERBA, G., DEMIRKAN, A., TAJES, J. F., STEINTHORSDOTTIR, V., WUTTKE, M., LECOEUR, C., PREUSS, M., BIELAK, L. F., GRAFF, M., HIGHLAND, H. M., JUSTICE, A. E., LIU, D. J., MAROULI, E., PELOSO, G. M., WARREN, H. R., EXOME, B. P. C., CONSORTIUM, M., CONSORTIUM, G., AFAQ, S., AFZAL, S., AHLQVIST, E., ALMGREN, P., AMIN, N., BANG, L. B., BERTONI, A. G., BOMBIERI, C., BORK-JENSEN, J., BRANDSLUND, I., BRODY, J. A., BURTT, N. P., CANOUIL, M., CHEN, Y. I., CHO, Y. S., CHRISTENSEN, C., EASTWOOD, S. V., ECKARDT, K. U., FISCHER, K., GAMBARO, G., GIEDRAITIS, V., GROVE, M. L., DE HAAN, H. G., HACKINGER, S., HAI, Y., HAN, S., TYBJAERG-HANSEN, A., HIVERT, M. F., ISOMAA, B., JAGER, S., JORGENSEN, M. E., JORGENSEN, T., KARAJAMAKI, A., KIM, B. J., KIM, S. S., KOISTINEN, H. A., KOVACS, P., KRIEBEL, J., KRONENBERG, F., LALL, K., LANGE, L. A., LEE, J. J.,

LEHNE, B., LI, H., LIN, K. H., et al. 2018. Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes. *Nat Genet,* 50**,** 559-571.

MALASPINA, D., CORCORAN, C., KLEINHAUS, K. R., PERRIN, M. C., FENNIG, S., NAHON, D., FRIEDLANDER, Y. & HARLAP, S. 2008. Acute maternal stress in pregnancy and schizophrenia in offspring: a cohort prospective study. *BMC Psychiatry,* 8**,** 71.

MANNING, A. K., HIVERT, M. F., SCOTT, R. A., GRIMSBY, J. L., BOUATIA-NAJI, N., CHEN, H., RYBIN, D., LIU, C. T., BIELAK, L. F., PROKOPENKO, I., AMIN, N., BARNES, D., CADBY, G., HOTTENGA, J. J., INGELSSON, E., JACKSON, A. U., JOHNSON, T., KANONI, S., LADENVALL, C., LAGOU, V., LAHTI, J., LECOEUR, C., LIU, Y., MARTINEZ-LARRAD, M. T., MONTASSER, M. E., NAVARRO, P., PERRY, J. R., RASMUSSEN-TORVIK, L. J., SALO, P., SATTAR, N., SHUNGIN, D., STRAWBRIDGE, R. J., TANAKA, T., VAN DUIJN, C. M., AN, P., DE ANDRADE, M., ANDREWS, J. S., ASPELUND, T., ATALAY, M., AULCHENKO, Y., BALKAU, B., BANDINELLI, S., BECKMANN, J. S., BEILBY, J. P., BELLIS, C., BERGMAN, R. N., BLANGERO, J., BOBAN, M., BOEHNKE, M., BOERWINKLE, E., BONNYCASTLE, L. L., BOOMSMA, D. I., BORECKI, I. B., BÖTTCHER, Y., BOUCHARD, C., BRUNNER, E., BUDIMIR, D., CAMPBELL, H., CARLSON, O., CHINES, P. S., CLARKE, R., COLLINS, F. S., CORBATÓN-ANCHUELO, A., COUPER, D., DE FAIRE, U., DEDOUSSIS, G. V., DELOUKAS, P., DIMITRIOU, M., EGAN, J. M., EIRIKSDOTTIR, G., ERDOS, M. R., ERIKSSON, J. G., EURY, E., FERRUCCI, L., FORD, I., FOROUHI, N. G., FOX, C. S., FRANZOSI, M. G., FRANKS, P. W., FRAYLING, T. M., FROGUEL, P., GALAN, P., DE GEUS, E., GIGANTE, B., GLAZER, N. L., GOEL, A., GROOP, L., GUDNASON, V., HALLMANS, G., HAMSTEN, A., HANSSON, O., HARRIS, T. B., HAYWARD, C., HEATH, S., HERCBERG, S., HICKS, A. A., HINGORANI, A., HOFMAN, A., HUI, J., HUNG, J., et al. 2012. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat Genet,* 44**,** 659-69.

MANOLIO, T. A., COLLINS, F. S., COX, N. J., GOLDSTEIN, D. B., HINDORFF, L. A., HUNTER, D. J., MCCARTHY, M. I., RAMOS, E. M., CARDON, L. R., CHAKRAVARTI, A., CHO, J. H., GUTTMACHER, A. E., KONG, A., KRUGLYAK, L., MARDIS, E., ROTIMI, C. N., SLATKIN, M., VALLE, D., WHITTEMORE, A. S., BOEHNKE, M., CLARK, A. G., EICHLER, E. E., GIBSON, G., HAINES, J. L., MACKAY, T. F.,

MCCARROLL, S. A. & VISSCHER, P. M. 2009. Finding the missing heritability of complex diseases. *Nature,* 461**,** 747-53.

MARQUES-VIDAL, P., MAZOYER, E., BONGARD, V., GOURDY, P., RUIDAVETS, J. B., DROUET, L. & FERRIERES, J. 2002. Prevalence of insulin resistance syndrome in southwestern France and its relationship with inflammatory and hemostatic markers. *Diabetes Care,* 25**,** 1371-7.

MARSHALL, M., LEWIS, S., LOCKWOOD, A., DRAKE, R., JONES, P. & CROUDACE, T. 2005. Association between duration of untreated psychosis and outcome in cohorts of first-episode patients: a systematic review. *Arch Gen Psychiatry,* 62**,** 975-83.

MARTIN, B. C., WARRAM, J. H., KROLEWSKI, A. S., BERGMAN, R. N., SOELDNER, J. S. & KAHN, C. R. 1992. Role of glucose and insulin resistance in development of type 2 diabetes mellitus: results of a 25-year follow-up study. *Lancet,* 340**,** 925-9.

MARTIN, D. J., UL-HAQ, Z., NICHOLL, B. I., CULLEN, B., EVANS, J., GILL, J. M., ROBERTS, B., GALLACHER, J., MACKAY, D., MCINTOSH, A., HOTOPF, M., CRADDOCK, N., DEARY, I. J., PELL, J. P. & SMITH, D. J. 2016. Cardiometabolic disease and features of depression and bipolar disorder: population-based, cross-sectional study. *Br J Psychiatry,* 208**,** 343-51.

MATAR, H. E., ALMERIE, M. Q., MAKHOUL, S., XIA, J. & HUMPHREYS, P. 2014. Pericyazine for schizophrenia. *Cochrane Database Syst Rev***,** CD007479.

MATHIESON, I. & MCVEAN, G. 2012. Differential confounding of rare and common variants in spatially structured populations. *Nat Genet,* 44**,** 243-6.

MATTSSON, M., MAHER, G. M., BOLAND, F., FITZGERALD, A. P., MURRAY, D. M. & BIESMA, R. 2019. Group-based trajectory modelling for BMI trajectories in childhood: A systematic review. *Obes Rev,* 20**,** 998-1015.

MAUDSLEY, H. 1895. *The Pathology Of Mind,* London, Facsimile Edition.

MAZUMDER, B., ALMOND, D., PARK, K., CRIMMINS, E. M. & FINCH, C. E. 2010. Lingering prenatal effects of the 1918 influenza pandemic on cardiovascular disease. *J Dev Orig Health Dis,* 1**,** 26-34.

MCCLELLAN, J. & WERRY, J. 1997. Practice parameters for the assessment and treatment of children and adolescents with schizophrenia. American Academy of Child and Adolescent Psychiatry. *J Am Acad Child Adolesc Psychiatry,* 36**,** 177S-93S.

MCCREADIE, R., MACDONALD, E., BLACKLOCK, C., TILAK-SINGH, D., WILES, D., HALLIDAY, J. & PATERSON, J. 1998. Dietary intake of schizophrenic patients in Nithsdale, Scotland: case-control study. *BMJ,* 317**,** 784-5.

MCCREADIE, R. G. & SCOTTISH SCHIZOPHRENIA LIFESTYLE, G. 2003. Diet, smoking and cardiovascular risk in people with schizophrenia: descriptive study. *Br J Psychiatry,* 183**,** 534-9.

MCCRORY, M. A., HARBAUGH, A. G., APPEADU, S. & ROBERTS, S. B. 2019. Fast-Food Offerings in the United States in 1986, 1991, and 2016 Show Large Increases in Food Variety, Portion Size, Dietary Energy, and Selected Micronutrients. *J Acad Nutr Diet,* 119**,** 923-933.

MCDAID, D., PARK, A. Counting the costs: the economic impacts of poor physical health in people with mental health problems. International Diabetes Federation, 2015 Brussels, Belgium.

MCGINNIS, J. M. & FOEGE, W. H. 1993. Actual causes of death in the United States. *JAMA,* 270**,** 2207-12.

MCGLASHAN, T. H. 1999. Duration of untreated psychosis in first-episode schizophrenia: marker or determinant of course? *Biol Psychiatry,* 46**,** 899-907.

MCGORRY, P. D., KILLACKEY, E. & YUNG, A. 2008. Early intervention in psychosis: concepts, evidence and future directions. *World Psychiatry,* 7**,** 148-56.

MCGOWAN, S., LAWRENCE, A. D., SALES, T., QUESTED, D. & GRASBY, P. 2004. Presynaptic dopaminergic dysfunction in schizophrenia: a positron emission tomographic [18F]fluorodopa study. *Arch Gen Psychiatry,* 61**,** 134-42.

MCGRATH, J., SAHA, S., CHANT, D. & WELHAM, J. 2008. Schizophrenia: a concise overview of incidence, prevalence, and mortality. *Epidemiol Rev,* 30**,** 67-76.

MCKENZIE, S. K., IMLACH GUNASEKARA, F., RICHARDSON, K. & CARTER, K. 2014. Do changes in socioeconomic factors lead to changes in mental health? Findings from three waves of a population based panel study. *J Epidemiol Community Health,* 68**,** 253-60.

MCLAUGHLIN, T., REAVEN, G., ABBASI, F., LAMENDOLA, C., SAAD, M., WATERS, D., SIMON, J. & KRAUSS, R. M. 2005. Is there a simple way to identify insulin-resistant individuals at increased risk of cardiovascular disease? *Am J Cardiol,* 96**,** 399-404.

MERLOT, E., COURET, D. & OTTEN, W. 2008. Prenatal stress, fetal imprinting and immunity. *Brain Behav Immun,* 22**,** 42-51.

METCALF, S. A., JONES, P. B., NORDSTROM, T., TIMONEN, M., MAKI, P., MIETTUNEN, J., JAASKELAINEN, E., JARVELIN, M. R., STOCHL, J., MURRAY, G. K., VEIJOLA, J. & KHANDAKER, G. M. 2017. Serum C-reactive protein in adolescence and risk of schizophrenia in adulthood: A prospective birth cohort study. *Brain Behav Immun,* 59**,** 253-259.

MILLER, B. J., BUCKLEY, P., SEABOLT, W., MELLOR, A. & KIRKPATRICK, B. 2011. Meta-analysis of cytokine alterations in schizophrenia: clinical status and antipsychotic effects. *Biol Psychiatry,* 70**,** 663-71.

MILLER, B. J., CULPEPPER, N. & RAPAPORT, M. H. 2014. C-reactive protein levels in schizophrenia: a review and meta-analysis. *Clin Schizophr Relat Psychoses,* 7**,** 223-30.

MILLER, M. A., KRUISBRINK, M., WALLACE, J., JI, C. & CAPPUCCIO, F. P. 2018. Sleep duration and incidence of obesity in infants, children, and adolescents: a systematic review and meta-analysis of prospective studies. *Sleep,* 41.

MIRETTI, M. M., WALSH, E. C., KE, X., DELGADO, M., GRIFFITHS, M., HUNT, S., MORRISON, J., WHITTAKER, P., LANDER, E. S., CARDON, L. R., BENTLEY, D. R., RIOUX, J. D., BECK, S. & DELOUKAS, P. 2005. A high-resolution linkage-disequilibrium map of the human major histocompatibility complex and first generation of tag single-nucleotide polymorphisms. *Am J Hum Genet,* 76**,** 634-46.

MISIAK, B., BARTOLI, F., STRAMECKI, F., SAMOCHOWIEC, J., LIS, M., KASZNIA, J., JAROSZ, K. & STANCZYKIEWICZ, B. 2019. Appetite regulating hormones in first-episode psychosis: A systematic review and meta-analysis. *Neurosci Biobehav Rev,* 102**,** 362-370.

MISIAK, B., STANCZYKIEWICZ, B., LACZMANSKI, L. & FRYDECKA, D. 2017. Lipid profile disturbances in antipsychotic-naive patients with first-episode non-affective psychosis: A systematic review and meta-analysis. *Schizophr Res*.

MITCHELL, A. J., VANCAMPFORT, D., DE HERDT, A., YU, W. & DE HERT, M. 2013a. Is the prevalence of metabolic syndrome and metabolic abnormalities increased in early schizophrenia? A comparative meta-analysis of first episode, untreated and treated patients. *Schizophr Bull,* 39**,** 295-305.

MITCHELL, A. J., VANCAMPFORT, D., SWEERS, K., VAN WINKEL, R., YU, W. & DE HERT, M. 2013b. Prevalence of metabolic syndrome and metabolic abnormalities in schizophrenia and related disorders--a systematic review and meta-analysis. *Schizophr Bull,* 39**,** 306-18.

MOEBUS, S., GORES, L., LOSCH, C. & JOCKEL, K. H. 2011. Impact of time since last caloric intake on blood glucose levels. *Eur J Epidemiol,* 26**,** 719-28.

MOHER, D., LIBERATI, A., TETZLAFF, J., ALTMAN, D. G. & GROUP, P. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med,* 6**,** e1000097.

MONS, U., MUEZZINLER, A., GELLERT, C., SCHOTTKER, B., ABNET, C. C., BOBAK, M., DE GROOT, L., FREEDMAN, N. D., JANSEN, E., KEE, F., KROMHOUT, D.,

KUULASMAA, K., LAATIKAINEN, T., O'DOHERTY, M. G., BUENO-DE-MESQUITA, B., ORFANOS, P., PETERS, A., VAN DER SCHOUW, Y. T., WILSGAARD, T., WOLK, A., TRICHOPOULOU, A., BOFFETTA, P., BRENNER, H. & CONSORTIUM, C. 2015. Impact of smoking and smoking cessation on cardiovascular events and mortality among older adults: meta-analysis of individual participant data from prospective cohort studies of the CHANCES consortium. *BMJ,* 350**,** h1551.

MONTEIRO, C. A., GOMES, F. S. & CANNON, G. 2010. The snack attack. *Am J Public Health,* 100**,** 975-81.

MONTEIRO, C. A., MOUBARAC, J. C., CANNON, G., NG, S. W. & POPKIN, B. 2013. Ultra-processed products are becoming dominant in the global food system. *Obes Rev,* 14 Suppl 2**,** 21-8.

MONTEIRO, R. & AZEVEDO, I. 2010. Chronic inflammation in obesity and the metabolic syndrome. *Mediators Inflamm,* 2010.

MONTEMAGNI, C., BELLINO, S., BRACALE, N., BOZZATELLO, P. & ROCCA, P. 2020. Models Predicting Psychosis in Patients With High Clinical Risk: A Systematic Review. *Front Psychiatry,* 11**,** 223.

MONTI, J. M. & MONTI, D. 2004. Sleep in schizophrenia patients and the effects of antipsychotic drugs. *Sleep Med Rev,* 8**,** 133-48.

MOONS, K. G., ROYSTON, P., VERGOUWE, Y., GROBBEE, D. E. & ALTMAN, D. G. 2009. Prognosis and prognostic research: what, why, and how? *BMJ,* 338**,** b375.

MOORE, H., JENTSCH, J. D., GHAJARNIA, M., GEYER, M. A. & GRACE, A. A. 2006. A neurobehavioral systems analysis of adult rats exposed to methylazoxymethanol acetate on E17: implications for the neuropathology of schizophrenia. *Biol Psychiatry,* 60**,** 253-64.

MORGAN, C., ABDUL-AL, R., LAPPIN, J. M., JONES, P., FEARON, P., LEESE, M., CROUDACE, T., MORGAN, K., DAZZAN, P., CRAIG, T., LEFF, J., MURRAY, R. & GROUP, A. S. 2006. Clinical and social determinants of duration of untreated psychosis in the AESOP first-episode psychosis study. *Br J Psychiatry,* 189**,** 446-52.

MORRIONE, A. 2000. Grb10 proteins in insulin-like growth factor and insulin receptor signaling (review). *Int J Mol Med,* 5**,** 151-4.

MORRIS, T. T., DAVIES, N. M., HEMANI, G. & SMITH, G. D. 2020. Population phenomena inflate genetic associations of complex social traits. *Sci Adv,* 6**,** eaay0328.

MURGUIA-ROMERO, M., JIMENEZ-FLORES, J. R., SIGRIST-FLORES, S. C., ESPINOZA-CAMACHO, M. A., JIMENEZ-MORALES, M., PINA, E., MENDEZ-CRUZ, A. R., VILLALOBOS-MOLINA, R. & REAVEN, G. M. 2013. Plasma triglyceride/HDL-

cholesterol ratio, insulin resistance, and cardiometabolic risk in young adults. *J Lipid Res,* 54**,** 2795-9.

MUTHEN, L. K., MUTHEN, B.O. 2017. Mplus User's Guide. Eighth Edition.

NANRI, A., NAKAGAWA, T., KUWAHARA, K., YAMAMOTO, S., HONDA, T., OKAZAKI, H., UEHARA, A., YAMAMOTO, M., MIYAMOTO, T., KOCHI, T., EGUCHI, M., MURAKAMI, T., SHIMIZU, C., SHIMIZU, M., TOMITA, K., NAGAHAMA, S., IMAI, T., NISHIHARA, A., SASAKI, N., HORI, A., SAKAMOTO, N., NISHIURA, C., TOTSUZAKI, T., KATO, N., FUKASAWA, K., HUANHUAN, H., AKTER, S., KUROTANI, K., KABE, I., MIZOUE, T., SONE, T., DOHI, S. & JAPAN EPIDEMIOLOGY COLLABORATION ON OCCUPATIONAL HEALTH STUDY, G. 2015. Development of Risk Score for Predicting 3-Year Incidence of Type 2 Diabetes: Japan Epidemiology Collaboration on Occupational Health Study. *PLoS One,* 10**,** e0142779.

NASH, A. I. 2017. Crosstalk between insulin and dopamine signaling: A basis for the metabolic effects of antipsychotic drugs. *J Chem Neuroanat,* 83-84**,** 59-68.

NATARAJAN, L., PU, M., FAN, J., LEVINE, R. A., PATTERSON, R. E., THOMSON, C. A., ROCK, C. L. & PIERCE, J. P. 2010. Measurement error of dietary self-report in intervention trials. *Am J Epidemiol,* 172**,** 819-27.

NATH, A. P., RITCHIE, S. C., GRINBERG, N. F., TANG, H. H., HUANG, Q. Q., TEO, S. M., AHOLA-OLLI, A. V., WURTZ, P., HAVULINNA, A. S., SANTALAHTI, K., PITKANEN, N., LEHTIMAKI, T., KAHONEN, M., LYYTIKAINEN, L. P., RAITOHARJU, E., SEPPALA, I., SARIN, A. P., RIPATTI, S., PALOTIE, A., PEROLA, M., VIIKARI, J. S., JALKANEN, S., MAKSIMOW, M., SALMI, M., WALLACE, C., RAITAKARI, O. T., SALOMAA, V., ABRAHAM, G., KETTUNEN, J. & INOUYE, M. 2019. Multivariate Genome-wide Association Analysis of a Cytokine Network Reveals Variants with Widespread Immune, Haematological, and Cardiometabolic Pleiotropy. *Am J Hum Genet,* 105**,** 1076-1090.

NATIONAL CENTER FOR HEALTH STATISTICS, D. O. H. I. S., , 2012. Crude and Age-Adjusted Percentage of Civilian, Noninstitutionalized Adults Living with Diagnosed Diabetes United States 1980-2010. Atlanta, GA.: National Center for Chronic Disease Prevention and Health Promotion.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE. 2003. Clinical Knowledge Summaries: CVD Risk Assessment and Management. London, UK.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE. 2010. Cardiovascular disease: risk assessment and reduction, including lipid modification. Clinical Guideline [CG181]. London, UK.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE. 2013. Psychosis and schizophrenia in children and young people: recognition and management.

NATTO, Z. S., YAGHMOOR, W., ALSHAERI, H. K. & VAN DYKE, T. E. 2019. Omega-3 Fatty Acids Effects on Inflammatory Biomarkers and Lipid Profiles among Diabetic and Cardiovascular Disease Patients: A Systematic Review and Meta-Analysis. *Sci Rep,* 9**,** 18867.

NAYLOR, C. P., M.; MCDAID, D.; KNAPP, M.; FOSSEY, M.; GALEA, A. 2012. Long-Term Conditions and Mental Health: The Cost of Comorbidities.

NESTEL, P., LYU, R., LOW, L. P., SHEU, W. H., NITIYANANT, W., SAITO, I. & TAN, C. E. 2007. Metabolic syndrome: recent prevalence in East and Southeast Asian populations. *Asia Pac J Clin Nutr,* 16**,** 362-7.

NETTIS, M. A., PERGOLA, G., KOLLIAKOU, A., O'CONNOR, J., BONACCORSO, S., DAVID, A., GAUGHRAN, F., DI FORTI, M., MURRAY, R. M., MARQUES, T. R., BLASI, G., BERTOLINO, A., PARIANTE, C. M., DAZZAN, P. & MONDELLI, V. 2019. Metabolic-inflammatory status as predictor of clinical outcome at 1-year follow-up in patients with first episode psychosis. *Psychoneuroendocrinology,* 99**,** 145-153.

NI, G., MOSER, G., SCHIZOPHRENIA WORKING GROUP OF THE PSYCHIATRIC GENOMICS, C., WRAY, N. R. & LEE, S. H. 2018. Estimation of Genetic Correlation via Linkage Disequilibrium Score Regression and Genomic Restricted Maximum Likelihood. *Am J Hum Genet,* 102**,** 1185-1194.

NIELSEN, P. R., MORTENSEN, P. B., DALMAN, C., HENRIKSEN, T. B., PEDERSEN, M. G., PEDERSEN, C. B. & AGERBO, E. 2013. Fetal growth and schizophrenia: a nested case-control and case-sibling study. *Schizophr Bull,* 39**,** 1337-42.

NILES, C. 2013. Examining the deinstitutionalization movement in North America. *Health Tomorrow,* 1**,** 54-83.

NODA, H., MARUYAMA, K., ISO, H., DOHI, S., TERAI, T., FUJIOKA, S., GOTO, K., HORIE, S., NAKANO, S., HIROBE, K. & "SAN-YU-KAI", M. S. P. C. O. T. J. A. O. O. P. 2010. Prediction of myocardial infarction using coronary risk scores among Japanese male workers: 3M Study. *J Atheroscler Thromb,* 17**,** 452-9.

NORDESTGAARD, B. G. & VARBO, A. 2014. Triglycerides and cardiovascular disease. *Lancet,* 384**,** 626-635.

NORTHSTONE, K., LEWCOCK, M., GROOM, A., BOYD, A., MACLEOD, J., TIMPSON, N. & WELLS, N. 2019. The Avon Longitudinal Study of Parents and Children (ALSPAC): an update on the enrolled sample of index children in 2019. *Wellcome Open Res,* 4**,** 51.

NOTARAS, M., HILL, R. & VAN DEN BUUSE, M. 2015. A role for the BDNF gene Val66Met polymorphism in schizophrenia? A comprehensive review. *Neurosci Biobehav Rev,* 51**,** 15-30.

NOUR, M. M. & HOWES, O. D. 2015. Interpreting the neurodevelopmental hypothesis of schizophrenia in the context of normal brain development and ageing. *Proc Natl Acad Sci U S A,* 112**,** E2745.

NURJONO, M., TAY, Y. H. & LEE, J. 2014. The relationship between serum brain-derived neurotrophic factor (BDNF) and cardiometabolic indices in schizophrenia. *Schizophr Res,* 157**,** 244-8.

O'RAHILLY, S., HATTERSLEY, A., VAAG, A. & GRAY, H. 1994. Insulin resistance as the major cause of impaired glucose tolerance: a self-fulfilling prophesy? *Lancet,* 344**,** 585-9.

ODA, E. 2012. Metabolic syndrome: its history, mechanisms, and limitations. *Acta Diabetol,* 49**,** 89-95.

ODGAARD-JENSEN, J., VIST, G. E., TIMMER, A., KUNZ, R., AKL, E. A., SCHUNEMANN, H., BRIEL, M., NORDMANN, A. J., PREGNO, S. & OXMAN, A. D. 2011. Randomisation to protect against selection bias in healthcare trials. *Cochrane Database Syst Rev***,** MR000012.

OGUNDIMU, E. O., ALTMAN, D. G. & COLLINS, G. S. 2016. Adequate sample size for developing prediction models is not simply related to events per variable. *J Clin Epidemiol,* 76**,** 175-82.

OGURTSOVA, K., DA ROCHA FERNANDES, J. D., HUANG, Y., LINNENKAMP, U., GUARIGUATA, L., CHO, N. H., CAVAN, D., SHAW, J. E. & MAKAROFF, L. E. 2017. IDF Diabetes Atlas: Global estimates for the prevalence of diabetes for 2015 and 2040. *Diabetes Res Clin Pract,* 128**,** 40-50.

OJEDA, N. B., GRIGORE, D. & ALEXANDER, B. T. 2008. Role of fetal programming in the development of hypertension. *Future Cardiol,* 4**,** 163-74.

OLDROYD, J., BANERJEE, M., HEALD, A. & CRUICKSHANK, K. 2005. Diabetes and ethnic minorities. *Postgrad Med J,* 81**,** 486-90.

OLFSON, M. 2009. Antipsychotic prescriptions for children and adolescents in the UK increased from 1993 to 2005. *Evid Based Ment Health,* 12**,** 30.

OSBORN, D. P., HARDOON, S., OMAR, R. Z., HOLT, R. I., KING, M., LARSEN, J., MARSTON, L., MORRIS, R. W., NAZARETH, I., WALTERS, K. & PETERSEN, I. 2015. Cardiovascular risk prediction models for people with severe mental illness: results from the prediction and management of cardiovascular risk in people with severe mental illnesses (PRIMROSE) research program. *JAMA Psychiatry,* 72**,** 143-51.

OSIMO, E. F., PERRY, B. I., CARDINAL, R. N., LYNALL, M. E., LEWIS, J., KUDCHADKAR, A., MURRAY, G. K., PEREZ, J., JONES, P. B. & KHANDAKER, G. M. 2021. Inflammatory and cardiometabolic markers at presentation with first episode psychosis and long-term clinical outcomes: A longitudinal study using electronic health records. *Brain Behav Immun,* 91**,** 117-127.

PADMANABHAN, J. L., NANDA, P., TANDON, N., MOTHI, S. S., BOLO, N., MCCARROLL, S., CLEMENTZ, B. A., GERSHON, E. S., PEARLSON, G. D., SWEENEY, J. A., TAMMINGA, C. A. & KESHAVAN, M. S. 2016. Polygenic risk for type 2 diabetes mellitus among individuals with psychosis and their relatives. *J Psychiatr Res,* 77**,** 52-8.

PANTOJA-TORRES, B., TORO-HUAMANCHUMO, C. J., URRUNAGA-PASTOR, D., GUARNIZO-POMA, M., LAZARO-ALCANTARA, H., PAICO-PALACIOS, S., DEL CARMEN RANILLA-SEGUIN, V., BENITES-ZAPATA, V. A., INSULIN, R. & METABOLIC SYNDROME RESEARCH, G. 2019. High triglycerides to HDL-cholesterol ratio is associated with insulin resistance in normal-weight healthy adults. *Diabetes Metab Syndr,* 13**,** 382-388.

PARDINAS, A. F., HOLMANS, P., POCKLINGTON, A. J., ESCOTT-PRICE, V., RIPKE, S., CARRERA, N., LEGGE, S. E., BISHOP, S., CAMERON, D., HAMSHERE, M. L., HAN, J., HUBBARD, L., LYNHAM, A., MANTRIPRAGADA, K., REES, E., MACCABE, J. H., MCCARROLL, S. A., BAUNE, B. T., BREEN, G., BYRNE, E. M., DANNLOWSKI, U., ELEY, T. C., HAYWARD, C., MARTIN, N. G., MCINTOSH, A. M., PLOMIN, R., PORTEOUS, D. J., WRAY, N. R., CABALLERO, A., GESCHWIND, D. H., HUCKINS, L. M., RUDERFER, D. M., SANTIAGO, E., SKLAR, P., STAHL, E. A., WON, H., AGERBO, E., ALS, T. D., ANDREASSEN, O. A., BAEKVAD-HANSEN, M., MORTENSEN, P. B., PEDERSEN, C. B., BORGLUM, A. D., BYBJERG-GRAUHOLM, J., DJUROVIC, S., DURMISHI, N., PEDERSEN, M. G., GOLIMBET, V., GROVE, J., HOUGAARD, D. M., MATTHEISEN, M., MOLDEN, E., MORS, O., NORDENTOFT, M., PEJOVIC-MILOVANCEVIC, M., SIGURDSSON, E., SILAGADZE, T., HANSEN, C. S., STEFANSSON, K., STEFANSSON, H., STEINBERG, S., TOSATO, S., WERGE, T., CONSORTIUM, G., CONSORTIUM, C., COLLIER, D. A., RUJESCU, D., KIROV, G., OWEN, M. J., O'DONOVAN, M. C., WALTERS, J. T. R., CONSORTIUM, G., CONSORTIUM, C., CONSORTIUM, G. & CONSORTIUM, C. 2018. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat Genet,* 50**,** 381-389.

PARK, Y., LIM, J., LEE, J. & KIM, S. G. 2009. Erythrocyte fatty acid profiles can predict acute non-fatal myocardial infarction. *Br J Nutr,* 102**,** 1355-61.

PAVLOU, M., AMBLER, G., SEAMAN, S. R., GUTTMANN, O., ELLIOTT, P., KING, M. & OMAR, R. Z. 2015. How to develop a more accurate risk prediction model when there are few events. *BMJ,* 351**,** h3868.

PAYNTER, N. P., CHASMAN, D. I., BURING, J. E., SHIFFMAN, D., COOK, N. R. & RIDKER, P. M. 2009. Cardiovascular disease risk prediction with and without knowledge of genetic variation at chromosome 9p21.3. *Ann Intern Med,* 150**,** 65-72.

PAYNTER, N. P., MAZER, N. A., PRADHAN, A. D., GAZIANO, J. M., RIDKER, P. M. & COOK, N. R. 2011. Cardiovascular risk prediction in diabetic men and women using hemoglobin A1c vs diabetes as a high-risk equivalent. *Arch Intern Med,* 171**,** 1712-8.

PEARSALL, R., SMITH, D. J. & GEDDES, J. R. 2019. Pharmacological and behavioural interventions to promote smoking cessation in adults with schizophrenia and bipolar disorders: a systematic review and meta-analysis of randomised trials. *BMJ Open,* 9**,** e027389.

PEDERSEN, J. M., MORTENSEN, E. L., MEINCKE, R. H., PETERSEN, G. L., BUDTZ-JORGENSEN, E., BRUNNSGAARD, H., SORENSEN, H. J. & LUND, R. 2019. Maternal infections during pregnancy and offspring midlife inflammation. *Matern Health Neonatol Perinatol,* 5**,** 4.

PEDUZZI, P., CONCATO, J., KEMPER, E., HOLFORD, T. R. & FEINSTEIN, A. R. 1996. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol,* 49**,** 1373-9.

PENCINA, M. J., D'AGOSTINO, R. B., LARSON, M. G., MASSARO, J. M. & VASAN, R. S. 2009. Predicting the 30-year risk of cardiovascular disease: the framingham heart study. *Circulation,* 119**,** 3078-84.

PENESOVA, A., CIZMAROVA, E., BELAN, V., BLAZICEK, P., IMRICH, R., VLCEK, M., VIGAS, M., SELKO, D., KOSKA, J. & RADIKOVA, Z. 2011. Insulin resistance in young, lean male subjects with essential hypertension. *J Hum Hypertens,* 25**,** 391-400.

PENNINX, B. W., BEEKMAN, A. T., HONIG, A., DEEG, D. J., SCHOEVERS, R. A., VAN EIJK, J. T. & VAN TILBURG, W. 2001. Depression and cardiac mortality: results from a community-based longitudinal study. *Arch Gen Psychiatry,* 58**,** 221-7.

PENTTILA, M., JAASKELAINEN, E., HIRVONEN, N., ISOHANNI, M. & MIETTUNEN, J. 2014. Duration of untreated psychosis as predictor of long-term outcome in schizophrenia: systematic review and meta-analysis. *Br J Psychiatry,* 205**,** 88-94.

PERALA, J., SUVISAARI, J., SAARNI, S. I., KUOPPASALMI, K., ISOMETSA, E., PIRKOLA, S., PARTONEN, T., TUULIO-HENRIKSSON, A., HINTIKKA, J., KIESEPPA, T.,

HARKANEN, T., KOSKINEN, S. & LONNQVIST, J. 2007. Lifetime prevalence of psychotic and bipolar I disorders in a general population. *Arch Gen Psychiatry,* 64**,** 19-28.

PEREZ-SANTIAGO, J., DIEZ-ALARCIA, R., CALLADO, L. F., ZHANG, J. X., CHANA, G., WHITE, C. H., GLATT, S. J., TSUANG, M. T., EVERALL, I. P., MEANA, J. J. & WOELK, C. H. 2012. A combined analysis of microarray gene expression studies of the human prefrontal cortex identifies genes implicated in schizophrenia. *J Psychiatr Res,* 46**,** 1464-74.

PERRY, B. I., BURGESS, S., JONES, H. J., ZAMMIT, S., UPTHEGROVE, R., MASON, A. M., DAY, F. R., LANGENBERG, C., WAREHAM, N. J., JONES, P. B. & KHANDAKER, G. M. 2021a. The potential shared role of inflammation in insulin resistance and schizophrenia: A bidirectional two-sample mendelian randomization study. *PLoS Med,* 18**,** e1003455.

PERRY, B. I., JONES, H. J., RICHARDSON, T. G., ZAMMIT, S., WAREHAM, N. J., LEWIS, G., JONES, P. B. & KHANDAKER, G. M. 2020a. Common mechanisms for type 2 diabetes and psychosis: Findings from a prospective birth cohort. *Schizophr Res,* 223**,** 227-235.

PERRY, B. I., KHANDAKER, G. M., MARWAHA, S., THOMPSON, A., ZAMMIT, S., SINGH, S. P. & UPTHEGROVE, R. 2020b. Insulin resistance and obesity, and their association with depression in relatively young people: findings from a large UK birth cohort. *Psychol Med,* 50**,** 556-565.

PERRY, B. I., MCINTOSH, G., WEICH, S., SINGH, S. & REES, K. 2016. The association between first-episode psychosis and abnormal glycaemic control: systematic review and meta-analysis. *Lancet Psychiatry,* 3**,** 1049-1058.

PERRY, B. I., STOCHL, J., UPTHEGROVE, R., ZAMMIT, S., WAREHAM, N., LANGENBERG, C., WINPENNY, E., DUNGER, D., JONES, P. B. & KHANDAKER, G. M. 2021b. Longitudinal Trends in Childhood Insulin Levels and Body Mass Index and Associations With Risks of Psychosis and Depression in Young Adults. *JAMA Psychiatry,* 78**,** 416-425.

PERRY, B. I., UPTHEGROVE, R., CRAWFORD, O., JANG, S., LAU, E., MCGILL, I., CARVER, E., JONES, P. B. & KHANDAKER, G. M. 2020c. Cardiometabolic risk prediction algorithms for young people with psychosis: a systematic review and exploratory analysis. *Acta Psychiatr Scand,* 142**,** 215-232.

PERRY, B. I., UPTHEGROVE, R., THOMPSON, A., MARWAHA, S., ZAMMIT, S., SINGH, S. P. & KHANDAKER, G. 2018. Dysglycaemia, Inflammation and Psychosis: Findings From the UK ALSPAC Birth Cohort. *Schizophr Bull*.

PETERS, S. A., SINGHATEH, Y., MACKAY, D., HUXLEY, R. R. & WOODWARD, M. 2016. Total cholesterol as a risk factor for coronary heart disease and stroke in women compared with men: A systematic review and meta-analysis. *Atherosclerosis,* 248**,** 123-31.

PETRUZZELLI, M. G., MARGARI, M., PESCHECHERA, A., DE GIAMBATTISTA, C., DE GIACOMO, A., MATERA, E. & MARGARI, F. 2018. Hyperprolactinemia and insulin resistance in drug naive patients with early onset first episode psychosis. *BMC Psychiatry,* 18**,** 246.

PIANO, M. R. 2017. Alcohol's Effects on the Cardiovascular System. *Alcohol Res,* 38**,** 219-241.

PIANO, M. R. & PHILLIPS, S. A. 2014. Alcoholic cardiomyopathy: pathophysiologic insights. *Cardiovasc Toxicol,* 14**,** 291-308.

PILLINGER, T., BECK, K., GOBJILA, C., DONOCIK, J. G., JAUHAR, S. & HOWES, O. D. 2017a. Impaired Glucose Homeostasis in First-Episode Schizophrenia: A Systematic Review and Meta-analysis. *JAMA Psychiatry,* 74**,** 261-269.

PILLINGER, T., BECK, K., STUBBS, B. & HOWES, O. D. 2017b. Cholesterol and triglyceride levels in first-episode psychosis: systematic review and meta-analysis. *Br J Psychiatry,* 211**,** 339-349.

PILLINGER, T., MCCUTCHEON, R. A., VANO, L., MIZUNO, Y., ARUMUHAM, A., HINDLEY, G., BECK, K., NATESAN, S., EFTHIMIOU, O., CIPRIANI, A. & HOWES, O. D. 2020. Comparative effects of 18 antipsychotics on metabolic function in patients with schizophrenia, predictors of metabolic dysregulation, and association with psychopathology: a systematic review and network meta-analysis. *Lancet Psychiatry,* 7**,** 64-77.

PISANU, C. & SQUASSINA, A. 2019. Treatment-Resistant Schizophrenia: Insights From Genetic Studies and Machine Learning Approaches. *Front Pharmacol,* 10**,** 617.

PLANA-RIPOLL, O., PEDERSEN, C. B., AGERBO, E., HOLTZ, Y., ERLANGSEN, A., CANUDAS-ROMO, V., ANDERSEN, P. K., CHARLSON, F. J., CHRISTENSEN, M. K., ERSKINE, H. E., FERRARI, A. J., IBURG, K. M., MOMEN, N., MORTENSEN, P. B., NORDENTOFT, M., SANTOMAURO, D. F., SCOTT, J. G., WHITEFORD, H. A., WEYE, N., MCGRATH, J. J. & LAURSEN, T. M. 2019. A comprehensive analysis of mortality-related health metrics associated with mental disorders: a nationwide, register-based cohort study. *Lancet,* 394**,** 1827-1835.

PLANA-RIPOLL, O., WEYE, N., MOMEN, N. C., CHRISTENSEN, M. K., IBURG, K. M., LAURSEN, T. M. & MCGRATH, J. J. 2020. Changes Over Time in the Differential Mortality Gap in Individuals With Mental Disorders. *JAMA Psychiatry*.

POLIMANTI, R., GELERNTER, J. & STEIN, D. J. 2017. Genetically determined schizophrenia is not associated with impaired glucose homeostasis. *Schizophr Res*.

POLUBRIAGINOF, F., SALMASIAN, H., ALBERT, D. A. & VAWDREY, D. K. 2017. Challenges with Collecting Smoking Status in Electronic Health Records. *AMIA Annu Symp Proc,* 2017**,** 1392-1400.

POPEJOY, A. B. & FULLERTON, S. M. 2016. Genomics is failing on diversity. *Nature,* 538**,** 161-164.

POPKIN, B. M. & DUFFEY, K. J. 2010. Does hunger and satiety drive eating anymore? Increasing eating occasions and decreasing time between eating occasions in the United States. *Am J Clin Nutr,* 91**,** 1342-7.

POULTON, R., CASPI, A., MOFFITT, T. E., CANNON, M., MURRAY, R. & HARRINGTON, H. 2000. Children's self-reported psychotic symptoms and adult schizophreniform disorder: a 15-year longitudinal study. *Arch Gen Psychiatry,* 57**,** 1053-8.

PRADHAN, A. D. 2014. Sex differences in the metabolic syndrome: implications for cardiovascular health in women. *Clin Chem,* 60**,** 44-52.

PRADHAN, A. D., MANSON, J. E., RIFAI, N., BURING, J. E. & RIDKER, P. M. 2001. C-reactive protein, interleukin 6, and risk of developing type 2 diabetes mellitus. *JAMA,* 286**,** 327-34.

PRICE, A. L., PATTERSON, N. J., PLENGE, R. M., WEINBLATT, M. E., SHADICK, N. A. & REICH, D. 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet,* 38**,** 904-9.

PULIT, S. L., STONEMAN, C., MORRIS, A. P., WOOD, A. R., GLASTONBURY, C. A., TYRRELL, J., YENGO, L., FERREIRA, T., MAROULI, E., JI, Y., YANG, J., JONES, S., BEAUMONT, R., CROTEAU-CHONKA, D. C., WINKLER, T. W., CONSORTIUM, G., HATTERSLEY, A. T., LOOS, R. J. F., HIRSCHHORN, J. N., VISSCHER, P. M., FRAYLING, T. M., YAGHOOTKAR, H. & LINDGREN, C. M. 2019. Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. *Hum Mol Genet,* 28**,** 166-174.

PURCELL, S., NEALE, B., TODD-BROWN, K., THOMAS, L., FERREIRA, M. A., BENDER, D., MALLER, J., SKLAR, P., DE BAKKER, P. I., DALY, M. J. & SHAM, P. C. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics,* 81**,** 559-75.

PURCELL, S. M., WRAY, N. R., STONE, J. L., VISSCHER, P. M., O'DONOVAN, M. C., SULLIVAN, P. F., SKLAR, P. & INTERNATIONAL SCHIZOPHRENIA CONSORTIUM 2009. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature,* 460**,** 748-52.

PYLYPCHUK, R., WELLS, S., KERR, A., POPPE, K., RIDDELL, T., HARWOOD, M., EXETER, D., MEHTA, S., GREY, C., WU, B. P., METCALF, P., WARREN, J., HARRISON, J., MARSHALL, R. & JACKSON, R. 2018. Cardiovascular disease risk prediction equations in 400 000 primary care patients in New Zealand: a derivation and validation study. *Lancet,* 391**,** 1897-1907.

R CORE TEAM. 2017. R: A Language and Environment for Statistical Computing. Available: https://www.R-project.org/.

RADER, D. J. & HOVINGH, G. K. 2014. HDL and cardiovascular disease. *Lancet,* 384**,** 618-625.

RAISON, C. L., RUTHERFORD, R. E., WOOLWINE, B. J., SHUO, C., SCHETTLER, P., DRAKE, D. F., HAROON, E. & MILLER, A. H. 2013. A randomized controlled trial of the tumor necrosis factor antagonist infliximab for treatment-resistant depression: the role of baseline inflammatory biomarkers. *JAMA Psychiatry,* 70**,** 31-41.

RALEIGH, V. 2020. What is happening to life expectancy in the UK?

RAM, N. & GRIMM, K. J. 2009. Growth Mixture Modeling: A Method for Identifying Differences in Longitudinal Change Among Unobserved Groups. *Int J Behav Dev,* 33**,** 565-576.

RATHMANN, W., STRASSBURGER, K., BONGAERTS, B., KUSS, O., MUSSIG, K., BURKART, V., SZENDROEDI, J., KOTZKA, J., KNEBEL, B., AL-HASANI, H., RODEN, M. & GROUP, G. D. S. 2019. A variant of the glucose transporter gene SLC2A2 modifies the glycaemic response to metformin therapy in recently diagnosed type 2 diabetes. *Diabetologia,* 62**,** 286-291.

REAVEN, G. M. 1988. Banting lecture 1988. Role of insulin resistance in human disease. *Diabetes,* 37**,** 1595-607.

REAVEN, G. M. 2001. Syndrome x: a short history. *Ochsner J,* 3**,** 124-5.

REAVEN, G. M. 2006. The metabolic syndrome: is this diagnosis necessary? *Am J Clin Nutr,* 83**,** 1237-47.

REINIKAINEN, J., LAATIKAINEN, T., KARVANEN, J. & TOLONEN, H. 2015. Lifetime cumulative risk factors predict cardiovascular disease mortality in a 50-year follow-up study in Finland. *Int J Epidemiol,* 44**,** 108-16.

RENAULT, V. M., RAFALSKI, V. A., MORGAN, A. A., SALIH, D. A., BRETT, J. O., WEBB, A. E., VILLEDA, S. A., THEKKAT, P. U., GUILLEREY, C., DENKO, N. C., PALMER, T. D., BUTTE, A. J. & BRUNET, A. 2009. FoxO3 regulates neural stem cell homeostasis. *Cell Stem Cell,* 5**,** 527-39.

RETHORST, C. D., BERNSTEIN, I. & TRIVEDI, M. H. 2014. Inflammation, obesity, and metabolic syndrome in depression: analysis of the 2009-2010 National Health and Nutrition Examination Survey (NHANES). *J Clin Psychiatry,* 75**,** e1428-32.

RIDKER, P. M., BURING, J. E., RIFAI, N. & COOK, N. R. 2007. Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score. *JAMA,* 297**,** 611-9.

RIDKER, P. M., PAYNTER, N. P., RIFAI, N., GAZIANO, J. M. & COOK, N. R. 2008. C-reactive protein and parental history improve global cardiovascular risk prediction: the Reynolds Risk Score for men. *Circulation,* 118**,** 2243-51, 4p following 2251.

RILEY, R. D., SNELL, K. I., ENSOR, J., BURKE, D. L., HARRELL, F. E., JR., MOONS, K. G. & COLLINS, G. S. 2019. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. *Stat Med,* 38**,** 1276-1296.

RILEY, R. D., VAN DER WINDT, D.A., CROFT, P., MOONS, K.G.M. 2019. *Prognosis Research in Healthcare,* Oxford, UK, Oxford University Press.

RINAUDO, P. & WANG, E. 2012. Fetal programming and metabolic syndrome. *Annu Rev Physiol,* 74**,** 107-30.

ROBINSON, C. A., AGARWAL, G. & NERENBERG, K. 2011. Validating the CANRISK prognostic model for assessing diabetes risk in Canada's multi-ethnic population. *Chronic Dis Inj Can,* 32**,** 19-31.

RODGERS, M., SOWDEN, A., PETTICREW, M., ARAI, L., ROBERTS, H., BRITTEN, N., POPAY, J. 2009. Testing Methodological Guidanc on the Conduct of Narrative Synthesis in Systematic Reviews: Effectiveness of Interventions to Promote Smoke Alarm Ownership and Function. *Evaluation,* 15**,** 47-71.

ROERECKE, M. & REHM, J. 2014. Cause-specific mortality risk in alcohol use disorder treatment patients: a systematic review and meta-analysis. *Int J Epidemiol,* 43**,** 906-19.

ROLLAND-CACHERA, M. F. & PENEAU, S. 2013. Growth trajectories associated with adult obesity. *World Rev Nutr Diet,* 106**,** 127-34.

ROYAL COLLEGE OF PSYCHIATRISTS. 2020. Early Intervention in Psychosis: Recording and reporting on NICE-recommended interventions using SNOMED CT codes. London, UK.

ROYUELA, A., MACIAS, J.A., GIL-VERONA, J.A., PASTOR, J.F., MANIEGA, M.A., ALONSO, J., ROMAN, J.M., DE PAZ, F., BARBOSA, M., RAMI-GONZALEZ, L., BOGET, T. 2002. Sleep in Schizophrenia: A Preliminary Study Using The Pittsburgh Sleep Quality Index. *Neurobiology of Sleep-Wakefulness Cycle,* 2**,** 37-39.

RUDIN, E. 1916. *Zer Verebung Und Neuentslehung der Dementia Praecox,* Berlin, Germany, Springer.

RUSSELL, A., CIUFOLINI, S., GARDNER-SOOD, P., BONACCORSO, S., GAUGHRAN, F., DAZZAN, P., PARIANTE, C. M. & MONDELLI, V. 2015. Inflammation and metabolic

changes in first episode psychosis: preliminary results from a longitudinal study. *Brain Behav Immun,* 49**,** 25-9.

SAGUD, M., VUKSAN-CUSA, B., JAKSIC, N., MIHALJEVIC-PELES, A., ROJNIC KUZMAN, M. & PIVAC, N. 2018. Smoking in Schizophrenia: an Updated Review. *Psychiatr Danub,* 30**,** 216-223.

SAHA, S., CHANT, D. & MCGRATH, J. 2007. A systematic review of mortality in schizophrenia: is the differential mortality gap worsening over time? *Arch Gen Psychiatry,* 64**,** 1123-31.

SAHA, S., CHANT, D., WELHAM, J. & MCGRATH, J. 2005. A systematic review of the prevalence of schizophrenia. *PLoS Med,* 2**,** e141.

SAKLAYEN, M. G. 2018. The Global Epidemic of the Metabolic Syndrome. *Curr Hypertens Rep,* 20**,** 12.

SALEM, R. O. & LAPOSATA, M. 2005. Effects of alcohol on hemostasis. *Am J Clin Pathol,* 123 Suppl**,** S96-105.

SALMAN, D., FAROOQI, M., MCGREGOR, A. & MAJEED, A. 2019. Time spent being sedentary: an emerging risk factor for poor health. *Br J Gen Pract,* 69**,** 278-279.

SAMUEL, V. T. & SHULMAN, G. I. 2016. The pathogenesis of insulin resistance: integrating signaling pathways and substrate flux. *J Clin Invest,* 126**,** 12-22.

SANDERSON, E., DAVEY SMITH, G., WINDMEIJER, F. & BOWDEN, J. 2018. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. *Int J Epidemiol*.

SANSBURY, F. H., FLANAGAN, S. E., HOUGHTON, J. A., SHUIXIAN SHEN, F. L., AL-SENANI, A. M., HABEB, A. M., ABDULLAH, M., KARIMINEJAD, A., ELLARD, S. & HATTERSLEY, A. T. 2012. SLC2A2 mutations can cause neonatal diabetes, suggesting GLUT2 may have a role in human insulin secretion. *Diabetologia,* 55**,** 2381-5.

SAUDER, K. A., STAMATOIU, A. V., LESHCHINSKAYA, E., RINGHAM, B. M., GLUECK, D. H. & DABELEA, D. 2019. Cord Blood Vitamin D Levels and Early Childhood Blood Pressure: The Healthy Start Study. *J Am Heart Assoc,* 8**,** e011485.

SAVAGE, D. B., PETERSEN, K. F. & SHULMAN, G. I. 2007. Disordered lipid metabolism and the pathogenesis of insulin resistance. *Physiol Rev,* 87**,** 507-20.

SAXENA, R., HIVERT, M. F., LANGENBERG, C., TANAKA, T., PANKOW, J. S., VOLLENWEIDER, P., LYSSENKO, V., BOUATIA-NAJI, N., DUPUIS, J., JACKSON, A. U., KAO, W. H., LI, M., GLAZER, N. L., MANNING, A. K., LUAN, J., STRINGHAM, H. M., PROKOPENKO, I., JOHNSON, T., GRARUP, N., BOESGAARD, T. W., LECOEUR, C., SHRADER, P., O'CONNELL, J., INGELSSON, E., COUPER, D. J., RICE, K., SONG, K., ANDREASEN, C. H., DINA, C., KÖTTGEN, A., LE BACQUER, O., PATTOU, F.,

TANEERA, J., STEINTHORSDOTTIR, V., RYBIN, D., ARDLIE, K., SAMPSON, M., QI, L., VAN HOEK, M., WEEDON, M. N., AULCHENKO, Y. S., VOIGHT, B. F., GRALLERT, H., BALKAU, B., BERGMAN, R. N., BIELINSKI, S. J., BONNEFOND, A., BONNYCASTLE, L. L., BORCH-JOHNSEN, K., BÖTTCHER, Y., BRUNNER, E., BUCHANAN, T. A., BUMPSTEAD, S. J., CAVALCANTI-PROENÇA, C., CHARPENTIER, G., CHEN, Y. D., CHINES, P. S., COLLINS, F. S., CORNELIS, M., J CRAWFORD, G., DELPLANQUE, J., DONEY, A., EGAN, J. M., ERDOS, M. R., FIRMANN, M., FOROUHI, N. G., FOX, C. S., GOODARZI, M. O., GRAESSLER, J., HINGORANI, A., ISOMAA, B., JØRGENSEN, T., KIVIMAKI, M., KOVACS, P., KROHN, K., KUMARI, M., LAURITZEN, T., LÉVY-MARCHAL, C., MAYOR, V., MCATEER, J. B., MEYRE, D., MITCHELL, B. D., MOHLKE, K. L., MORKEN, M. A., NARISU, N., PALMER, C. N., PAKYZ, R., PASCOE, L., PAYNE, F., PEARSON, D., RATHMANN, W., SANDBAEK, A., SAYER, A. A., SCOTT, L. J., SHARP, S. J., SIJBRANDS, E., SINGLETON, A., SISCOVICK, D. S., SMITH, N. L., SPARSØ, T., et al. 2010. Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge. *Nat Genet,* 42**,** 142-8.

SCHILLER, C. E., MELTZER-BRODY, S. & RUBINOW, D. R. 2015. The role of reproductive hormones in postpartum depression. *CNS Spectr,* 20**,** 48-59.

SCHIZOPHRENIA WORKING GROUP OF THE PSYCHIATRIC GENOMICS, C. 2014. Biological insights from 108 schizophrenia-associated genetic loci. *Nature,* 511**,** 421-7.

SCHIZOPHRENIA WORKING GROUP OF THE PSYCHIATRIC GENOMICS CONSORTIUM 2014. Biological insights from 108 schizophrenia-associated genetic loci. *Nature,* 511**,** 421-7.

SCHMITT, P., MANDEL, J., GUEDJ, M. 2015. A Comparison of Six Methods for Missing Data Imputation. *Journal of Biometrics & Biostatistics,* 6.

SCHULINGKAMP, R. J., PAGANO, T. C., HUNG, D. & RAFFA, R. B. 2000. Insulin receptors and insulin action in the brain: review and clinical implications. *Neurosci Biobehav Rev,* 24**,** 855-72.

SCHWINGSHACKL, L., HOFFMANN, G., LAMPOUSI, A. M., KNUPPEL, S., IQBAL, K., SCHWEDHELM, C., BECHTHOLD, A., SCHLESINGER, S. & BOEING, H. 2017. Food groups and risk of type 2 diabetes mellitus: a systematic review and meta-analysis of prospective studies. *Eur J Epidemiol,* 32**,** 363-375.

SCOTT, E. M., CARPENTER, J. S., IORFINO, F., CROSS, S. P. M., HERMENS, D. F., GEHUE, J., WILSON, C., WHITE, D., NAISMITH, S. L., GUASTELLA, A. J. & HICKIE, I. B. 2019. What is the prevalence, and what are the clinical correlates, of insulin resistance in

young people presenting for mental health care? A cross-sectional study. *BMJ Open,* 9**,** e025674.

SCOTT, R. A., LAGOU, V., WELCH, R. P., WHEELER, E., MONTASSER, M. E., LUAN, J., MAGI, R., STRAWBRIDGE, R. J., REHNBERG, E., GUSTAFSSON, S., KANONI, S., RASMUSSEN-TORVIK, L. J., YENGO, L., LECOEUR, C., SHUNGIN, D., SANNA, S., SIDORE, C., JOHNSON, P. C., JUKEMA, J. W., JOHNSON, T., MAHAJAN, A., VERWEIJ, N., THORLEIFSSON, G., HOTTENGA, J. J., SHAH, S., SMITH, A. V., SENNBLAD, B., GIEGER, C., SALO, P., PEROLA, M., TIMPSON, N. J., EVANS, D. M., POURCAIN, B. S., WU, Y., ANDREWS, J. S., HUI, J., BIELAK, L. F., ZHAO, W., HORIKOSHI, M., NAVARRO, P., ISAACS, A., O'CONNELL, J. R., STIRRUPS, K., VITART, V., HAYWARD, C., ESKO, T., MIHAILOV, E., FRASER, R. M., FALL, T., VOIGHT, B. F., RAYCHAUDHURI, S., CHEN, H., LINDGREN, C. M., MORRIS, A. P., RAYNER, N. W., ROBERTSON, N., RYBIN, D., LIU, C. T., BECKMANN, J. S., WILLEMS, S. M., CHINES, P. S., JACKSON, A. U., KANG, H. M., STRINGHAM, H. M., SONG, K., TANAKA, T., PEDEN, J. F., GOEL, A., HICKS, A. A., AN, P., MULLER-NURASYID, M., FRANCO-CERECEDA, A., FOLKERSEN, L., MARULLO, L., JANSEN, H., OLDEHINKEL, A. J., BRUINENBERG, M., PANKOW, J. S., NORTH, K. E., FOROUHI, N. G., LOOS, R. J., EDKINS, S., VARGA, T. V., HALLMANS, G., OKSA, H., ANTONELLA, M., NAGARAJA, R., TROMPET, S., FORD, I., BAKKER, S. J., KONG, A., KUMARI, M., GIGANTE, B., HERDER, C., MUNROE, P. B., CAULFIELD, M., ANTTI, J., MANGINO, M., SMALL, K., MILJKOVIC, I., et al. 2012. Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat Genet,* 44**,** 991-1005.

SENTISSI, O., VIALA, A., BOURDEL, M. C., KAMINSKI, F., BELLISLE, F., OLIE, J. P. & POIRIER, M. F. 2009. Impact of antipsychotic treatments on the motivation to eat: preliminary results in 153 schizophrenic patients. *Int Clin Psychopharmacol,* 24**,** 257-64.

SHAH, N. R. & BRAVERMAN, E. R. 2012. Measuring adiposity in patients: the utility of body mass index (BMI), percent body fat, and leptin. *PLoS One,* 7**,** e33308.

SHEPHERD, M., WATT, D., FALLOON, I. & SMEETON, N. 1989. The natural history of schizophrenia: a five-year follow-up study of outcome and prediction in a representative sample of schizophrenics. *Psychol Med Monogr Suppl,* 15**,** 1-46.

SHI, H., MANCUSO, N., SPENDLOVE, S. & PASANIUC, B. 2017. Local Genetic Correlation Gives Insights into the Shared Genetic Architecture of Complex Traits. *Am J Hum Genet,* 101**,** 737-751.

SHIN, J. A., LEE, J. H., LIM, S. Y., HA, H. S., KWON, H. S., PARK, Y. M., LEE, W. C., KANG, M. I., YIM, H. W., YOON, K. H. & SON, H. Y. 2013. Metabolic syndrome as a predictor of type 2 diabetes, and its clinical interpretations and usefulness. *J Diabetes Investig,* 4**,** 334-43.

SHULMAN, G. I. 2000. Cellular mechanisms of insulin resistance. *J Clin Invest,* 106**,** 171-6.

SIMEONE, J. C., WARD, A. J., ROTELLA, P., COLLINS, J. & WINDISCH, R. 2015. An evaluation of variation in published estimates of schizophrenia prevalence from 1990 horizontal line 2013: a systematic literature review. *BMC Psychiatry,* 15**,** 193.

SINGH, T., WALTERS, J. T. R., JOHNSTONE, M., CURTIS, D., SUVISAARI, J., TORNIAINEN, M., REES, E., IYEGBE, C., BLACKWOOD, D., MCINTOSH, A. M., KIROV, G., GESCHWIND, D., MURRAY, R. M., DI FORTI, M., BRAMON, E., GANDAL, M., HULTMAN, C. M., SKLAR, P., STUDY, I., CONSORTIUM, U. K., PALOTIE, A., SULLIVAN, P. F., O'DONOVAN, M. C., OWEN, M. J. & BARRETT, J. C. 2017. The contribution of rare variants to risk of schizophrenia in individuals with and without intellectual disability. *Nat Genet,* 49**,** 1167-1173.

SIRUGO, G., WILLIAMS, S. M. & TISHKOFF, S. A. 2019. The Missing Diversity in Human Genetic Studies. *Cell,* 177**,** 26-31.

SKINNER, A. C., PERRIN, E. M. & SKELTON, J. A. 2016. Prevalence of obesity and severe obesity in US children, 1999-2014. *Obesity (Silver Spring),* 24**,** 1116-23.

SLACK, J. & EVANS, K. A. 1966. The increased risk of death from ischaemic heart disease in first degree relatives of 121 men and 96 women with ischaemic heart disease. *J Med Genet,* 3**,** 239-57.

SLOPEN, N., KUBZANSKY, L. D., MCLAUGHLIN, K. A. & KOENEN, K. C. 2013. Childhood adversity and inflammatory processes in youth: a prospective study. *Psychoneuroendocrinology,* 38**,** 188-200.

SLUIK, D., BROUWER-BROLSMA, E. M., DE VRIES, J. H., GEELEN, A. & FESKENS, E. J. 2016. Associations of alcoholic beverage preference with cardiometabolic and lifestyle factors: the NQplus study. *BMJ Open,* 6**,** e010437.

SMELAND, O. B., FREI, O., KAUPPI, K., HILL, W. D., LI, W., WANG, Y., KRULL, F., BETTELLA, F., ERIKSEN, J. A., WITOELAR, A., DAVIES, G., FAN, C. C., THOMPSON, W. K., LAM, M., LENCZ, T., CHEN, C. H., UELAND, T., JONSSON, E. G., DJUROVIC, S., DEARY, I. J., DALE, A. M., ANDREASSEN, O. A. & NEURO, C. C. W. G. 2017. Identification of Genetic Loci Jointly Influencing Schizophrenia Risk and the Cognitive Traits of Verbal-Numerical Reasoning, Reaction Time, and General Cognitive Function. *JAMA Psychiatry,* 74**,** 1065-1075.

SMITH, G. D. 2010. Mendelian Randomization for Strengthening Causal Inference in Observational Studies: Application to Gene x Environment Interactions. *Perspect Psychol Sci,* 5**,** 527-45.

SMITH, G. D. & EBRAHIM, S. 2003. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol,* 32**,** 1-22.

SMITH, M., HOPKINS, D., PEVELER, R. C., HOLT, R. I., WOODWARD, M. & ISMAIL, K. 2008. First- v. second-generation antipsychotics and risk for diabetes in schizophrenia: systematic review and meta-analysis. *Br J Psychiatry,* 192**,** 406-11.

SMITH, S. M. & VALE, W. W. 2006. The role of the hypothalamic-pituitary-adrenal axis in neuroendocrine responses to stress. *Dialogues Clin Neurosci,* 8**,** 383-95.

SO, H. C., CHAU, K. L., AO, F. K., MO, C. H. & SHAM, P. C. 2019. Exploring shared genetic bases and causal relationships of schizophrenia and bipolar disorder with 28 cardiovascular and metabolic traits. *Psychol Med,* 49**,** 1286-1298.

SOARES, C. N. & ZITEK, B. 2008. Reproductive hormone sensitivity and risk for depression across the female life cycle: a continuum of vulnerability? *J Psychiatry Neurosci,* 33**,** 331-43.

SORENSEN, H. J., MORTENSEN, E. L., REINISCH, J. M. & MEDNICK, S. A. 2006. Height, weight and body mass index in early adulthood and risk of schizophrenia. *Acta Psychiatr Scand,* 114**,** 49-54.

SPERTUS, J., HORVITZ-LENNON, M., ABING, H. & NORMAND, S. L. 2018. Risk of weight gain for specific antipsychotic drugs: a meta-analysis. *NPJ Schizophr,* 4**,** 12.

SPEYER, H., CHRISTIAN BRIX NORGAARD, H., BIRK, M., KARLSEN, M., STORCH JAKOBSEN, A., PEDERSEN, K., HJORTHOJ, C., PISINGER, C., GLUUD, C., MORS, O., KROGH, J. & NORDENTOFT, M. 2016. The CHANGE trial: no superiority of lifestyle coaching plus care coordination plus treatment as usual compared to treatment as usual alone in reducing risk of cardiovascular disease in adults with schizophrenia spectrum disorders and abdominal obesity. *World Psychiatry,* 15**,** 155-65.

SROUR, B., FEZEU, L. K., KESSE-GUYOT, E., ALLES, B., MEJEAN, C., ANDRIANASOLO, R. M., CHAZELAS, E., DESCHASAUX, M., HERCBERG, S., GALAN, P., MONTEIRO, C. A., JULIA, C. & TOUVIER, M. 2019. Ultra-processed food intake and risk of cardiovascular disease: prospective cohort study (NutriNet-Sante). *BMJ,* 365**,** l1451.

STALEY, J. R., BLACKSHAW, J., KAMAT, M. A., ELLIS, S., SURENDRAN, P., SUN, B. B., PAUL, D. S., FREITAG, D., BURGESS, S., DANESH, J., YOUNG, R. &

BUTTERWORTH, A. S. 2016. PhenoScanner: a database of human genotype-phenotype associations. *Bioinformatics,* 32**,** 3207-3209.

STARRENBURG, F. C. & BOGERS, J. P. 2009. How can antipsychotics cause Diabetes Mellitus? Insights based on receptor-binding profiles, humoral factors and transporter proteins. *Eur Psychiatry,* 24**,** 164-70.

STEFANIS, N. C., HANSSEN, M., SMIRNIS, N. K., AVRAMOPOULOS, D. A., EVDOKIMIDIS, I. K., STEFANIS, C. N., VERDOUX, H. & VAN OS, J. 2002. Evidence that three dimensions of psychosis have a distribution in the general population. *Psychol Med,* 32**,** 347-58.

STEINER, J., BERGER, M., GUEST, P. C., DOBROWOLNY, H., WESTPHAL, S., SCHILTZ, K. & SARNYAI, Z. 2017. Assessment of Insulin Resistance Among Drug-Naive Patients With First-Episode Schizophrenia in the Context of Hormonal Stress Axis Activation. *JAMA Psychiatry,* 74**,** 968-970.

STEYERBERG, E. W., EIJKEMANS, M. J., HARRELL, F. E., JR. & HABBEMA, J. D. 2000. Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. *Stat Med,* 19**,** 1059-79.

STEYERBERG, E. W. & VERGOUWE, Y. 2014. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J,* 35**,** 1925-31.

STILO, S. A. & MURRAY, R. M. 2010. The epidemiology of schizophrenia: replacing dogma with knowledge. *Dialogues Clin Neurosci,* 12**,** 305-15.

STRAMECKI, F., KOTOWICZ, K. D., PIOTROWSKI, P., FRYDECKA, D., RYMASZEWSKA, J., BESZLEJ, J. A., SAMOCHOWIEC, J., JABLONSKI, M., WRONSKI, M., MOUSTAFA, A. A. & MISIAK, B. 2018. Assessment of the Association Between Cigarette Smoking and Cognitive Performance in Patients With Schizophrenia-Spectrum Disorders: A Case-Control Study. *Front Psychiatry,* 9**,** 642.

STRINGER, S., KAHN, R. S., DE WITTE, L. D., OPHOFF, R. A. & DERKS, E. M. 2014. Genetic liability for schizophrenia predicts risk of immune disorders. *Schizophr Res,* 159**,** 347-52.

SUBRAMANIAN, J. & SIMON, R. 2013. Overfitting in prediction models - is it a problem only in high dimensions? *Contemp Clin Trials,* 36**,** 636-41.

SULLIVAN, P. F., KENDLER, K. S. & NEALE, M. C. 2003. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Arch Gen Psychiatry,* 60**,** 1187-92.

SULLIVAN, S. A., KOUNALI, D., CANNON, M., DAVID, A. S., FLETCHER, P. C., HOLMANS, P., JONES, H., JONES, P. B., LINDEN, D. E. J., LEWIS, G., OWEN, M. J., O'DONOVAN, M., RAMMOS, A., THOMPSON, A., WOLKE, D., HERON, J. & ZAMMIT, S. 2020. A Population-Based Cohort Study Examining the Incidence and Impact

of Psychotic Experiences From Childhood to Adulthood, and Prediction of Psychotic Disorder. *Am J Psychiatry*, appiajp201919060654.

SUNDQUIST, K., WINKLEBY, M., LI, X., JI, J., HEMMINKI, K. & SUNDQUIST, J. 2011. Familial [corrected] transmission of coronary heart disease: a cohort study of 80,214 Swedish adoptees linked to their biological and adoptive parents. *Am Heart J,* 162**,** 317-23.

SWERDLOW, D. I., HOLMES, M. V., KUCHENBAECKER, K. B., ENGMANN, J. E., SHAH, T., SOFAT, R., GUO, Y., CHUNG, C., PEASEY, A., PFISTER, R., MOOIJAART, S. P., IRELAND, H. A., LEUSINK, M., LANGENBERG, C., LI, K. W., PALMEN, J., HOWARD, P., COOPER, J. A., DRENOS, F., HARDY, J., NALLS, M. A., LI, Y. R., LOWE, G., STEWART, M., BIELINSKI, S. J., PETO, J., TIMPSON, N. J., GALLACHER, J., DUNLOP, M., HOULSTON, R., TOMLINSON, I., TZOULAKI, I., LUAN, J., BOER, J. M., FOROUHI, N. G., ONLAND-MORET, N. C., VAN DER SCHOUW, Y. T., SCHNABEL, R. B., HUBACEK, J. A., KUBINOVA, R., BACEVICIENE, M., TAMOSIUNAS, A., PAJAK, A., TOPOR-MADRY, R., MALYUTINA, S., BALDASSARRE, D., SENNBLAD, B., TREMOLI, E., DE FAIRE, U., FERRUCCI, L., BANDENELLI, S., TANAKA, T., MESCHIA, J. F., SINGLETON, A., NAVIS, G., MATEO LEACH, I., BAKKER, S. J., GANSEVOORT, R. T., FORD, I., EPSTEIN, S. E., BURNETT, M. S., DEVANEY, J. M., JUKEMA, J. W., WESTENDORP, R. G., JAN DE BORST, G., VAN DER GRAAF, Y., DE JONG, P. A., MAILAND-VAN DER ZEE, A. H., KLUNGEL, O. H., DE BOER, A., DOEVENDANS, P. A., STEPHENS, J. W., EATON, C. B., ROBINSON, J. G., MANSON, J. E., FOWKES, F. G., FRAYLING, T. M., PRICE, J. F., WHINCUP, P. H., MORRIS, R. W., LAWLOR, D. A., SMITH, G. D., BEN-SHLOMO, Y., REDLINE, S., LANGE, L. A., KUMARI, M., WAREHAM, N. J., VERSCHUREN, W. M., BENJAMIN, E. J., WHITTAKER, J. C., HAMSTEN, A., DUDBRIDGE, F., DELANEY, J. A., WONG, A., KUH, D., HARDY, R., CASTILLO, B. A., CONNOLLY, J. J., VAN DER HARST, P., BRUNNER, E. J., et al. 2012. The interleukin-6 receptor as a target for prevention of coronary heart disease: a mendelian randomisation analysis. *Lancet,* 379**,** 1214-24.

TABAK, A. G., HERDER, C., RATHMANN, W., BRUNNER, E. J. & KIVIMAKI, M. 2012. Prediabetes: a high-risk state for diabetes development. *Lancet,* 379**,** 2279-90.

TAKADA, T., NIJMAN, S., DENAXAS, S., SNELL, K. I. E., UIJL, A., NGUYEN, T. L., ASSELBERGS, F. W. & DEBRAY, T. P. A. 2021. Internal-external cross-validation helped to evaluate the generalizability of prediction models in large clustered datasets. *J Clin Epidemiol*.

TANI, M., AKASHI, N., HORI, K., KONISHI, K., KITAJIMA, Y., TOMIOKA, H., INAMOTO, A., HIRATA, A., TOMITA, A., KOGANEMARU, T., TAKAHASHI, A. & HACHISU, M. 2015. Anticholinergic Activity and Schizophrenia. *Neurodegener Dis,* 15**,** 168-74.

TASCI, I., KABUL, H. K. & AYDOGDU, A. 2012. Brain derived neurotrophic factor (BDNF) in cardiometabolic physiology and diseases. *Anadolu Kardiyol Derg,* 12**,** 684-8.

THALHAMER, T., MCGRATH, M. A. & HARNETT, M. M. 2008. MAPKs and their relevance to arthritis and inflammation. *Rheumatology (Oxford),* 47**,** 409-14.

THARA, R. 2004. Twenty-year course of schizophrenia: the Madras Longitudinal Study. *Can J Psychiatry,* 49**,** 564-9.

TIENARI, P., WYNNE, L. C., MORING, J., LAHTI, I., NAARALA, M., SORRI, A., WAHLBERG, K. E., SAARENTO, O., SEITAMAA, M., KALEVA, M. & ET AL. 1994. The Finnish adoptive family study of schizophrenia. Implications for family research. *Br J Psychiatry Suppl***,** 20-6.

TIIKKAJA, S., SANDIN, S., MALKI, N., MODIN, B., SPAREN, P. & HULTMAN, C. M. 2013. Social class, social mobility and risk of psychiatric disorder--a population-based longitudinal study. *PLoS One,* 8**,** e77975.

TILLIN, T., FOROUHI, N., JOHNSTON, D. G., MCKEIGUE, P. M., CHATURVEDI, N. & GODSLAND, I. F. 2005. Metabolic syndrome and coronary heart disease in South Asians, African-Caribbeans and white Europeans: a UK population-based cross-sectional study. *Diabetologia,* 48**,** 649-56.

TIMONEN, M., RAJALA, U., JOKELAINEN, J., KEINANEN-KIUKAANNIEMI, S., MEYER-ROCHOW, V. B. & RASANEN, P. 2006. Depressive symptoms and insulin resistance in young adult males: results from the Northern Finland 1966 birth cohort. *Mol Psychiatry,* 11**,** 929-33.

TOMASIK, J., LAGO, S. G., VAZQUEZ-BOURGON, J., PAPIOL, S., SUAREZ-PINILLA, P., CRESPO-FACORRO, B. & BAHN, S. 2019. Association of Insulin Resistance With Schizophrenia Polygenic Risk Score and Response to Antipsychotic Treatment. *JAMA Psychiatry*.

TOSATO, S., BONETTO, C., TOMASSI, S., ZANARDINI, R., FARAVELLI, C., BRUSCHI, C., D'AGOSTINO, A., MINELLI, A., SCOCCO, P., LASALVIA, A., FURLATO, K., IMBESI, M., PRETI, A., RUGGERI, M., GENNARELLI, M., GET UP, G. & BOCCHIO-CHIAVETTO, L. 2020. Childhood trauma and glucose metabolism in patients with first-episode psychosis. *Psychoneuroendocrinology,* 113**,** 104536.

TOWNSEND, D. K., MCGREGOR, K., WU, E., CIALKOWSKI, K., HAUB, M. D. & BARSTOW, T. J. 2018. Insulin resistance and metabolic syndrome criteria in lean, normoglycemic college-age subjects. *Diabetes Metab Syndr,* 12**,** 609-616.

TURICCHI, J., O'DRISCOLL, R., HORGAN, G., DUARTE, C., PALMEIRA, A. L., LARSEN, S. C., HEITMANN, B. L. & STUBBS, J. 2020. Weekly, seasonal and holiday body weight fluctuation patterns among individuals engaged in a European multi-centre behavioural weight loss maintenance intervention. *PLoS One,* 15**,** e0232152.

TURPEINEN, H., RAITOHARJU, E., OKSANEN, A., OKSALA, N., LEVULA, M., LYYTIKAINEN, L. P., JARVINEN, O., CREEMERS, J. W., KAHONEN, M., LAAKSONEN, R., PELTO-HUIKKO, M., LEHTIMAKI, T. & PESU, M. 2011. Proprotein convertases in human atherosclerotic plaques: the overexpression of FURIN and its substrate cytokines BAFF and APRIL. *Atherosclerosis,* 219**,** 799-806.

TYLEE, D. S., SUN, J., HESS, J. L., TAHIR, M. A., SHARMA, E., MALIK, R., WORRALL, B. B., LEVINE, A. J., MARTINSON, J. J., NEJENTSEV, S., SPEED, D., FISCHER, A., MICK, E., WALKER, B. R., CRAWFORD, A., GRANT, S. F. A., POLYCHRONAKOS, C., BRADFIELD, J. P., SLEIMAN, P. M. A., HAKONARSON, H., ELLINGHAUS, E., ELDER, J. T., TSOI, L. C., TREMBATH, R. C., BARKER, J. N., FRANKE, A., DEHGHAN, A., ME RESEARCH, T., INFLAMMATION WORKING GROUP OF THE, C. C., CONSORTIUM, M. C. O. T. I. S. G., NETHERLANDS TWIN, R., NEURO, C. W. G., OBSESSIVE, C., TOURETTE SYNDROME WORKING GROUP OF THE PSYCHIATRIC GENOMICS, C., FARAONE, S. V. & GLATT, S. J. 2018. Genetic correlations among psychiatric and immune-related phenotypes based on genome-wide association data. *Am J Med Genet B Neuropsychiatr Genet,* 177**,** 641-657.

TYRRELL, J., MULUGETA, A., WOOD, A. R., ZHOU, A., BEAUMONT, R. N., TUKE, M. A., JONES, S. E., RUTH, K. S., YAGHOOTKAR, H., SHARP, S., THOMPSON, W. D., JI, Y., HARRISON, J., FREATHY, R. M., MURRAY, A., WEEDON, M. N., LEWIS, C., FRAYLING, T. M. & HYPPONEN, E. 2019. Using genetics to understand the causal influence of higher BMI on depression. *Int J Epidemiol,* 48**,** 834-848.

TYRRELL, J., ZHENG, J., BEAUMONT, R., HINTON, K., RICHARDSON, T. G., WOOD, A. R., DAVEY SMITH, G., FRAYLING, T. M. & TILLING, K. 2021. Genetic predictors of participation in optional components of UK Biobank. *Nat Commun,* 12**,** 886.

UNGER, J. W., LIVINGSTON, J. N. & MOSS, A. M. 1991. Insulin receptors in the central nervous system: localization, signalling mechanisms and functional aspects. *Prog Neurobiol,* 36**,** 343-62.

UPTHEGROVE, R., MANZANARES-TESON, N. & BARNES, N. M. 2014. Cytokine function in medication-naive first episode psychosis: a systematic review and meta-analysis. *Schizophr Res,* 155**,** 101-8.

VAN CALSTER, B., MCLERNON, D. J., VAN SMEDEN, M., WYNANTS, L., STEYERBERG, E. W., TOPIC GROUP 'EVALUATING DIAGNOSTIC, T. & PREDICTION MODELS' OF THE, S. I. 2019. Calibration: the Achilles heel of predictive analytics. *BMC Med,* 17**,** 230.

VAN DER HARST, P. & VERWEIJ, N. 2018. Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. *Circ Res,* 122**,** 433-443.

VAN DER PLOEG, T., NIEBOER, D. & STEYERBERG, E. W. 2016. Modern modeling techniques had limited external validity in predicting mortality from traumatic brain injury. *J Clin Epidemiol,* 78**,** 83-89.

VAN GAAL, L. F., MERTENS, I. L. & DE BLOCK, C. E. 2006. Mechanisms linking obesity with cardiovascular disease. *Nature,* 444**,** 875-80.

VAN MELLE, J. P., DE JONGE, P., SPIJKERMAN, T. A., TIJSSEN, J. G., ORMEL, J., VAN VELDHUISEN, D. J., VAN DEN BRINK, R. H. & VAN DEN BERG, M. P. 2004. Prognostic association of depression following myocardial infarction with mortality and cardiovascular events: a meta-analysis. *Psychosom Med,* 66**,** 814-22.

VAN OS, J., LINSCOTT, R. J., MYIN-GERMEYS, I., DELESPAUL, P. & KRABBENDAM, L. 2009. A systematic review and meta-analysis of the psychosis continuum: evidence for a psychosis proneness-persistence-impairment model of psychotic disorder. *Psychol Med,* 39**,** 179-95.

VANCAMPFORT, D., FIRTH, J., SCHUCH, F. B., ROSENBAUM, S., MUGISHA, J., HALLGREN, M., PROBST, M., WARD, P. B., GAUGHRAN, F., DE HERT, M., CARVALHO, A. F. & STUBBS, B. 2017. Sedentary behavior and physical activity levels in people with schizophrenia, bipolar disorder and major depressive disorder: a global systematic review and meta-analysis. *World Psychiatry,* 16**,** 308-315.

VANCAMPFORT, D., STUBBS, B., MITCHELL, A. J., DE HERT, M., WAMPERS, M., WARD, P. B., ROSENBAUM, S. & CORRELL, C. U. 2015. Risk of metabolic syndrome and its components in people with schizophrenia and related psychotic disorders, bipolar disorder and major depressive disorder: a systematic review and meta-analysis. *World Psychiatry,* 14**,** 339-47.

VARGHESE, D., SCOTT, J., WELHAM, J., BOR, W., NAJMAN, J., O'CALLAGHAN, M., WILLIAMS, G. & MCGRATH, J. 2011. Psychotic-like experiences in major depression

and anxiety disorders: a population-based survey in young adults. *Schizophr Bull,* 37**,** 389-93.

VERBANCK, M., CHEN, C. Y., NEALE, B. & DO, R. 2018. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet,* 50**,** 693-698.

VICKERS, A. J. & ELKIN, E. B. 2006. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making,* 26**,** 565-74.

VICKERS, A. J., VAN CALSTER, B. & STEYERBERG, E. W. 2019. A simple, step-by-step guide to interpreting decision curve analysis. *Diagn Progn Res,* 3**,** 18.

VIEIRA, B. A., LUFT, V. C., SCHMIDT, M. I., CHAMBLESS, L. E., CHOR, D., BARRETO, S. M. & DUNCAN, B. B. 2016. Timing and Type of Alcohol Consumption and the Metabolic Syndrome - ELSA-Brasil. *PLoS One,* 11**,** e0163044.

VITA, A., MINELLI, A., BARLATI, S., DESTE, G., GIACOPUZZI, E., VALSECCHI, P., TURRINA, C. & GENNARELLI, M. 2019. Treatment-Resistant Schizophrenia: Genetic and Neuroimaging Correlates. *Front Pharmacol,* 10**,** 402.

VOGELEZANG, S., BRADFIELD, J. P., AHLUWALIA, T. S., CURTIN, J. A., LAKKA, T. A., GRARUP, N., SCHOLZ, M., VAN DER MOST, P. J., MONNEREAU, C., STERGIAKOULI, E., HEISKALA, A., HORIKOSHI, M., FEDKO, I. O., VILOR-TEJEDOR, N., COUSMINER, D. L., STANDL, M., WANG, C. A., VIIKARI, J., GELLER, F., INIGUEZ, C., PITKANEN, N., CHESI, A., BACELIS, J., YENGO, L., TORRENT, M., NTALLA, I., HELGELAND, O., SELZAM, S., VONK, J. M., ZAFARMAND, M. H., HEUDE, B., FAROOQI, I. S., ALYASS, A., BEAUMONT, R. N., HAVE, C. T., RZEHAK, P., BILBAO, J. R., SCHNURR, T. M., BARROSO, I., BONNELYKKE, K., BEILIN, L. J., CARSTENSEN, L., CHARLES, M. A., CHAWES, B., CLEMENT, K., CLOSA-MONASTEROLO, R., CUSTOVIC, A., ERIKSSON, J. G., ESCRIBANO, J., GROEN-BLOKHUIS, M., GROTE, V., GRUSZFELD, D., HAKONARSON, H., HANSEN, T., HATTERSLEY, A. T., HOLLENSTED, M., HOTTENGA, J. J., HYPPONEN, E., JOHANSSON, S., JORO, R., KAHONEN, M., KARHUNEN, V., KIESS, W., KNIGHT, B. A., KOLETZKO, B., KUHNAPFEL, A., LANDGRAF, K., LANGHENDRIES, J. P., LEHTIMAKI, T., LEINONEN, J. T., LI, A., LINDI, V., LOWRY, E., BUSTAMANTE, M., MEDINA-GOMEZ, C., MELBYE, M., MICHAELSEN, K. F., MORGEN, C. S., MORI, T. A., NIELSEN, T. R. H., NIINIKOSKI, H., OLDEHINKEL, A. J., PAHKALA, K., PANOUTSOPOULOU, K., PEDERSEN, O., PENNELL, C. E., POWER, C., REIJNEVELD, S. A., RIVADENEIRA, F., SIMPSON, A., SLY, P. D., STOKHOLM, J., TEO, K. K., THIERING, E., TIMPSON, N. J.,

UITTERLINDEN, A. G., VAN BEIJSTERVELDT, C. E. M., VAN SCHAIK, B. D. C., VAUDEL, M., VERDUCI, E., et al. 2020. Novel loci for childhood body mass index and shared heritability with adult cardiometabolic traits. *PLoS Genet,* 16**,** e1008718.

VON KORFF, M., KATON, W., LIN, E. H., SIMON, G., CIECHANOWSKI, P., LUDMAN, E., OLIVER, M., RUTTER, C. & YOUNG, B. 2005. Work disability among individuals with diabetes. *Diabetes Care,* 28**,** 1326-32.

VOSS, R., CULLEN, P., SCHULTE, H. & ASSMANN, G. 2002. Prediction of risk of coronary events in middle-aged men in the Prospective Cardiovascular Münster Study (PROCAM) using neural networks. *Int J Epidemiol,* 31**,** 1253-62; discussion 1262-64.

WAHLBECK, K., FORSEN, T., OSMOND, C., BARKER, D. J. & ERIKSSON, J. G. 2001. Association of schizophrenia with low maternal body mass index, small size at birth, and thinness during childhood. *Arch Gen Psychiatry,* 58**,** 48-52.

WALLACE, T. M., LEVY, J. C. & MATTHEWS, D. R. 2004. Use and abuse of HOMA modeling. *Diabetes Care,* 27**,** 1487-95.

WALTERS, R. K., POLIMANTI, R., JOHNSON, E. C., MCCLINTICK, J. N., ADAMS, M. J., ADKINS, A. E., ALIEV, F., BACANU, S. A., BATZLER, A., BERTELSEN, S., BIERNACKA, J. M., BIGDELI, T. B., CHEN, L. S., CLARKE, T. K., CHOU, Y. L., DEGENHARDT, F., DOCHERTY, A. R., EDWARDS, A. C., FONTANILLAS, P., FOO, J. C., FOX, L., FRANK, J., GIEGLING, I., GORDON, S., HACK, L. M., HARTMANN, A. M., HARTZ, S. M., HEILMANN-HEIMBACH, S., HERMS, S., HODGKINSON, C., HOFFMANN, P., JAN HOTTENGA, J., KENNEDY, M. A., ALANNE-KINNUNEN, M., KONTE, B., LAHTI, J., LAHTI-PULKKINEN, M., LAI, D., LIGTHART, L., LOUKOLA, A., MAHER, B. S., MBAREK, H., MCINTOSH, A. M., MCQUEEN, M. B., MEYERS, J. L., MILANESCHI, Y., PALVIAINEN, T., PEARSON, J. F., PETERSON, R. E., RIPATTI, S., RYU, E., SACCONE, N. L., SALVATORE, J. E., SANCHEZ-ROIGE, S., SCHWANDT, M., SHERVA, R., STREIT, F., STROHMAIER, J., THOMAS, N., WANG, J. C., WEBB, B. T., WEDOW, R., WETHERILL, L., WILLS, A. G., ANDME RESEARCH, T., BOARDMAN, J. D., CHEN, D., CHOI, D. S., COPELAND, W. E., CULVERHOUSE, R. C., DAHMEN, N., DEGENHARDT, L., DOMINGUE, B. W., ELSON, S. L., FRYE, M. A., GABEL, W., HAYWARD, C., ISING, M., KEYES, M., KIEFER, F., KRAMER, J., KUPERMAN, S., LUCAE, S., LYNSKEY, M. T., MAIER, W., MANN, K., MANNISTO, S., MULLER-MYHSOK, B., MURRAY, A. D., NURNBERGER, J. I., PALOTIE, A., PREUSS, U., RAIKKONEN, K., REYNOLDS, M. D., RIDINGER, M., SCHERBAUM, N., SCHUCKIT, M. A., SOYKA, M., TREUTLEIN,

J., WITT, S., et al. 2018. Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nat Neurosci,* 21**,** 1656-1669.

WANG, X., BAO, W., LIU, J., OUYANG, Y. Y., WANG, D., RONG, S., XIAO, X., SHAN, Z. L., ZHANG, Y., YAO, P. & LIU, L. G. 2013. Inflammatory markers and risk of type 2 diabetes: a systematic review and meta-analysis. *Diabetes Care,* 36**,** 166-75.

WARD, M. & DRUSS, B. 2015. The epidemiology of diabetes in psychotic disorders. *Lancet Psychiatry,* 2**,** 431-451.

WATERALL, J. 2019. Health Matters: Preventing Cardiovascular Disease. *Health Matters* [Online].

WEBB, T. R., ERDMANN, J., STIRRUPS, K. E., STITZIEL, N. O., MASCA, N. G., JANSEN, H., KANONI, S., NELSON, C. P., FERRARIO, P. G., KONIG, I. R., EICHER, J. D., JOHNSON, A. D., HAMBY, S. E., BETSHOLTZ, C., RUUSALEPP, A., FRANZEN, O., SCHADT, E. E., BJORKEGREN, J. L., WEEKE, P. E., AUER, P. L., SCHICK, U. M., LU, Y., ZHANG, H., DUBE, M. P., GOEL, A., FARRALL, M., PELOSO, G. M., WON, H. H., DO, R., VAN IPEREN, E., KRUPPA, J., MAHAJAN, A., SCOTT, R. A., WILLENBORG, C., BRAUND, P. S., VAN CAPELLEVEEN, J. C., DONEY, A. S., DONNELLY, L. A., ASSELTA, R., MERLINI, P. A., DUGA, S., MARZILIANO, N., DENNY, J. C., SHAFFER, C., EL-MOKHTARI, N. E., FRANKE, A., HEILMANN, S., HENGSTENBERG, C., HOFFMANN, P., HOLMEN, O. L., HVEEM, K., JANSSON, J. H., JOCKEL, K. H., KESSLER, T., KRIEBEL, J., LAUGWITZ, K. L., MAROULI, E., MARTINELLI, N., MCCARTHY, M. I., VAN ZUYDAM, N. R., MEISINGER, C., ESKO, T., MIHAILOV, E., ESCHER, S. A., ALVER, M., MOEBUS, S., MORRIS, A. D., VIRTAMO, J., NIKPAY, M., OLIVIERI, O., PROVOST, S., ALQARAWI, A., ROBERTSON, N. R., AKINSANSYA, K. O., REILLY, D. F., VOGT, T. F., YIN, W., ASSELBERGS, F. W., KOOPERBERG, C., JACKSON, R. D., STAHL, E., MULLER-NURASYID, M., STRAUCH, K., VARGA, T. V., WALDENBERGER, M., WELLCOME TRUST CASE CONTROL, C., ZENG, L., CHOWDHURY, R., SALOMAA, V., FORD, I., JUKEMA, J. W., AMOUYEL, P., KONTTO, J., INVESTIGATORS, M., NORDESTGAARD, B. G., FERRIERES, J., SALEHEEN, D., SATTAR, N., SURENDRAN, P., WAGNER, A., et al. 2017. Systematic Evaluation of Pleiotropy Identifies 6 Further Loci Associated With Coronary Artery Disease. *J Am Coll Cardiol,* 69**,** 823-836.

WEINTRAUB, W. S., DANIELS, S. R., BURKE, L. E., FRANKLIN, B. A., GOFF, D. C., JR., HAYMAN, L. L., LLOYD-JONES, D., PANDEY, D. K., SANCHEZ, E. J., SCHRAM, A. P., WHITSEL, L. P., AMERICAN HEART ASSOCIATION ADVOCACY

COORDINATING, C., COUNCIL ON CARDIOVASCULAR DISEASE IN THE, Y., COUNCIL ON THE KIDNEY IN CARDIOVASCULAR, D., COUNCIL ON, E., PREVENTION, COUNCIL ON CARDIOVASCULAR, N., COUNCIL ON, A., THROMBOSIS, VASCULAR, B., COUNCIL ON CLINICAL, C. & STROKE, C. 2011. Value of primordial and primary prevention for cardiovascular disease: a policy statement from the American Heart Association. *Circulation,* 124**,** 967-90.

WEISER, M., KNOBLER, H., LUBIN, G., NAHON, D., KRAVITZ, E., CASPI, A., NOY, S., KNOBLER, H. Y. & DAVIDSON, M. 2004. Body mass index and future schizophrenia in Israeli male adolescents. *J Clin Psychiatry,* 65**,** 1546-9.

WELCH, E., JANGMO, A., THORNTON, L. M., NORRING, C., VON HAUSSWOLFF-JUHLIN, Y., HERMAN, B. K., PAWASKAR, M., LARSSON, H. & BULIK, C. M. 2016. Treatment-seeking patients with binge-eating disorder in the Swedish national registers: clinical course and psychiatric comorbidity. *BMC Psychiatry,* 16**,** 163.

WEN, J., HAO, J., LIANG, Y., LI, S., CAO, K., LU, X., LU, X. & WANG, N. 2017. A non-invasive risk score for predicting incident diabetes among rural Chinese people: A village-based cohort study. *PLoS One,* 12**,** e0186172.

WENDER, P. H., ROSENTHAL, D., KETY, S. S., SCHULSINGER, F. & WELNER, J. 1974. Crossfostering. A research strategy for clarifying the role of genetic and experiential factors in the etiology of schizophrenia. *Arch Gen Psychiatry,* 30**,** 121-8.

WESTMAN, J., ERIKSSON, S. V., GISSLER, M., HALLGREN, J., PRIETO, M. L., BOBO, W. V., FRYE, M. A., ERLINGE, D., ALFREDSSON, L. & OSBY, U. 2018. Increased cardiovascular mortality in people with schizophrenia: a 24-year national register study. *Epidemiol Psychiatr Sci,* 27**,** 519-527.

WHEELER, E., LEONG, A., LIU, C. T., HIVERT, M. F., STRAWBRIDGE, R. J., PODMORE, C., LI, M., YAO, J., SIM, X., HONG, J., CHU, A. Y., ZHANG, W., WANG, X., CHEN, P., MARUTHUR, N. M., PORNEALA, B. C., SHARP, S. J., JIA, Y., KABAGAMBE, E. K., CHANG, L. C., CHEN, W. M., ELKS, C. E., EVANS, D. S., FAN, Q., GIULIANINI, F., GO, M. J., HOTTENGA, J. J., HU, Y., JACKSON, A. U., KANONI, S., KIM, Y. J., KLEBER, M. E., LADENVALL, C., LECOEUR, C., LIM, S. H., LU, Y., MAHAJAN, A., MARZI, C., NALLS, M. A., NAVARRO, P., NOLTE, I. M., ROSE, L. M., RYBIN, D. V., SANNA, S., SHI, Y., STRAM, D. O., TAKEUCHI, F., TAN, S. P., VAN DER MOST, P. J., VAN VLIET-OSTAPTCHOUK, J. V., WONG, A., YENGO, L., ZHAO, W., GOEL, A., MARTINEZ LARRAD, M. T., RADKE, D., SALO, P., TANAKA, T., VAN IPEREN, E. P. A., ABECASIS, G., AFAQ, S., ALIZADEH, B. Z., BERTONI, A. G., BONNEFOND, A., BOTTCHER, Y., BOTTINGER, E. P., CAMPBELL, H., CARLSON, O. D., CHEN, C. H.,

CHO, Y. S., GARVEY, W. T., GIEGER, C., GOODARZI, M. O., GRALLERT, H., HAMSTEN, A., HARTMAN, C. A., HERDER, C., HSIUNG, C. A., HUANG, J., IGASE, M., ISONO, M., KATSUYA, T., KHOR, C. C., KIESS, W., KOHARA, K., KOVACS, P., LEE, J., LEE, W. J., LEHNE, B., LI, H., LIU, J., LOBBENS, S., LUAN, J., LYSSENKO, V., MEITINGER, T., MIKI, T., MILJKOVIC, I., MOON, S., MULAS, A., MULLER, G., et al. 2017. Impact of common genetic determinants of Hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: A transethnic genome-wide meta-analysis. *PLoS Med,* 14**,** e1002383.

WHITE, I. R., ROYSTON, P. & WOOD, A. M. 2011. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med,* 30**,** 377-99.

WICKRAMASINGHE, C. D., AYERS, C. R., DAS, S., DE LEMOS, J. A., WILLIS, B. L. & BERRY, J. D. 2014. Prediction of 30-year risk for cardiovascular mortality by fitness and risk factor levels: the Cooper Center Longitudinal Study. *Circ Cardiovasc Qual Outcomes,* 7**,** 597-602.

WIEBE, N., YE, F., CRUMLEY, E. T., BELLO, A., STENVINKEL, P. & TONELLI, M. 2021. Temporal Associations Among Body Mass Index, Fasting Insulin, and Systemic Inflammation: A Systematic Review and Meta-analysis. *JAMA Netw Open,* 4**,** e211263.

WILKINSON, J., ARNOLD, K. F., MURRAY, E. J., VAN SMEDEN, M., CARR, K., SIPPY, R., DE KAMPS, M., BEAM, A., KONIGORSKI, S., LIPPERT, C., GILTHORPE, M. S. & TENNANT, P. W. G. 2020. Time to reality check the promises of machine learning-powered precision medicine. *Lancet Digit Health,* 2**,** e677-e680.

WILLER, C. J., SCHMIDT, E. M., SENGUPTA, S., PELOSO, G. M., GUSTAFSSON, S., KANONI, S., GANNA, A., CHEN, J., BUCHKOVICH, M. L., MORA, S., BECKMANN, J. S., BRAGG-GRESHAM, J. L., CHANG, H. Y., DEMIRKAN, A., DEN HERTOG, H. M., DO, R., DONNELLY, L. A., EHRET, G. B., ESKO, T., FEITOSA, M. F., FERREIRA, T., FISCHER, K., FONTANILLAS, P., FRASER, R. M., FREITAG, D. F., GURDASANI, D., HEIKKILA, K., HYPPONEN, E., ISAACS, A., JACKSON, A. U., JOHANSSON, A., JOHNSON, T., KAAKINEN, M., KETTUNEN, J., KLEBER, M. E., LI, X., LUAN, J., LYYTIKAINEN, L. P., MAGNUSSON, P. K. E., MANGINO, M., MIHAILOV, E., MONTASSER, M. E., MULLER-NURASYID, M., NOLTE, I. M., O'CONNELL, J. R., PALMER, C. D., PEROLA, M., PETERSEN, A. K., SANNA, S., SAXENA, R., SERVICE, S. K., SHAH, S., SHUNGIN, D., SIDORE, C., SONG, C., STRAWBRIDGE, R. J., SURAKKA, I., TANAKA, T., TESLOVICH, T. M., THORLEIFSSON, G., VAN DEN HERIK, E. G., VOIGHT, B. F., VOLCIK, K. A., WAITE, L. L., WONG, A., WU, Y., ZHANG, W., ABSHER, D., ASIKI, G., BARROSO, I., BEEN, L. F., BOLTON, J. L.,

BONNYCASTLE, L. L., BRAMBILLA, P., BURNETT, M. S., CESANA, G., DIMITRIOU, M., DONEY, A. S. F., DORING, A., ELLIOTT, P., EPSTEIN, S. E., INGI EYJOLFSSON, G., GIGANTE, B., GOODARZI, M. O., GRALLERT, H., GRAVITO, M. L., GROVES, C. J., HALLMANS, G., HARTIKAINEN, A. L., HAYWARD, C., HERNANDEZ, D., HICKS, A. A., HOLM, H., HUNG, Y. J., ILLIG, T., JONES, M. R., KALEEBU, P., KASTELEIN, J. J. P., KHAW, K. T., KIM, E., et al. 2013. Discovery and refinement of loci associated with lipid levels. *Nat Genet,* 45**,** 1274-1283.

WILMOT, E. & IDRIS, I. 2014. Early onset type 2 diabetes: risk factors, clinical impact and management. *Ther Adv Chronic Dis,* 5**,** 234-44.

WILSON, P. W., D'AGOSTINO, R. B., PARISE, H., SULLIVAN, L. & MEIGS, J. B. 2005. Metabolic syndrome as a precursor of cardiovascular disease and type 2 diabetes mellitus. *Circulation,* 112**,** 3066-72.

WINDSOR-SHELLARD, B., HORTON, M., SCANLON, S., MANDERS, S., 2020. Adult Smoking Habits in the UK: 2019.

WOLFERS, T., DOAN, N. T., KAUFMANN, T., ALNAES, D., MOBERGET, T., AGARTZ, I., BUITELAAR, J. K., UELAND, T., MELLE, I., FRANKE, B., ANDREASSEN, O. A., BECKMANN, C. F., WESTLYE, L. T. & MARQUAND, A. F. 2018. Mapping the Heterogeneous Phenotype of Schizophrenia and Bipolar Disorder Using Normative Models. *JAMA Psychiatry,* 75**,** 1146-1155.

WOLFF, R. F., MOONS, K. G. M., RILEY, R. D., WHITING, P. F., WESTWOOD, M., COLLINS, G. S., REITSMA, J. B., KLEIJNEN, J., MALLETT, S. & GROUPDAGGER, P. 2019. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. *Ann Intern Med,* 170**,** 51-58.

WONG, C. K., SIU, S. C., WAN, E. Y., JIAO, F. F., YU, E. Y., FUNG, C. S., WONG, K. W., LEUNG, A. Y. & LAM, C. L. 2016. Simple non-laboratory- and laboratory-based risk assessment algorithms and nomogram for detecting undiagnosed diabetes mellitus. *J Diabetes,* 8**,** 414-21.

WOO, H. J., YU, C., KUMAR, K. & REIFMAN, J. 2017. Large-scale interaction effects reveal missing heritability in schizophrenia, bipolar disorder and posttraumatic stress disorder. *Transl Psychiatry,* 7**,** e1089.

WOOTTON, R. E., RICHMOND, R. C., STUIJFZAND, B. G., LAWN, R. B., SALLIS, H. M., TAYLOR, G. M. J., HEMANI, G., JONES, H. J., ZAMMIT, S., DAVEY SMITH, G. & MUNAFO, M. R. 2020. Evidence for causal effects of lifetime smoking on risk for depression and schizophrenia: a Mendelian randomisation study. *Psychol Med,* 50**,** 2435-2443.

WORLD HEALTH ORGANIZATION 2018. The top 10 causes of death. World Health Organization, Geneva.

WORLD HEALTH ORGANIZATION, I. D. F. 2006. Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia. http://www.who.int/diabetes/publications/diagnosis_diabetes2006/en/.

WU, C. Y., CHANG, C. K., ROBSON, D., JACKSON, R., CHEN, S. J., HAYES, R. D. & STEWART, R. 2013. Evaluation of smoking status identification using electronic health records and open-text information in a large mental health case register. *PLoS One,* 8**,** e74262.

WU, S. H., NEALE, M. C., ACTON, A. J., JR., CONSIDINE, R. V., KRASNOW, R. E., REED, T. & DAI, J. 2014. Genetic and environmental influences on the prospective correlation between systemic inflammation and coronary heart disease death in male twins. *Arterioscler Thromb Vasc Biol,* 34**,** 2168-74.

WYNANTS, L., VAN CALSTER, B., COLLINS, G. S., RILEY, R. D., HEINZE, G., SCHUIT, E., BONTEN, M. M. J., DAHLY, D. L., DAMEN, J. A. A., DEBRAY, T. P. A., DE JONG, V. M. T., DE VOS, M., DHIMAN, P., HALLER, M. C., HARHAY, M. O., HENCKAERTS, L., HEUS, P., KAMMER, M., KREUZBERGER, N., LOHMANN, A., LUIJKEN, K., MA, J., MARTIN, G. P., MCLERNON, D. J., ANDAUR, C. L., REITSMA, J. B., SERGEANT, J. C., SHI, C., SKOETZ, N., SMITS, L. J. M., SNELL, K. I. E., SPERRIN, M., SPIJKER, R., STEYERBERG, E. W., TAKADA, T., TZOULAKI, I., VAN KUIJK, S. M. J., VAN BUSSEL, B., VAN ROYEN, F. S., VERBAKEL, J. Y., WALLISCH, C., WILKINSON, J., WOLFF, R., HOOFT, L., MOONS, K. G. M. & VAN SMEDEN, M. 2020. Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. *BMJ,* 369**,** m1328.

XI, L., XIAO, C., BANDSMA, R. H., NAPLES, M., ADELI, K. & LEWIS, G. F. 2011. C-reactive protein impairs hepatic insulin sensitivity and insulin signaling in rats: role of mitogen-activated protein kinases. *Hepatology,* 53**,** 127-35.

YAJNIK, C. S. 2010. Fetal programming of diabetes: still so much to learn! *Diabetes Care,* 33**,** 1146-8.

YANG, C. & WINKELMAN, J. W. 2006. Clinical significance of sleep EEG abnormalities in chronic schizophrenia. *Schizophr Res,* 82**,** 251-60.

YANG, L., CAO, C., KANTOR, E. D., NGUYEN, L. H., ZHENG, X., PARK, Y., GIOVANNUCCI, E. L., MATTHEWS, C. E., COLDITZ, G. A. & CAO, Y. 2019. Trends in Sedentary Behavior Among the US Population, 2001-2016. *JAMA,* 321**,** 1587-1597.

YATSUYA, H., ISO, H., LI, Y., YAMAGISHI, K., KOKUBO, Y., SAITO, I., SAWADA, N., INOUE, M. & TSUGANE, S. 2016. Development of a Risk Equation for the Incidence of Coronary Artery Disease and Ischemic Stroke for Middle-Aged Japanese- Japan Public Health Center-Based Prospective Study. *Circ J,* 80**,** 1386-95.

YATSUYA, H., ISO, H., YAMAGISHI, K., KOKUBO, Y., SAITO, I., SUZUKI, K., SAWADA, N., INOUE, M. & TSUGANE, S. 2013. Development of a point-based prediction model for the incidence of total stroke: Japan public health center study. *Stroke,* 44**,** 1295-302.

YILMAZ, Z., GOTTFREDSON, N. C., ZERWAS, S. C., BULIK, C. M. & MICALI, N. 2019. Developmental Premorbid Body Mass Index Trajectories of Adolescents With Eating Disorders in a Longitudinal Population Cohort. *J Am Acad Child Adolesc Psychiatry,* 58**,** 191-199.

YOUNIS, A., YOUNIS, A., TZUR, B., PELED, Y., SHLOMO, N., GOLDENBERG, I., FISMAN, E. Z., TENENBAUM, A. & KLEMPFNER, R. 2016. Metabolic syndrome is independently associated with increased 20-year mortality in patients with stable coronary artery disease. *Cardiovasc Diabetol,* 15**,** 149.

YUAN, S. & LARSSON, S. C. 2020. An atlas on risk factors for type 2 diabetes: a wide-angled Mendelian randomisation study. *Diabetologia,* 63**,** 2359-2371.

YUNG, A. R., YUEN, H. P., MCGORRY, P. D., PHILLIPS, L. J., KELLY, D., DELL'OLIO, M., FRANCEY, S. M., COSGRAVE, E. M., KILLACKEY, E., STANFORD, C., GODFREY, K. & BUCKBY, J. 2005. Mapping the onset of psychosis: the Comprehensive Assessment of At-Risk Mental States. *Aust N Z J Psychiatry,* 39**,** 964-71.

ZAKHARYAN, R. & BOYAJYAN, A. 2014. Brain-derived neurotrophic factor blood levels are decreased in schizophrenia patients and associate with rs6265 genotypes. *Clin Biochem,* 47**,** 1052-5.

ZAMMIT, S., KOUNALI, D., CANNON, M., DAVID, A. S., GUNNELL, D., HERON, J., JONES, P. B., LEWIS, S., SULLIVAN, S., WOLKE, D. & LEWIS, G. 2013. Psychotic Experiences and Psychotic Disorders at Age 18 in Relation to Psychotic Experiences at Age 12 in a Longitudinal Population-Based Cohort Study. *Am J Psychiatry,* 170**,** 742-750.

ZAMMIT, S., RASMUSSEN, F., FARAHMAND, B., GUNNELL, D., LEWIS, G., TYNELIUS, P. & BROBERT, G. P. 2007. Height and body mass index in young adulthood and risk of schizophrenia: a longitudinal study of 1 347 520 Swedish men. *Acta Psychiatr Scand,* 116**,** 378-85.

ZDRAVKOVIC, S., WIENKE, A., PEDERSEN, N. L., MARENBERG, M. E., YASHIN, A. I. & DE FAIRE, U. 2002. Heritability of death from coronary heart disease: a 36-year follow-up of 20 966 Swedish twins. *J Intern Med,* 252**,** 247-54.

ZHANG, C., QIN, Y. Y., CHEN, Q., JIANG, H., CHEN, X. Z., XU, C. L., MAO, P. J., HE, J. & ZHOU, Y. H. 2014. Alcohol intake and risk of stroke: a dose-response meta-analysis of prospective studies. *Int J Cardiol,* 174**,** 669-77.

ZHANG, Q., WU, K. H., HE, J. Y., ZENG, Y., GREENBAUM, J., XIA, X., LIU, H. M., LV, W. Q., LIN, X., ZHANG, W. D., XI, Y. L., SHI, X. Z., SUN, C. Q. & DENG, H. W. 2017. Novel Common Variants Associated with Obesity and Type 2 Diabetes Detected Using a cFDR Method. *Sci Rep,* 7**,** 16397.

ZHANG, T., XU, J., LI, S., BAZZANO, L. A., HE, J., WHELTON, P. K. & CHEN, W. 2019. Trajectories of childhood BMI and adult diabetes: the Bogalusa Heart Study. *Diabetologia,* 62**,** 70-77.

ZHANG, X., ZHANG, G., ZHANG, H., KARIN, M., BAI, H. & CAI, D. 2008. Hypothalamic IKKbeta/NF-kappaB and ER stress link overnutrition to energy imbalance and obesity. *Cell,* 135**,** 61-73.

ZHANG, X. F., ATTIA, J., D'ESTE, C., YU, X. H. & WU, X. G. 2005. A risk score predicted coronary heart disease and stroke in a Chinese cohort. *J Clin Epidemiol,* 58**,** 951-8.

ZHEUTLIN, A. B., DENNIS, J., KARLSSON LINNER, R., MOSCATI, A., RESTREPO, N., STRAUB, P., RUDERFER, D., CASTRO, V. M., CHEN, C. Y., GE, T., HUCKINS, L. M., CHARNEY, A., KIRCHNER, H. L., STAHL, E. A., CHABRIS, C. F., DAVIS, L. K. & SMOLLER, J. W. 2019. Penetrance and Pleiotropy of Polygenic Risk Scores for Schizophrenia in 106,160 Patients Across Four Health Care Systems. *Am J Psychiatry,* 176**,** 846-855.

ZHU, W., BIJUR, G. N., STYLES, N. A. & LI, X. 2004. Regulation of FOXO3a by brain-derived neurotrophic factor in differentiated human SH-SY5Y neuroblastoma cells. *Brain Res Mol Brain Res,* 126**,** 45-56.

ZIAADDINI, H., KHERADMAND, A. & VAHABI, M. 2009. Prevalence of cigarette smoking in schizophrenic patients compared to other hospital admitted psychiatric patients. *Addict Health,* 1**,** 38-42.

ZIMMET, P., MAGLIANO, D., MATSUZAWA, Y., ALBERTI, G. & SHAW, J. 2005. The metabolic syndrome: a global public health problem and a new definition. *J Atheroscler Thromb,* 12**,** 295-300.

ZOMER, E., OSBORN, D., NAZARETH, I., BLACKBURN, R., BURTON, A., HARDOON, S., HOLT, R. I. G., KING, M., MARSTON, L., MORRIS, S., OMAR, R., PETERSEN, I., WALTERS, K. & HUNTER, R. M. 2017. Effectiveness and cost-effectiveness of a cardiovascular risk prediction algorithm for people with severe mental illness (PRIMROSE). *BMJ Open,* 7**,** e018181.

The Relationship between Cardiometabolic Disorders and Schizophrenia: From Early-Life Origins to the Development of a Cardiometabolic Risk Prediction Algorithm For Young People with Psychosis

Appendices

Appendix B

Appendix B Tables

Appendix B Table 1: Odds Ratios for Associations of Fasting Insulin and Body Mass Index Trajectories with Metabolic Syndrome at Age 24

Trajectory	**Sample**	**Odds Ratio (95% C.I.)**		***p*-value**
		Unadjusted	**Adjusted for sex, ethnicity, social class, SDQ (7y), cumulative smoking, physical activity, alcohol and substance use, sleep and calorie intake**	
Fasting Insulin				
Class 1 – 'Stable Average'	4,939	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Minor Increase'	693	5.14 (3.01-8.09)	4.24 (2.34-8.21)	<0.001
Class 3 – 'Persistently High'	158	10.51 (4.82-22.18)	9.21 (3.77-20.15)	<0.001
BMI				
Class 1 – 'Stable Average'	8,383	1.00 [reference]	1.00 [reference]	-
Class 2 – 'Gradually Decreasing'	949	0.91 (0.57-1.48)	0.90 (0.55-1.46)	0.684
Class 3 – 'Puberty Onset – Minor Increase'	668	6.02 (4.32-8.38)	5.64 (3.31-9.01)	<0.001
Class 4 – 'Puberty Onset – Major Increase'	174	7.80 (3.67-13.54)	6.91 (3.20-12.87)	<0.001
Class 5 – 'Persistently High	289	11.65 (7.45-15.45)	10.62 (5.89-19.13)	<0.001

Appendix B Table 2: Odds Ratios for Sex-Stratified Associations of Fasting Insulin Trajectories with Continuous Psychosis and Depression Outcomes at Age 24 Years

Trajectory	Sample	Beta Coefficient (95% C.I.)		*p*-value[a]
		Unadjusted	Adjusted for sex, ethnicity, social class, SDQ (7y), cumulative smoking, physical activity, alcohol and substance use, sleep and calorie intake, negative/depressive symptoms	
Depressive Symptom Score at Age 24 (Males)				
Class 1 – 'Stable Average'	2,319	0.00 [reference]	0.00 [reference]	-
Class 2 – 'Minor Increase'	278	0.03 (-0.05, 0.08)	0.01 (-0.09, 0.10)	>0.999
Class 3 – 'Persistently High'	66	0.10 (-0.12, 0.23)	0.03 (-0.04, 0.10)	>0.999
Depressive Symptom Score at Age 24 (Females)				
Class 1 – 'Stable Average'	2,620	0.00 [reference]	0.00 [reference]	-
Class 2 – 'Minor Increase'	415	0.05 (-0.04, 0.09)	0.00 (-0.05, 0.06)	>0.999
Class 3 – 'Persistently High'	92	0.06 (-0.08, 0.15)	0.02 (-0.09, 0.14)	>0.999
Negative Psychotic Symptom Score at Age 24 (Males)				
Class 1 – 'Stable Average'	2,319	0.00 [reference]	0.00 [reference]	-
Class 2 – 'Minor Increase'	278	0.11 (0.02, 0.24)	0.08 (0.02, 0.15)	0.036
Class 3 – 'Persistently High'	66	0.23 (0.08, 0.38)	0.12 (0.03, 0.21)	0.021
Negative Psychotic Symptom Score at Age 24 (Females)				
Class 1 – 'Stable Average'	2,620	0.00 [reference]	0.00 [reference]	-
Class 2 – 'Minor Increase'	415	0.08 (0.03, 0.14)	0.04 (0.00, 0.07)	0.253
Class 3 – 'Persistently High'	92	0.15 (-0.03, 0.32)	0.03 (-0.04, 0.10)	0.59

[a] *p*-values adjusted for multiple testing using Holm-Bonferroni method

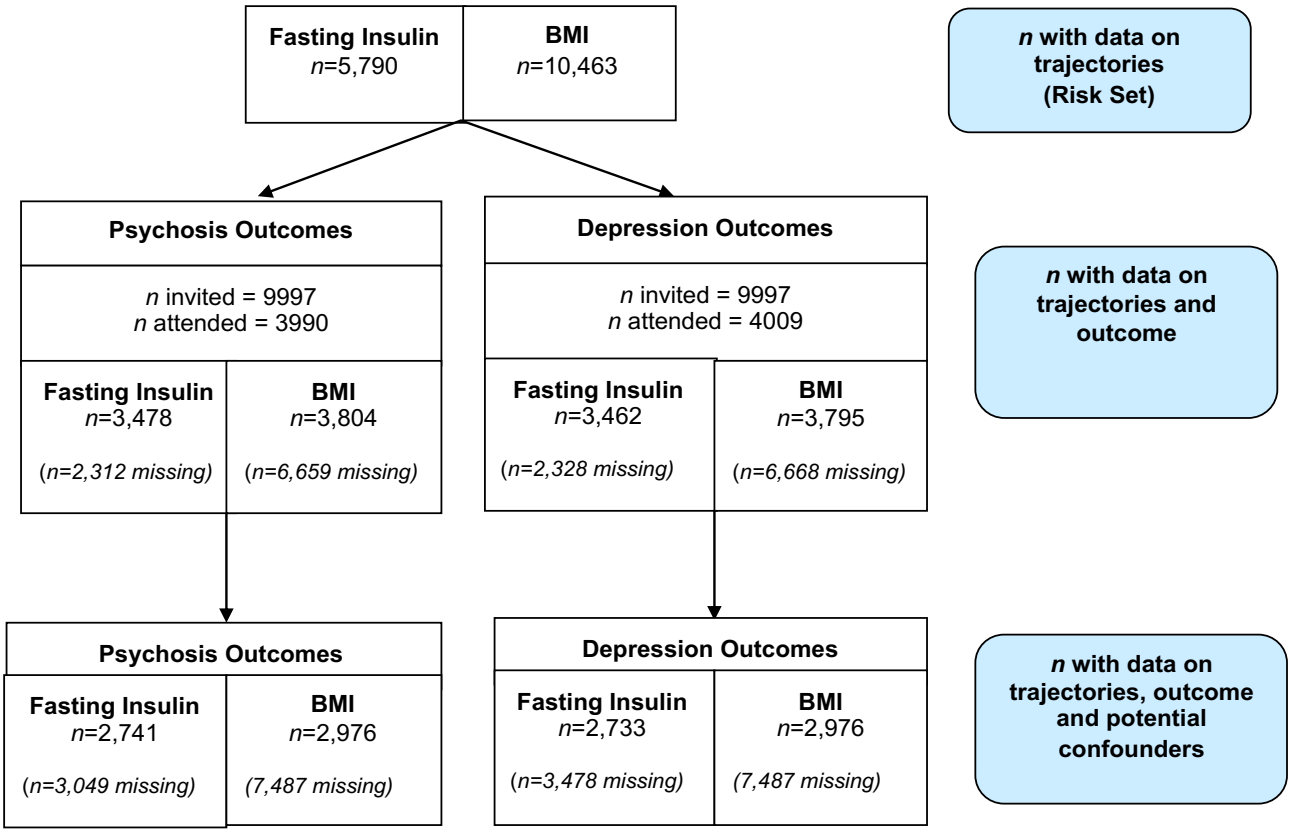
Appendix B Table 3: Odds Ratios for Sex-Stratified Associations of Body Mass Index Trajectories with Continuous Psychosis and Depression Outcomes at Age 24 Years

Trajectory	Sample	Beta Coefficient (95% C.I.)		*p*-value[a]
		Unadjusted	Adjusted for sex, ethnicity, social class, SDQ (7y), cumulative smoking, physical activity, alcohol and substance use, sleep and calorie intake, negative/depressive symptoms	
Depressive Symptom Score at Age 24 (Males)				
Class 1 – 'Stable Average'	4,164	0.00 [reference]	0.00 [reference]	-
Class 2 – 'Gradually Decreasing'	443	0.05 (-0.11, 0.21)	0.01 (-0.19, 0.22)	>0.999
Class 3 – 'Puberty Onset – Minor Increase'	311	-0.05 (-0.15, 0.20)	-0.02 (-0.10, 0.12)	>0.999
Class 4 – 'Puberty Onset – Major Increase'	105	0.08 (-0.16, 0.23)	0.03 (-0.08, 0.12)	>0.999
Class 5 – 'Persistently High	107	0.11 (-0.09, 0.21)	0.03 (-0.07, 0.11)	>0.999
Depressive Symptom Score at Age 24 (Females)				
Class 1 – 'Stable Average'	4,219	0.00 [reference]	0.00 [reference]	-
Class 2 – 'Gradually Decreasing'	506	0.02 (-0.06, 0.11)	0.01 (-0.10, 0.12)	>0.999
Class 3 – 'Puberty Onset – Minor Increase'	357	0.09 (0.02, 0.19)	0.06 (0.03, 0.09)	0.048
Class 4 – 'Puberty Onset – Major Increase'	184	0.15 (0.04, 0.26)	0.09 (0.04, 0.15)	0.046
Class 5 – 'Persistently High	67	0.18 (-0.08, 0.44)	0.03 (-0.07, 0.17)	>0.999
Negative Symptom Score at Age 24 (Males)				
Class 1 – 'Stable Average'	4,164	0.00 [reference]	0.00 [reference]	-
Class 2 – 'Gradually Decreasing'	443	0.10 (-0.12, 0.33)	0.05 (-0.12, 0.31)	>0.999
Class 3 – 'Puberty Onset – Minor Increase'	311	0.13 (0.08, 0.19)	0.04 (-0.03, 0.11)	>0.999
Class 4 – 'Puberty Onset – Major Increase'	105	0.21 (-0.12, 0.54)	0.12 (-0.13, 0.36)	>0.999
Class 5 – 'Persistently High	107	0.17 (-0.07, 0.42)	0.04 (-0.06, 0.15)	>0.999
Negative Symptom Score at Age 24 (Females)				
Class 1 – 'Stable Average'	4,219	0.00 [reference]	0.00 [reference]	-
Class 2 – 'Gradually Decreasing'	506	-0.03 (-0.10, 0.16)	0.01 (-0.09, 0.11)	>0.999
Class 3 – 'Puberty Onset – Minor Increase'	357	0.07 (0.00, 0.13)	0.03 (-0.07, 0.13)	>0.999
Class 4 – 'Puberty Onset – Major Increase'	184	0.16 (0.02, 0.30)	0.04 (-0.09, 0.18)	>0.999
Class 5 – 'Persistently High	67	0.11 (-0.05, 0.27)	0.04 (-0.09, 0.17)	>0.999

[a] *p*-values adjusted for multiple testing using Holm-Bonferroni method

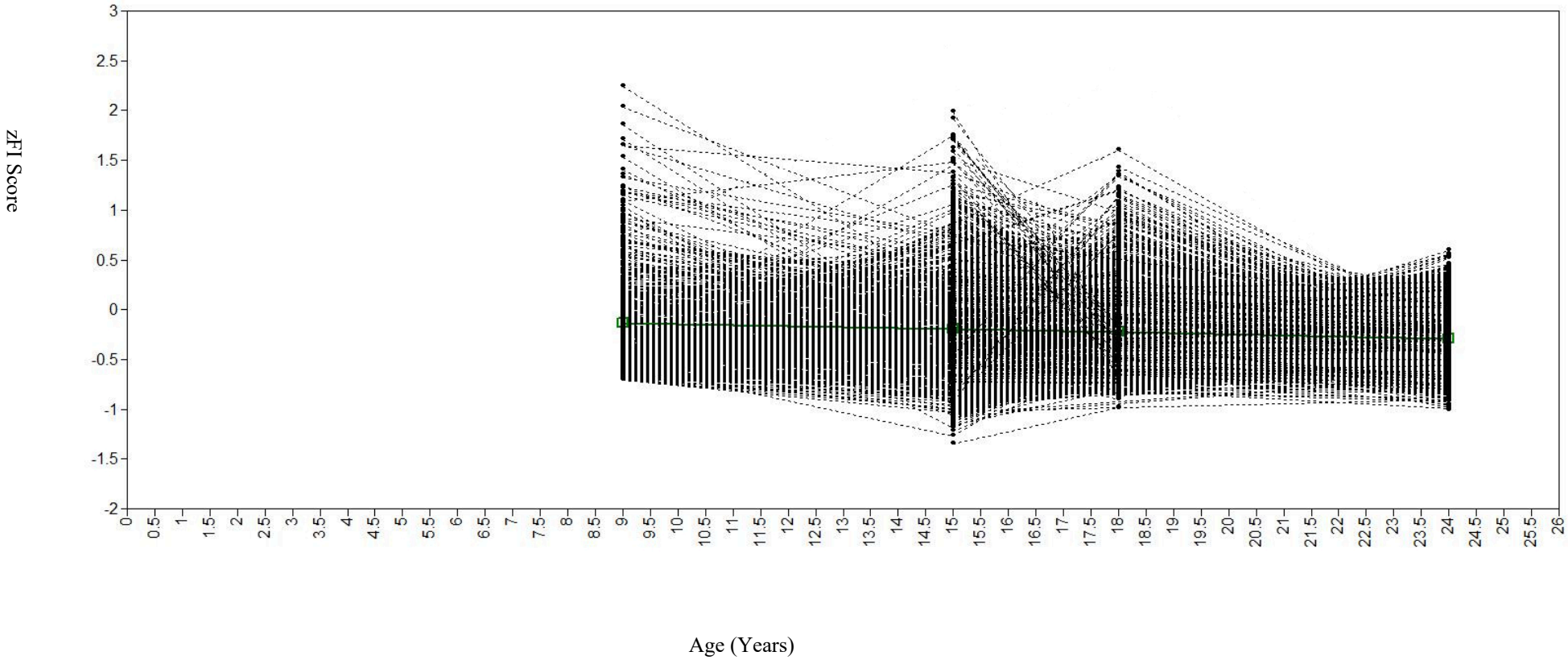
Appendix B Figures

Appendix B Figure 1: Flow Chart of Included Participants

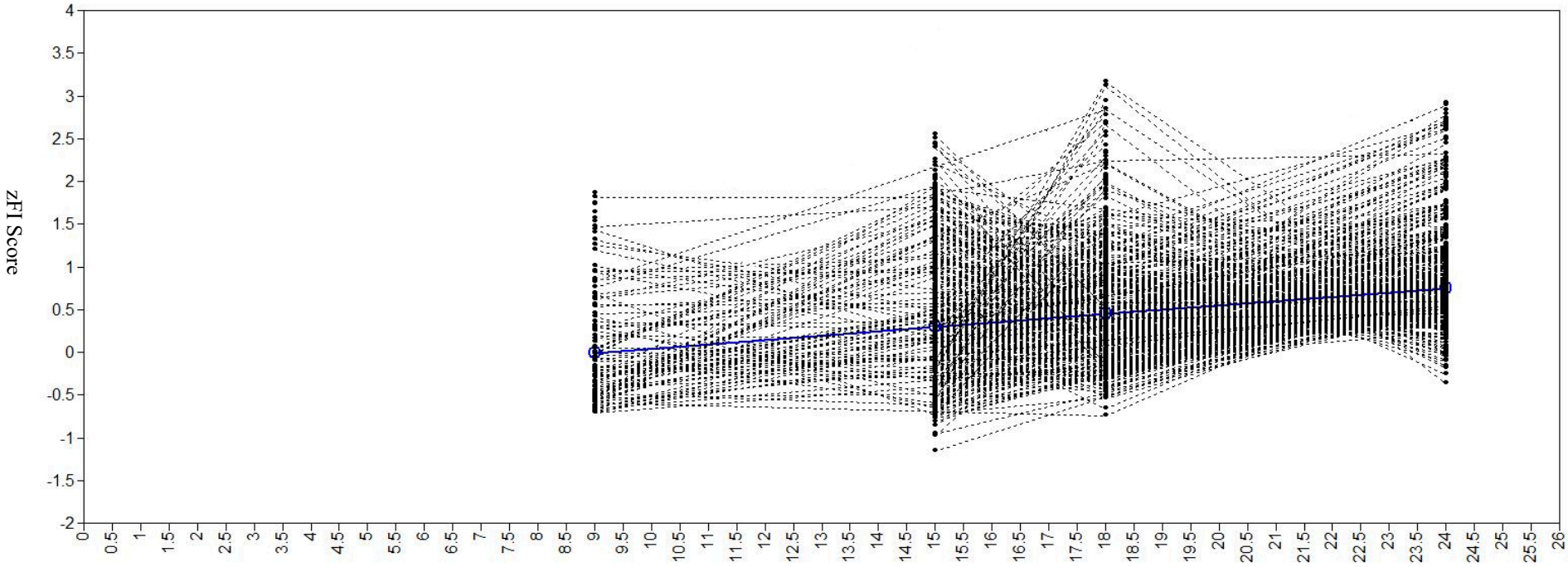


Appendix B Figure 2: Trajectory Means and Individual Values per Developmental Trajectory of Fasting Insulin

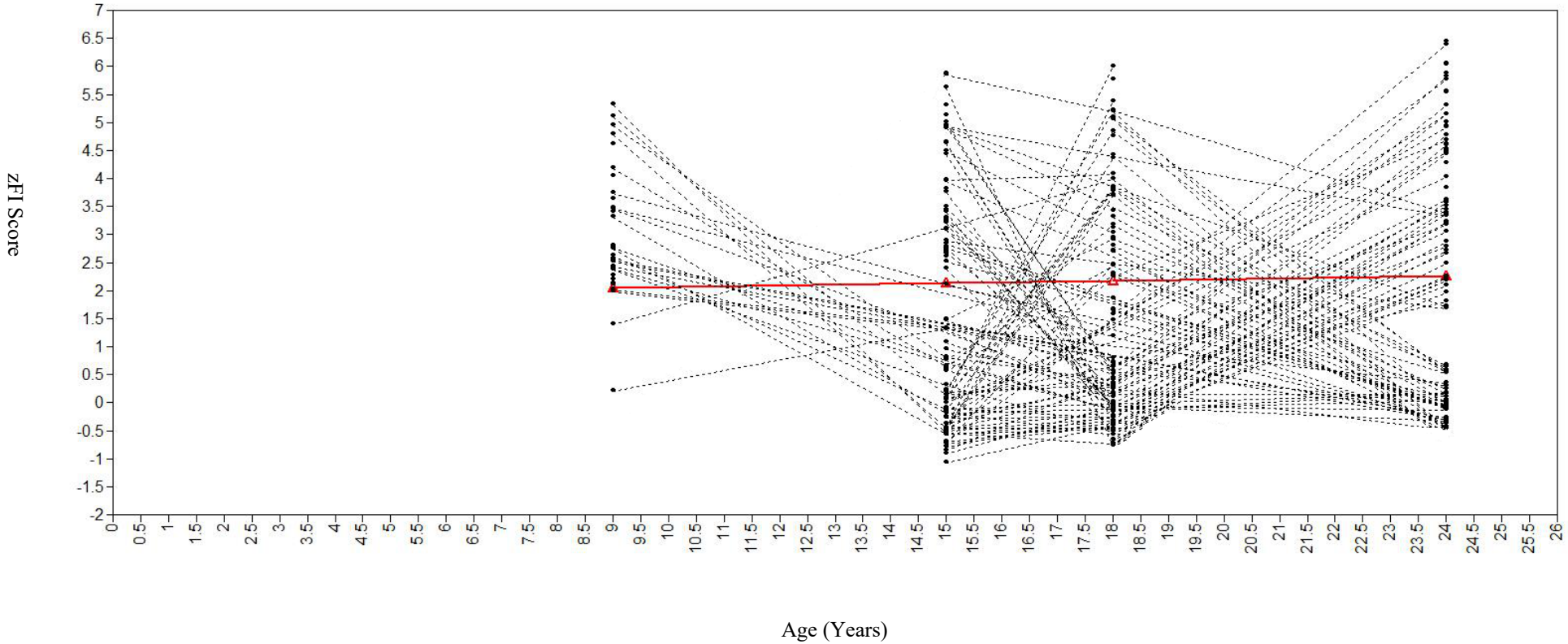
A. Class 1 – 'Stable Average'



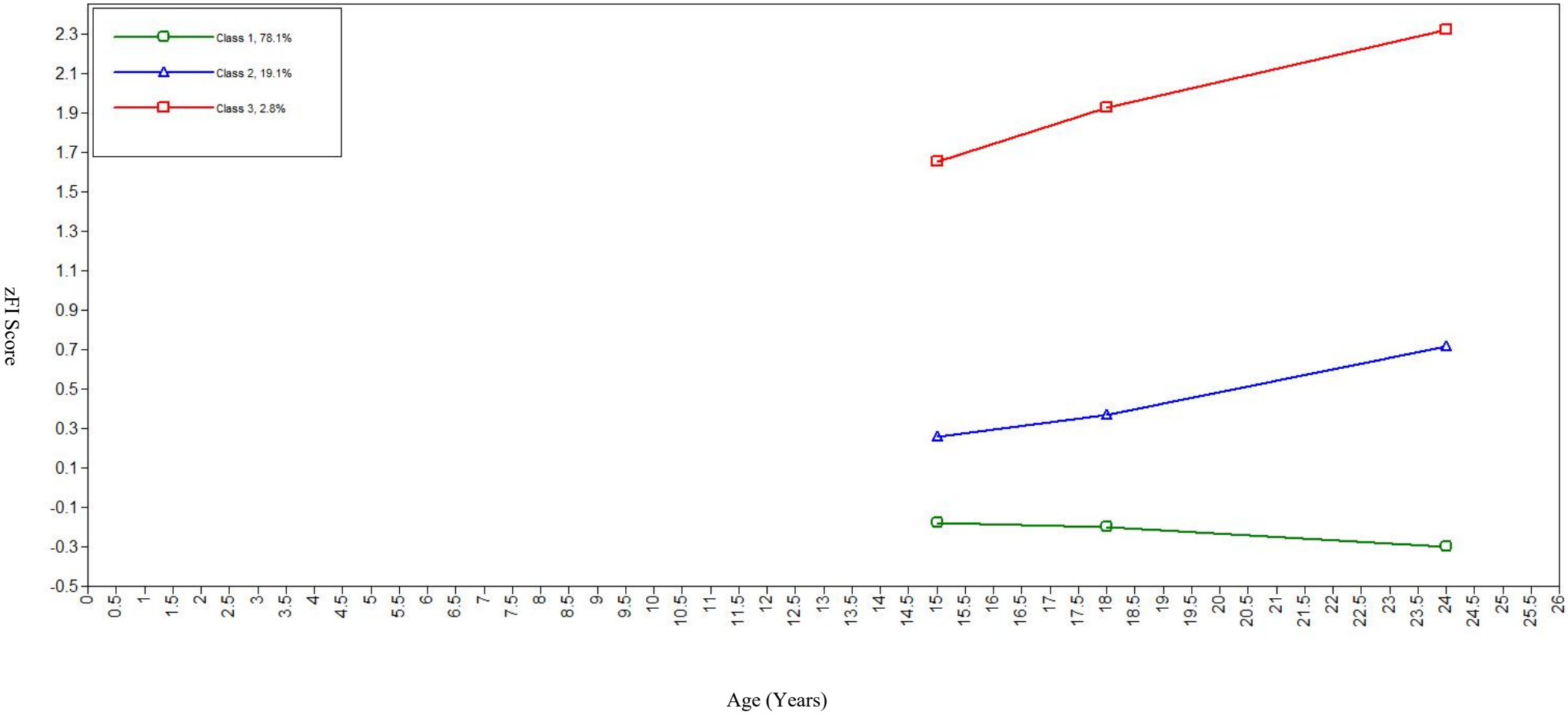
B. Class 2 – Minor Increase



C. Class 3 – Persistently High

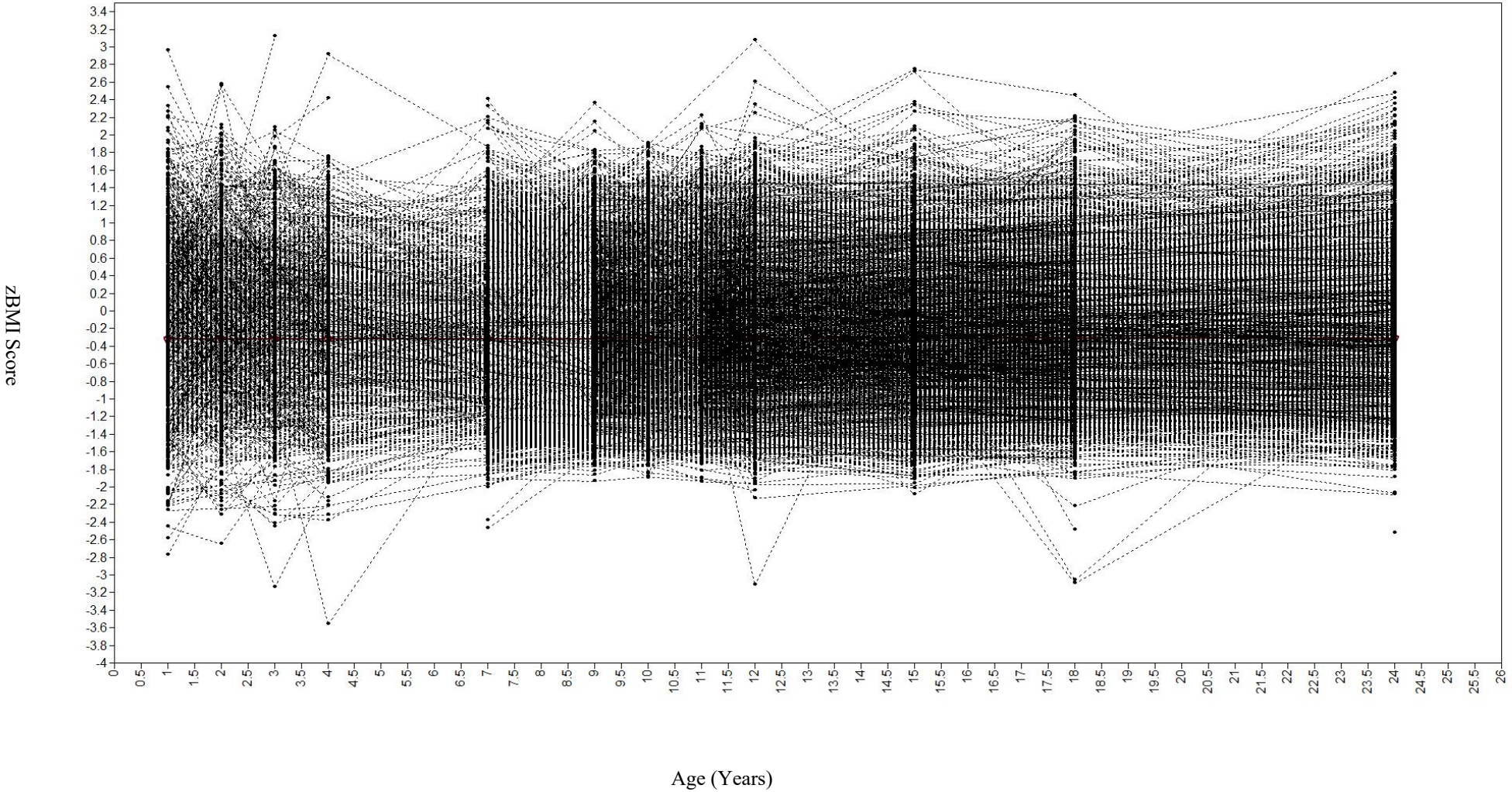


Appendix B Figure 3: Sensitivity Analysis Examining Trajectories of Fasting Insulin Between Ages 15-24

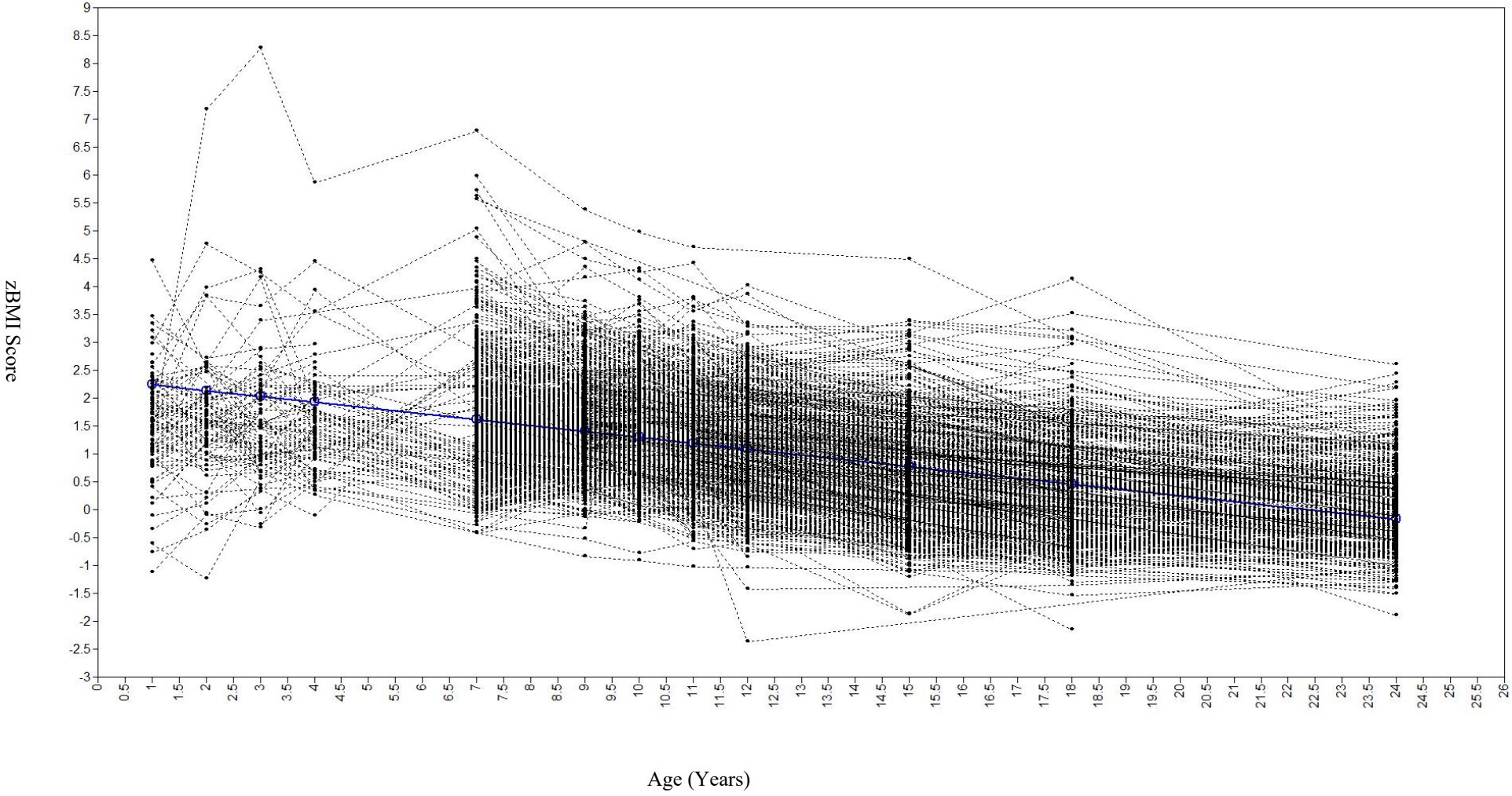


Appendix B Figure 4: Trajectory Means and Individual Values per Developmental Trajectory of Body Mass Index

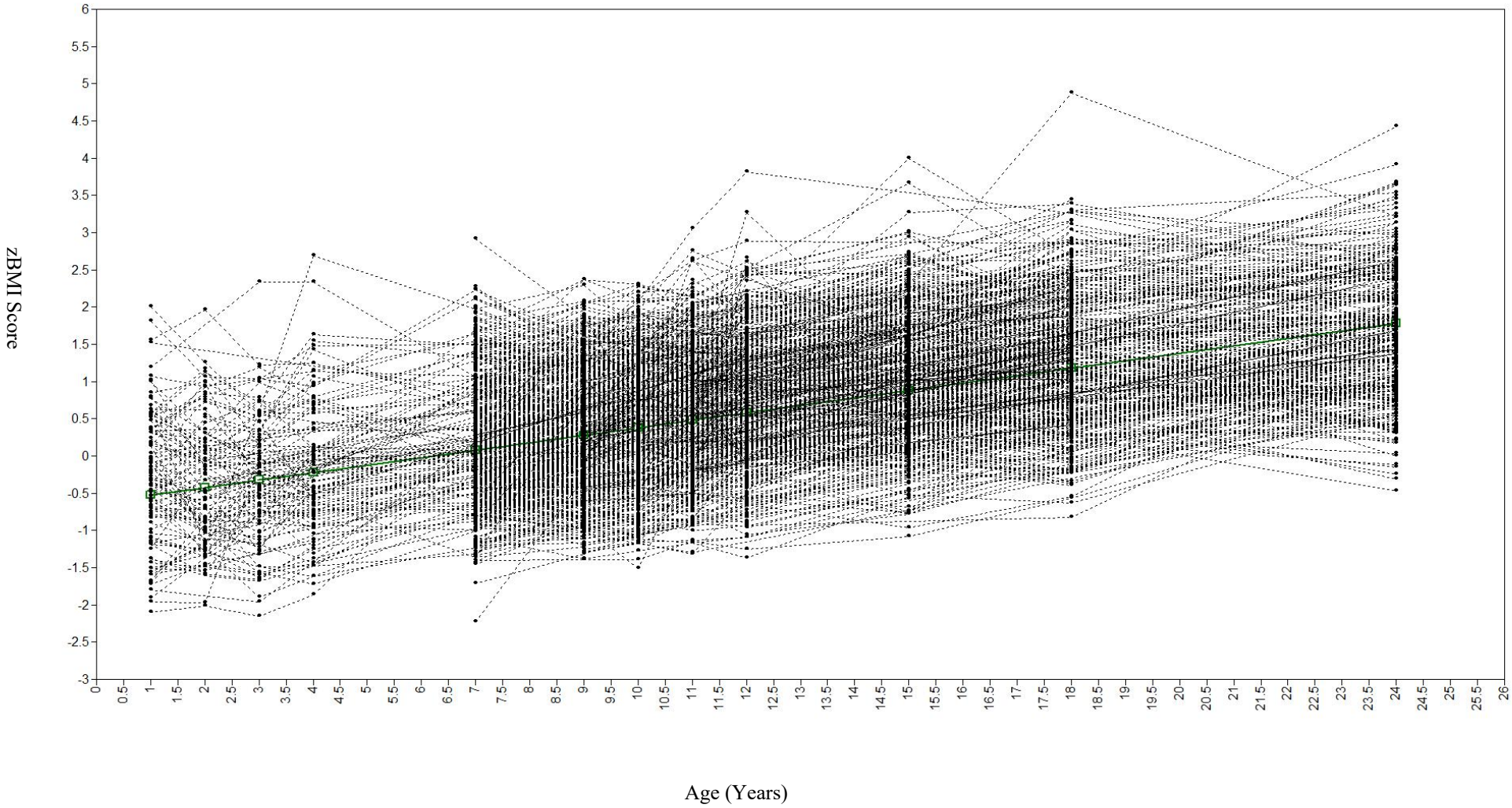
A. Class 1: Stable Average



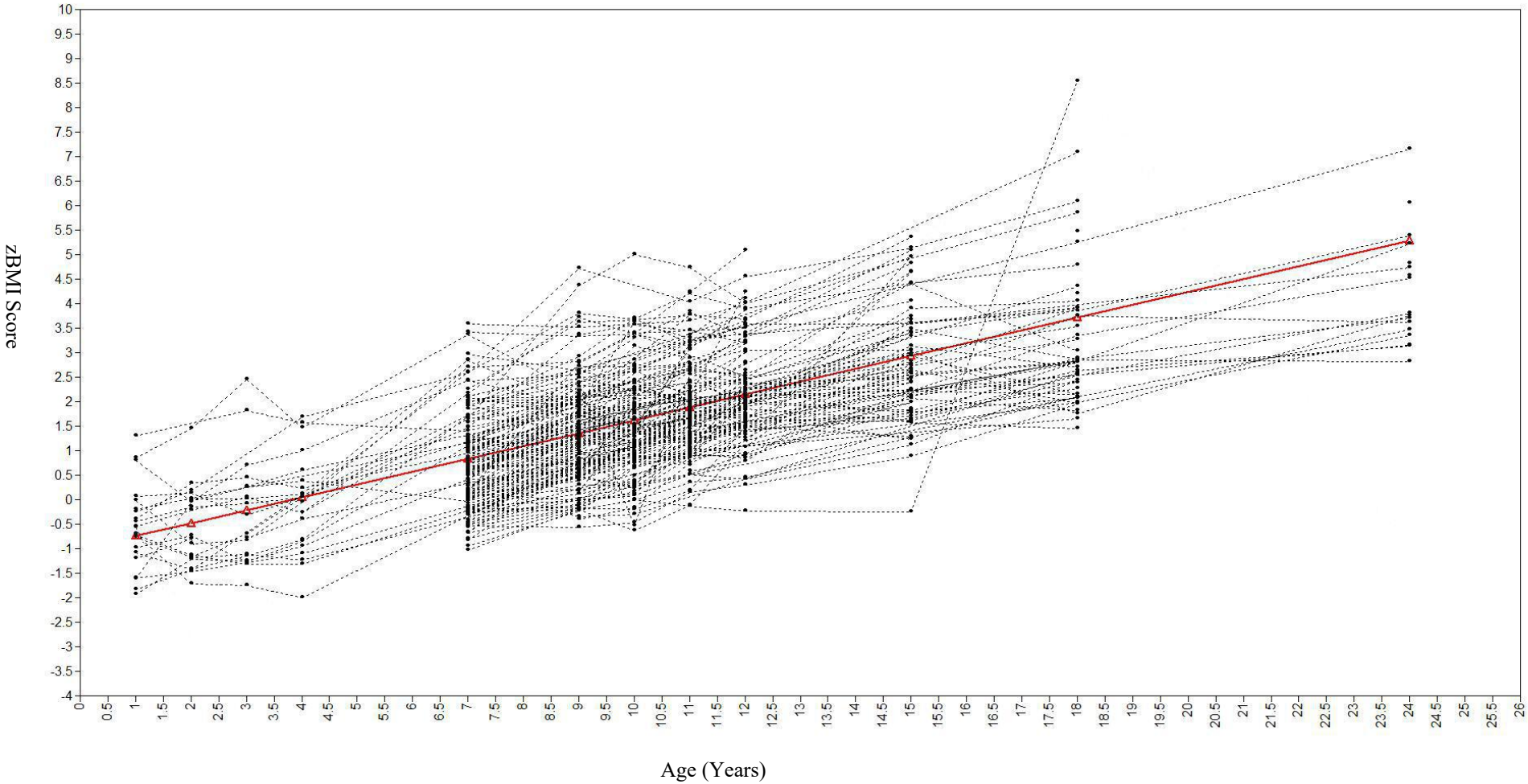
B. Class 2: Gradually Decreasing



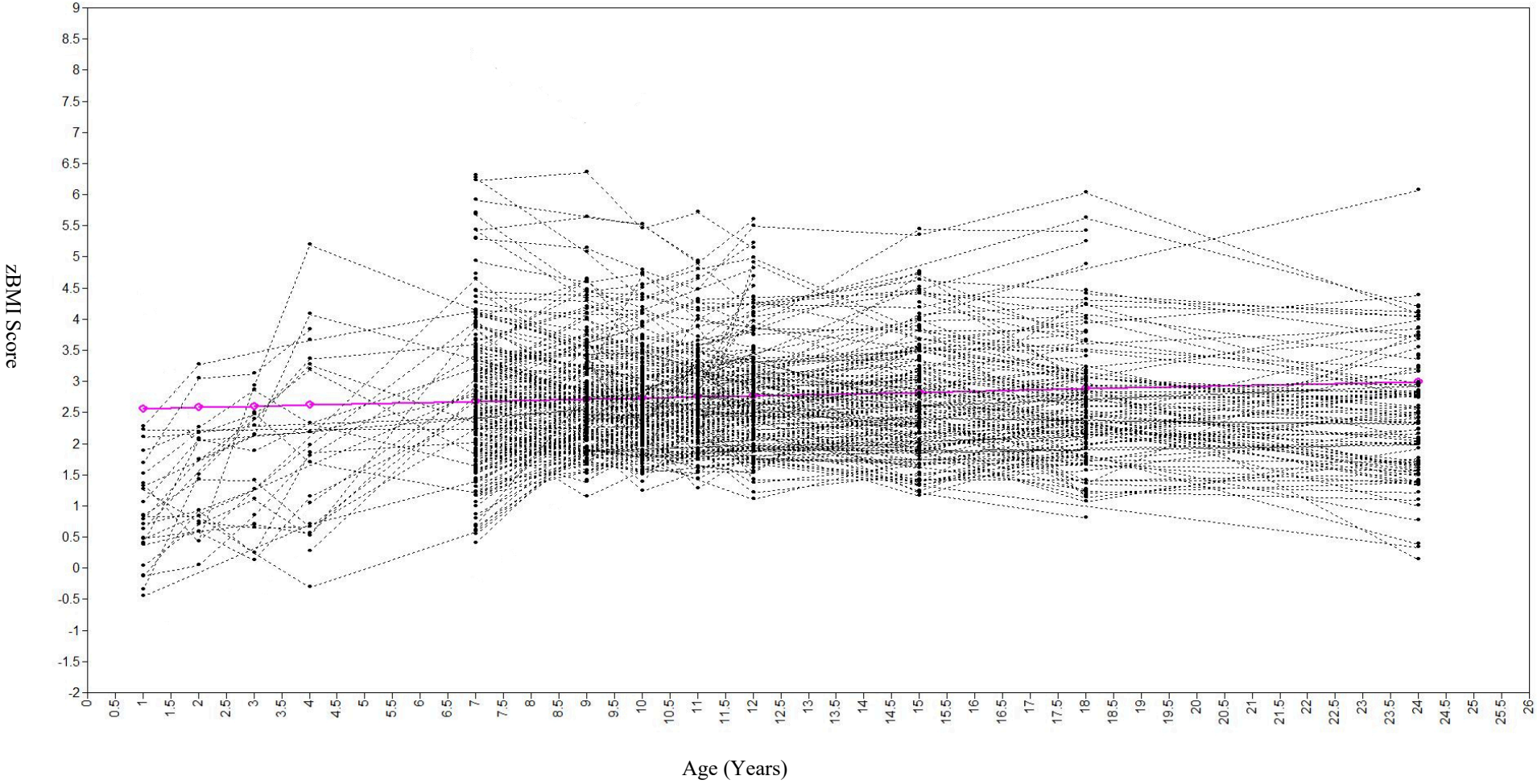
C. **Class 3: Puberty Onset – Minor Increase**



D. Class 4: Puberty Onset – Major Increase (1.9% of Sample)



E: Class 5: Persistently High



Appendix B Published Manuscript

JAMA Psychiatry | Original Investigation

Longitudinal Trends in Childhood Insulin Levels and Body Mass Index and Associations With Risks of Psychosis and Depression in Young Adults

Benjamin I. Perry, MRCPsych; Jan Stochl, PhD; Rachel Upthegrove, PhD; Stan Zammit, PhD; Nick Wareham, PhD; Claudia Langenberg, PhD; Eleanor Winpenny, PhD; David Dunger, PhD; Peter B. Jones, PhD; Golam M. Khandaker, PhD

IMPORTANCE Cardiometabolic disorders often occur concomitantly with psychosis and depression, contribute to high mortality rates, and are detectable from the onset of the psychiatric disorders. However, it is unclear whether longitudinal trends in cardiometabolic traits from childhood are associated with risks for adult psychosis and depression.

OBJECTIVE To examine whether specific developmental trajectories of fasting insulin (FI) levels and body mass index (BMI) from early childhood were longitudinally associated with psychosis and depression in young adults.

DESIGN, SETTING, AND PARTICIPANTS A cohort study from the Avon Longitudinal Study of Parents and Children, a prospective study including a population-representative British cohort of 14 975 individuals, was conducted using data from participants aged 1 to 24 years. Body mass index and FI level data were used for growth mixture modeling to delineate developmental trajectories, and associations with psychosis and depression were assessed. The study was conducted between July 15, 2019, and March 24, 2020.

EXPOSURES Fasting insulin levels were measured at 9, 15, 18, and 24 years, and BMI was measured at 1, 2, 3, 4, 7, 9, 10, 11, 12, 15, 18, and 24 years. Data on sex, race/ethnicity, paternal social class, childhood emotional and behavioral problems, and cumulative scores of sleep problems, average calorie intake, physical activity, smoking, and alcohol and substance use in childhood and adolescence were examined as potential confounders.

MAIN OUTCOMES AND MEASURES Psychosis risk (definite psychotic experiences, psychotic disorder, at-risk mental state status, and negative symptom score) depression risk (measured using the computerized Clinical Interview Schedule–Revised) were assessed at 24 years.

RESULTS From data available on 5790 participants (3132 [54.1%] female) for FI levels and data available on 10 463 participants (5336 [51.0%] female) for BMI, 3 distinct trajectories for FI levels and 5 distinct trajectories for BMI were noted, all of which were differentiated by mid-childhood. The persistently high FI level trajectory was associated with a psychosis at-risk mental state (adjusted odds ratio [aOR], 5.01; 95% CI, 1.76-13.19) and psychotic disorder (aOR, 3.22; 95% CI, 1.29-8.02) but not depression (aOR, 1.38; 95% CI, 0.75-2.54). A puberty-onset major increase in BMI was associated with depression (aOR, 4.46; 95% CI, 2.38-9.87) but not psychosis (aOR, 1.98; 95% CI, 0.56-7.79).

CONCLUSIONS AND RELEVANCE The cardiometabolic comorbidity of psychosis and depression may have distinct, disorder-specific early-life origins. Disrupted insulin sensitivity could be a shared risk factor for comorbid cardiometabolic disorders and psychosis. A puberty-onset major increase in BMI could be a risk factor or risk indicator for adult depression. These markers may represent targets for prevention and treatment of cardiometabolic disorders in individuals with psychosis and depression.

Supplemental content

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Benjamin I. Perry, MRCPsych, Inflammation and Psychiatry Research Group, Department of Psychiatry, University of Cambridge, Herchel Smith Building, Robinson Way, Cambridge, CB2 0SZ United Kingdom (bip20@medschl.cam.ac.uk).

JAMA Psychiatry. 2021;78(4):416-425. doi:10.1001/jamapsychiatry.2020.4180
Published online January 13, 2021. Corrected on April 7, 2021.

Cardiometabolic disorders often occur concomitantly with depression and schizophrenia,[1] leading to a reduced quality of life, increased health care costs,[2] and a shortened life expectancy.[3,4] Traditionally, this comorbidity has been attributed to chronic lifestyle factors (eg, physical inactivity or smoking) or the adverse effects of psychotropic medications.[5] However, meta-analyses report altered glucose-insulin homeostasis in relatively young, drug-naive patients with first-episode psychosis.[6,7] Similarly, reports from population-based longitudinal studies suggest a bidirectional association between depression and cardiovascular disease.[8,9] Together, this evidence suggests that cardiometabolic and psychiatric conditions may share pathophysiologic mechanisms. However, 2 key issues remain.

First, existing studies have predominantly included prevalent depression or psychosis cases and so cannot appropriately test the direction of association between cardiometabolic and psychiatric phenotypes.[10] Second, most longitudinal studies have included one-off measures of cardiometabolic indices, overlooking dynamic temporal changes in these markers.[11,12] Longitudinal repeated measurements could provide a more reliable measure of underlying homeostatic mechanisms and could identify population subgroups. For example, aberrant trajectories of childhood body mass index (BMI) are associated with adult cardiometabolic disorders.[13] Although cardiometabolic function encompasses a broad range of parameters, 2 pathways—insulin sensitivity and adiposity—are of particular interest regarding psychosis and depression. Genetic studies have indicated associations of BMI with depression[14] and fasting insulin (FI) levels with schizophrenia.[15] However, to our knowledge, no studies have examined whether FI level and BMI trajectories from childhood are associated with adult psychosis and depression.

Using data from the Avon Longitudinal Study of Parents and Children (ALSPAC) birth cohort, we aimed to (1) delineate longitudinal trajectories of FI level and BMI based on repeated measurements in individuals between ages 1 and 24 years, (2) examine the characteristics of identified trajectories, and (3) test associations with risks of psychosis and depression at 24 years in the total sample and by sex. We hypothesized that altered cardiometabolic development from childhood would be associated with increased risks for depression and psychosis in adulthood.

Key Points

Question Are longitudinal trends in insulin levels and body mass index from childhood associated with adult depression and psychosis?

Findings This cohort study of repeated-measure data from age 1 to 24 years in up to 10 463 individuals identified trajectories of fasting insulin levels and body mass index. Persistently high fasting insulin levels from age 9 years were associated with psychosis at 24 years, and puberty-onset body mass index increase was associated with depression at 24 years.

Meaning This study's findings suggest that changes in insulin sensitivity and adiposity starting from childhood may have disorder-specific associations with psychosis and depression and represent targets for prevention and treatment of cardiometabolic disorders in people with psychosis and depression.

Methods

Cohort and Sample

The ALSPAC initially recruited 14 541 pregnant residents in southwest England, with expected delivery dates between April 1, 1991, and December 31, 1992, resulting in 14 062 live births.[16-18] An additional 913 participants were recruited subsequently. Participants received financial compensation. Data were collected and managed using REDCap (University of Bristol[19,20]). Modeling of the trajectories was performed using 5790 participants for FI levels and 10 463 participants for BMI (eFigure 1 in the Supplement). Missing exposure data were handled using full-information maximum likelihood estimation (eMethods in the Supplement). Data were deidentified. The ALSPAC Ethics and Law Committee and local research ethics committees provided ethical approval for the ALSPAC cohort study. Ethical approval for the present study was obtained via the ALSPAC Executive Committee. Consent for biological samples was collected in accordance with the Human Tissue Act of 2004 covering England, Wales and Northern Ireland. Informed consent for all collected data was obtained from participants following the recommendations of the ALSPAC Ethics and Law Committee at the time. This study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline for cohort study.

Measurement of Exposures

Fasting insulin levels were measured at ages 9 (n = 894), 15 (n = 3484), 18 (n = 3286), and 24 (n = 3253) years, using an ultrasensitive automated microparticle enzyme immunoassay (Mercodia), which does not cross-react with proinsulin. Sensitivity of the immunoassay was 0.07 mU/L, and interassay and intraassay coefficients of variation were less than 6%. Fasting blood samples were drawn at 9 AM after a 10-hour fast, then spun and stored at −80 °C. There was no evidence of freeze-thaw cycles during storage.

Body mass index, calculated as weight in kilograms divided by height in meters squared, was measured at 1 (n = 1236), 2 (n = 1036), 3 (n = 1050), 4 (n = 1018), 7 (n = 8200), 9 (n = 7633), 10 (n = 7465), 11 (n = 7100), 12 (n = 6704), 15 (n = 5415), 18 (n = 5061), and 24 (n = 3975) years.

Psychiatric Outcomes at Age 24 Years

Psychotic experiences (PEs) were identified through the semi-structured Psychosis-Like Symptom Interview[21] conducted by trained psychology graduates and coded per the definitions in the Schedules for Clinical Assessment in Neuropsychiatry, version 2.0.[22] The Psychosis-Like Symptom Interview had good interrater (intraclass correlation: 0.81; 95% CI, 0.68-0.89) and test-retest (0.9; 95% CI, 0.83-0.95) reliability. Psychotic experiences occurring in the past 6 months covered the 3 main positive symptom domains: hallucinations, delusions, and thought interference. After cross-questioning, interviewers rated PEs

as absent, suspected, or definite. We included cases of definite PEs; the comparator group comprised individuals with suspected or absent PEs.

Cases of at-risk mental state were identified by mapping Psychosis-Like Symptom Interview data to Comprehensive Assessment of At-Risk Mental State (CAARMS) criteria.[23] Cases were defined as participants meeting CAARMS criteria for attenuated psychosis (symptoms not reaching the psychosis threshold owing to levels of intensity or frequency) or brief limited intermittent psychosis (frank psychotic symptoms that resolved spontaneously within 1 week).

Cases of psychotic disorder were defined[21] as definite PEs that were not attributable to sleep or fever, had occurred more than once per month during the previous 6 months, and were very distressing or negatively impactful on social/occupational functioning, and led to seeking of professional help. We also included participants meeting the criteria for CAARMS psychotic disorder (threshold psychotic symptoms occurring for >1 week).

Ten questions from the Community Assessment of Psychic Experiences questionnaire[24] were administered covering interest, motivation, emotional reactivity, pleasure, and sociability. Participants rated each item as 0 (never), 1 (sometimes), 2 (often), and 3 (always). We recoded the variables by scoring always and often as 1 and never and sometimes as 0, and then summed the values to result in a possible total score of 0 to 10.

Depression was measured using the computerized Clinical Interview Schedule-Revised.[25] The interview assesses symptoms of depression occurring in the past week and provides a diagnosis of depressive episode based on the *International Statistical Classification of Diseases, Tenth Revision* criteria, which we used as a binary outcome (codes F32.0-32.2). We also included a Clinical Interview Schedule-Revised depression severity score, comprising scores for mood, thoughts, fatigue, concentration, and sleep, as a continuous outcome.

For assessment of potential confounders, we included sex at birth, race/ethnicity, paternal social class, childhood emotional and behavioral problems (measured using the Strength and Difficulties Questionnaire[26] at age 7 years), and cumulative scores of smoking, physical activity, alcohol use, substance use, sleep problems, and average calorie intake between ages 7 and 24 years (eMethods in the Supplement).

Statistical Analysis

We standardized (z transformed) FI levels and BMI separately in males and females and then combined the sex-stratified z scores for each variable at each time point to delineate trajectories using curvilinear growth mixture modeling[27] (eMethods in the Supplement). We used z scores to measure the relative change in FI levels and BMI because BMI increases in all young people during early life. Because the sample size for FI levels at age 9 years was smaller, we repeated growth mixture modeling without age-9-years data and compared the characteristics of the resultant trajectories. Analyses were conducted using MPlus, version 8 (Muthén & Muthén), and R, version 3.6.0 (R Project for Statistical Computing). Two-tailed P values were corrected for multiple testing using the Holm-Bonferroni method[28] for the 6 psychiatric outcomes. A corrected P value <.05 was used as the threshold for significance. We estimated how participants overlapped between BMI and FI level trajectories (the most common and highest risk) using the φ correlation coefficient.

We used the 3-step method[29] to estimate associations of sociodemographic, lifestyle, and clinical factors with trajectory membership (eMethods in the Supplement). The 3-step method allows class separation unaffected by auxiliary variables, retains and includes information on class uncertainty, and is robust when entropy is greater than 0.60. Multinomial regression was used to estimate odds ratios (ORs) and 95% CIs for the associations of sociodemographic and lifestyle factors with FI level and BMI trajectories compared with the most common trajectory. We considered time-invariant (sex, ethnicity, social class at birth, family history of cardiovascular disease, gestational age, birth weight, and perinatal stressful life events) and time-variant (physical activity and smoking in adolescence and early adulthood) factors. Odds ratios represent the increase in the risk of membership of a particular trajectory category per SD increase in factor. Next, we examined the clinical phenotype of trajectories at 24 years, examining mean levels of commonly measured clinical and biochemical factors for participants grouped by most-likely trajectory membership (eMethods in the Supplement). Next, we used logistic regression to estimate the association of trajectory membership with an age-appropriate cardiometabolic outcome: metabolic syndrome at 24 years (eMethods in the Supplement).

Using the 3-step method, logistic regression was used to estimate ORs and 95% CIs for binary outcomes per trajectory, compared with the most common trajectory. Linear regression for continuous outcomes was used to estimate β coefficients and 95% CIs representing the SD increase in the risk of outcomes per trajectory. We tested associations for the total sample and separately by sex before and after adjusting for potential confounders. Regression models for negative symptoms were additionally adjusted for depressive symptoms, and vice versa.

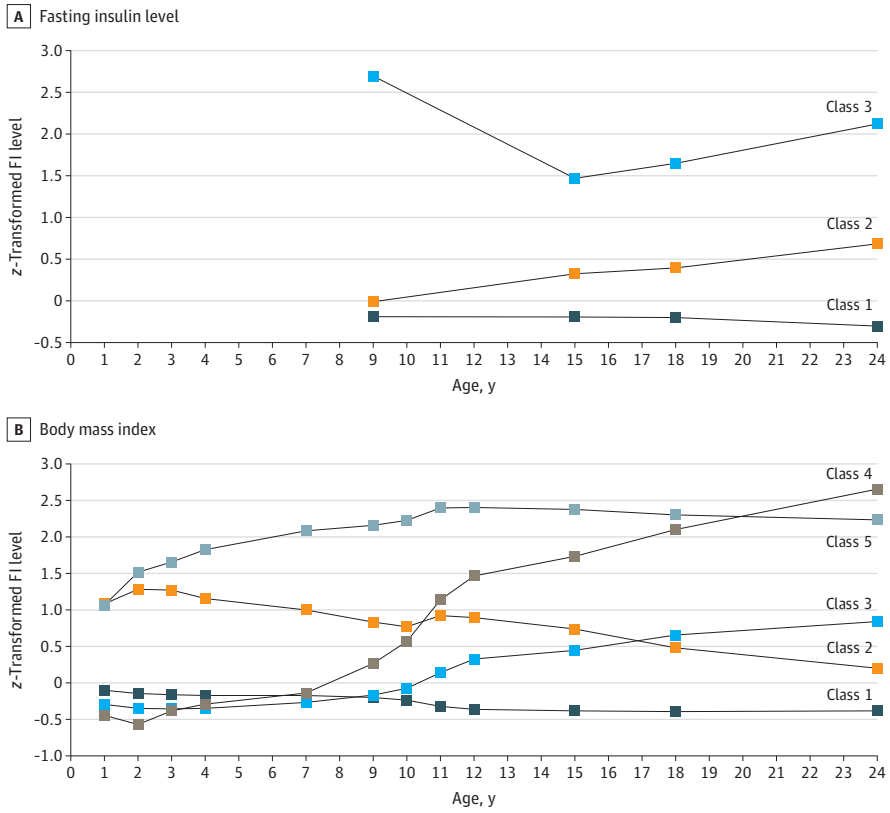
Results

Trajectories of FI Levels and BMI

Based on 5790 participants (2658 [45.9%] male, 3132 [54.1%] female), the 3-trajectory solution for FI levels was optimum, representing stable average (class 1: 4939 [77.8%]), minor increase (class 2: 693 [19.0%]), and persistently high (class 3: 158 [3.1%]) trajectories between ages 9 and 24 years (**Figure 1**A; eTable 1 and eFigure 2 in the Supplement). The trajectories were similar after excluding age-9-years data (eFigure 3 in the Supplement).

Based on 10 463 participants (5336 [51.0%] female, 5127 [49.0%] male) included in the analysis of BMI, the 5-trajectory solution was optimum, representing stable average (class 1: 8383 [71.1%]), gradually decreasing (class 2: 949 [7.0%]), puberty-onset minor increase (class 3: 668 [14.5%]), puberty-onset major increase (class 4: 174 [1.9%]), and persistently high (class 5: 289 [5.5%]) BMI trajectories between

Figure 1. Fasting Insulin Levels and Body Mass Index Trajectories in the Avon Longitudinal Study of Parents and Children



A, Fasting insulin levels measured at ages 9 to 24 years in 5790 participants. Class 1 (stable average) comprised 77.8% of the sample; class 2 (minor increase), 19.0%; and class 3 (persistently high), 3.1%. B, Body mass index measured at ages 1 to 24 years in 10 463 participants. Class 1 (stable average) comprised 71.1% of the sample; class 2 (gradually decreasing), 7.0%; class 3 (puberty-onset minor increase), 14.5%; class 4 (puberty-onset major increase), 1.9%; and class 5 (persistently high), 5.5%. Trajectories were delineated using growth mixture modeling at 4 time points for fasting insulin and 12 time points for body mass index. Nodes in the graph represent mean *z* scores for fasting insulin level or body mass index at each time point for each developmental trajectory.

ages 1 and 24 years (Figure 1B; eTable 2 and eFigure 4 in the Supplement).

The stable average FI level and BMI trajectories were statistically significantly correlated (r_{φ} = 0.233, P < .001), as were the persistently high trajectories (r_{φ} = 0.092, P < .001).

Both deviating FI level trajectories were associated with lower social class, family history of cardiometabolic disease, lower physical activity, and smoking in adolescence and early adulthood. Lower birth weight and more perinatal stressful life events were associated with the persistently high trajectory compared with the stable average trajectory (eTable 3 in the Supplement). The persistently high trajectory cohort also had mean FI, high-density lipoprotein cholesterol, triglyceride, and C-reactive protein levels outside of UK reference ranges at 24 years (eTable 4 in the Supplement). Deviating FI level trajectories were associated with metabolic syndrome at 24 years (adjusted OR [aOR] for the persistently high trajectory, 9.21; 95% CI, 3.77-20.15) (eTable 5 in the Supplement).

Deviating BMI trajectories were associated with lower social class, family history of cardiometabolic disease, more perinatal stressful life events, lower physical activity, and smoking in adolescence and early adulthood compared with the stable average trajectory. Higher birth weight was associated with the gradually decreasing and persistently high trajectories, whereas lower birth weight was weakly associated with both puberty-onset increase trajectories (eTable 6 in the Supplement). Deviating BMI trajectories were also associated with mean values of waist circumference and FI, high-density lipoprotein cholesterol, and C-reactive protein levels outside of UK reference ranges at 24 years (eTable 7 in the Supplement). All deviating BMI trajectories were associated with metabolic syndrome at 24 years (aOR for the persistently high trajectory, 10.62; 95% CI, 5.89-19.13) (eTable 5 in the Supplement).

Associations of FI and BMI Trajectories With Psychiatric Outcomes

The persistently high FI level trajectory was associated with the psychosis at-risk mental state (aOR, 5.01; 95% CI, 1.76-13.19), psychotic disorder (aOR, 3.22; 95% CI, 1.29-8.02), and negative symptoms (adjusted β, 0.07; 95% CI, 0.01-0.13) at age 24 years. Fasting insulin level trajectories were not associated with depression (aOR, 1.38; 95% CI, 0.75-2.54) (**Table 1**; **Figure 2**A; eTable 8 in the Supplement).

The puberty-onset major increase trajectory of BMI was associated with a higher risk of a depressive episode

Table 1. Psychosis and Depressive Outcomes at Age 24 Years Associated With Fasting Insulin Level Trajectories From Age 9 to 24 Years

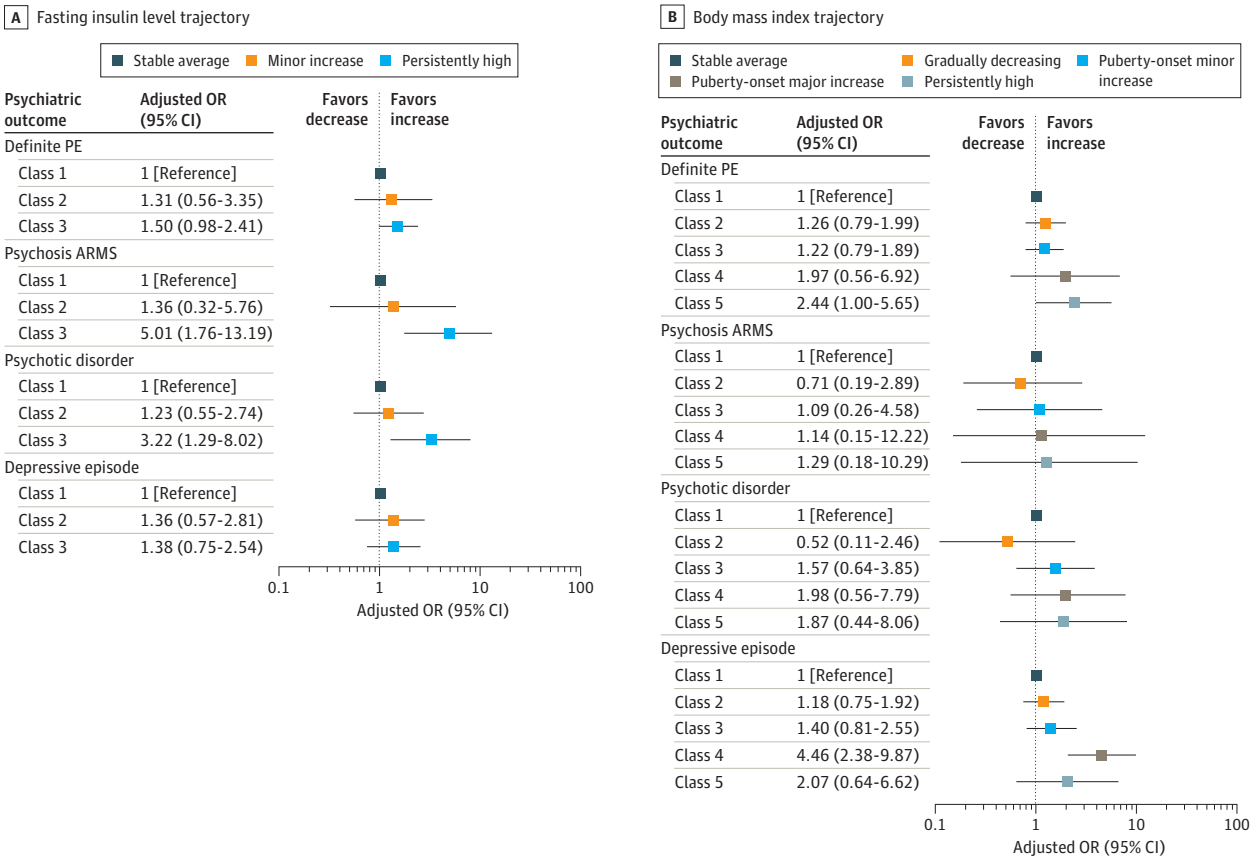
Trajectory and outcome at 24 y	Sample, No.	Odds ratio (95% CI)		*P* value[b]
		Unadjusted	Adjusted[a]	
Definite PE				
Class 1: stable average	4939	1 [Reference]	1 [Reference]	
Class 2: minor increase	693	1.48 (0.98-2.24)	1.31 (0.56-3.35)	>.99
Class 3: persistently high	158	1.88 (1.05-3.60)	1.50 (0.98-2.41)	.33
Psychosis at-risk mental state				
Class 1: stable average	4939	1 [Reference]	1 [Reference]	
Class 2: minor increase	693	1.59 (0.20-8.02)	1.36 (0.32-5.76)	>.99
Class 3: persistently high	158	6.33 (1.97-20.30)	5.01 (1.76-13.19)	.006
Psychotic disorder				
Class 1: stable average	4939	1.00 [reference]	1.00 [Reference]	
Class 2: minor increase	693	1.85 (0.70-4.88)	1.23 (0.55-2.74)	>.99
Class 3: persistently high	158	4.74 (1.67-13.42)	3.22 (1.29-8.02)	.05
Depressive episode				
Class 1: stable average	4939	1 [Reference]	1 [Reference]	
Class 2: minor increase	693	1.26 (0.73-2.67)	1.36 (0.57-2.81)	.88
Class 3: persistently high	158	1.31 (0.81-4.32)	1.38 (0.75-2.54)	.69

Abbreviation: PE, psychotic experience.

[a] Adjusted for sex, ethnicity, social class, Strength and Difficulties Questionnaire (measured at 7 years) findings, and cumulative scores for smoking, physical activity, alcohol and substance use, sleep problems, and calorie intake.

[b] *P* values adjusted for multiple testing using the Holm-Bonferroni method.

Figure 2. Associations of Fasting Insulin Levels and Body Mass Index Trajectories With Psychosis and Depressive Outcomes in the Avon Longitudinal Study of Parents and Children



Adjusted odds ratios (ORs) and 95% CIs showing associations of fasting insulin (A) and body mass index (B) trajectories from childhood with risk of psychosis and depression outcomes at age 24 years after adjusting for sex, ethnicity, social class, childhood emotional and behavioral problems, and cumulative scores of smoking, physical activity, alcohol and substance use, sleep problems, and calorie intake. ARMS indicates at-risk mental state; PE, psychotic experiences.

Table 2. Psychiatric Outcomes at Age 24 Years Associated With BMI Trajectories From Age 1 to 24 Years

Trajectory and outcome at 24 y	Sample, No.	Odds ratio (95% CI)		*P* value[b]
		Unadjusted	Adjusted[a]	
Definite PE				
Class 1: stable average	8383	1 [Reference]	1 [Reference]	
Class 2: gradually decreasing	949	1.43 (0.82-1.96)	1.26 (0.79-1.99)	>.99
Class 3: puberty-onset minor increase	668	1.66 (0.87-2.55)	1.22 (0.79-1.89)	>.99
Class 4: puberty-onset major increase	174	3.56 (0.87-11.54)	1.97 (0.56-6.92)	>.99
Class 5: persistently high	289	3.21 (1.01-9.11)	2.44 (1.00-5.65)	.37
Psychosis at-risk mental state				
Class 1: stable average	8383	1 [Reference]	1 [Reference]	
Class 2: gradually decreasing	949	0.49 (0.10-3.21)	0.71 (0.19-2.89)	>.99
Class 3: puberty-onset minor increase	668	1.12 (0.23-5.43)	1.09 (0.26-4.58)	>.99
Class 4: puberty-onset major increase	174	1.32 (0.10-13.11)	1.14 (0.15-12.22)	>.99
Class 5: persistently high	289	1.55 (0.44-3.21)	1.29 (0.18-10.29)	>.99
Psychotic disorder				
Class 1: stable average	8383	1 [Reference]	1 [Reference]	
Class 2: gradually decreasing	949	0.44 (0.21-2.03)	0.52 (0.11-2.46)	>.99
Class 3: puberty-onset minor increase	668	1.97 (0.60-3.46)	1.57 (0.64-3.85)	>.99
Class 4: puberty-onset major increase	174	2.14 (0.65-6.21)	1.98 (0.56-7.79)	>.99
Class 5: persistently high	289	3.11 (0.53-13.22)	1.87 (0.44-8.06)	>.99
Depressive episode				
Class 1: stable average	8383	1 [Reference]	1 [Reference]	
Class 2: gradually decreasing	949	1.33 (0.77-1.88)	1.18 (0.75-1.92)	>.99
Class 3: puberty-onset minor increase	668	1.69 (0.90-3.21)	1.40 (0.81-2.55)	>.99
Class 4: puberty-onset major increase	174	8.91 (4.21-17.12)	4.46 (2.38-9.87)	.006
Class 5: persistently high	289	3.01 (0.91-7.59)	2.07 (0.64-6.62)	>.99

Abbreviations: BMI, body mass index; PE, psychotic experience.

[a] Adjusted for sex, ethnicity, social class, Strength and Difficulties Questionnaire (measured at 7 years), and cumulative scores for smoking, physical activity, alcohol and substance use, sleep problems, and calorie intake.

[b] *P* values adjusted for multiple testing using the Holm-Bonferroni method.

(aOR, 4.46; 95% CI, 2.38-9.87) and depressive symptoms (adjusted β, 0.08; 95% CI, 0.03-0.14) at age 24 years. The puberty-onset minor increase trajectory was weakly associated with depressive symptoms at 24 years (adjusted β, 0.06; 95% CI, 0.01-0.11). Body mass index trajectories were not associated with psychosis outcomes (aOR for psychotic disorder in the puberty-onset major increase trajectory, 1.98; 95% CI, 0.56-7.79) (**Table 2** and Figure 2B; eTable 9 in the Supplement).

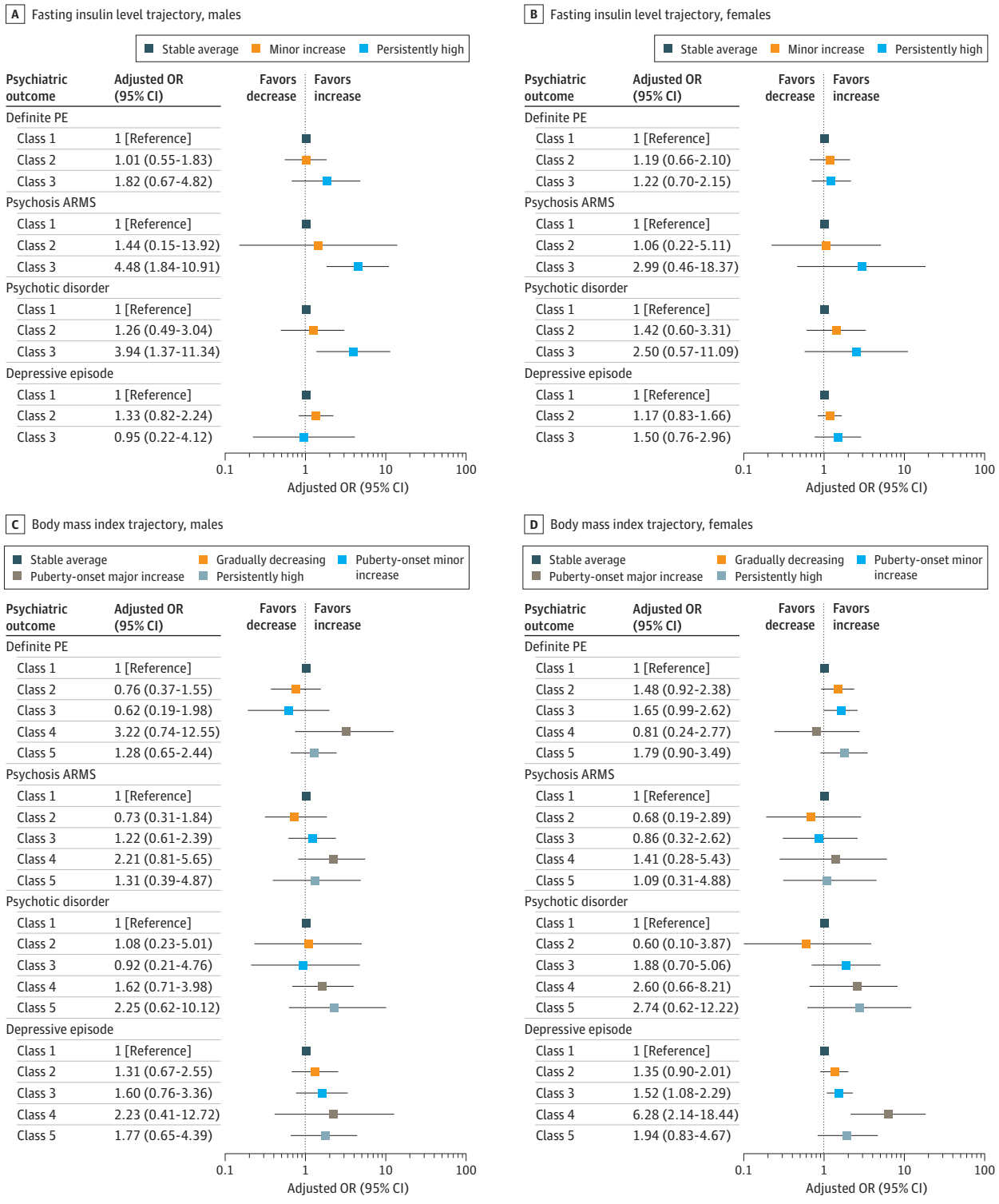
Sex-Stratified Associations of Risks for Psychiatric Outcomes

For FI trajectories, the pattern of association with risks for psychiatric outcomes in sex-stratified analysis was similar to the primary analysis. For example, point estimates for the association between the persistently high FI trajectory and psychotic disorder were similar in males (aOR, 3.94; 95% CI, 1.10-11.96) compared with females (aOR, 2.50; 95% CI, 0.57-11.09), and 95% CIs overlapped. There was no association between persistently high FI and depression in males (aOR, 0.95; 95% CI, 0.22-4.12) or females (aOR, 1.50; 95% CI, 0.76-2.96) (**Figure 3**; eTable 10 and eTable 11 in the Supplement). For BMI, point estimates for depression for both puberty-onset increase trajectories were larger in females. For example, for the puberty-onset major increase trajectory, the association for females (aOR, 6.28; 95% CI, 2.14-18.44) was stronger than for males (aOR, 2.23; 95% CI, 0.41-12.72). There was no significant association of BMI trajectories with psychosis outcomes. For example, there was no association between puberty-onset major BMI increase and psychotic disorder for males (aOR, 1.62; 95% CI, 0.71-3.98) or females (aOR, 2.60; 95% CI, 0.66-8.21) (Figure 3; eTable 12 and eTable 13 in the Supplement).

Discussion

We delineated FI level and BMI trajectories from early life, using prospective repeated measurements in a large population-representative birth cohort, and report distinct associations with psychosis and depression measured in adulthood. After adjusting for a number of relevant confounders, we found that persistently high FI levels from mid-childhood appeared to be associated with an increased risk of psychosis outcomes at age 24 years, while BMI increases around the age of puberty onset were associated with depression at age 24 years. Associations of BMI and FI level trajectories with cardiometabolic risk factors, such as social class, ethnicity, smoking, physical activity, and adult metabolic syndrome, suggest face validity to the identified trajectories. Although the last data point for BMI and FI levels overlapped with the outcome assessment, the trajectories were differentiated by mid-childhood, suggesting a temporal association between exposure and outcome. Evidence for the association of puberty-onset BMI increase and adult depression remained after adjusting for childhood emotional and behavioral problems, suggesting that

Figure 3. Sex-Stratified Associations of Fasting Insulin Levels and Body Mass Index Trajectories With Psychosis and Depressive Outcomes in the Avon Longitudinal Study of Parents and Children



Adjusted odds ratios (ORs) and 95% CIs showing associations of fasting insulin level trajectories in males (A) and females (B) and body mass index trajectories in males (C) and females (D) from childhood with risk of psychosis and depression outcomes at age 24 years after adjusting for sex, ethnicity, social class, childhood emotional and behavioral problems, and cumulative scores of smoking, physical activity, alcohol and substance use, sleep problems, and calorie intake. ARMS indicates at-risk mental state; PE, psychotic experience.

a reverse direction of association may not fully explain this finding. Although the same adjustment may be less capable of ruling out reverse direction of the association between persistently high FI levels and psychosis, it is unlikely that many participants had experienced psychosis before age 9 years, and so a reverse direction of association is unlikely.

We found consistent evidence for an association between FI level trajectories and psychosis outcomes. Effect sizes were largest in the persistently high trajectory, consistent with a dose-response relationship, and point estimates were larger in more clinically relevant outcomes. Our findings complement meta-analyses reporting altered glucose-insulin homeostasis in first-episode psychosis.[6,7] Moreover, our results suggest that disruptions to glucose-insulin homeostasis detectable at first-episode psychosis in adults may begin in childhood. The point estimates partly attenuated after adjustment for confounders, suggesting that malleable lifestyle factors, such as smoking, physical activity, and diet, should remain targets for reducing the risk of incident cardiometabolic disorders in young people with psychosis. We also found that participants classified into the persistently high FI level trajectory, who had the highest risk of psychosis, had mean BMI and fasting plasma glucose values within reference ranges at age 24 years. Therefore, the risk of incident cardiometabolic disorders in these individuals may not be detected in psychosis early-intervention services, since commonly measured physical indices may not identify them. Consequently, careful assessment and clinical considerations are needed to minimize the risk of cardiometabolic disorders in these individuals.

Our findings suggest that altered glucose-insulin homeostasis could be a shared mechanism for psychosis and type 2 diabetes, which could be genetic and/or environmental in origin. People with comorbid schizophrenia and type 2 diabetes have a higher genetic predisposition for both disorders compared with controls,[30] and genetic predisposition for schizophrenia is associated with insulin resistance in patients with schizophrenia.[31] In addition, we found that the persistently high FI trajectory, which had the strongest associations with psychosis outcomes, was also associated with lower birth weight and perinatal stressful life events. We noted similar patterns of association in BMI trajectories that were associated with depression. These findings are consistent with the fetal programming hypothesis,[32] which posits that disruption in early-life development can have broad influences on adult health.

Our findings regarding the association of BMI trajectories with depression at age 24 years are in line with meta-analyses[33,34] suggesting an association between BMI and risk of depression. Similar trajectories of BMI have been linked with adult type 2 diabetes,[35] obesity,[36] and coronary heart disease.[37] The character and composition of BMI trajectories we identified are consistent with those of previous studies, although our length of follow-up was longer than the follow-up of most previous studies.[38]

Our findings provide further insights into the link between BMI and depression,[34] suggesting that puberty-onset increases in BMI specifically are associated with risk of adult depression. This finding, together with the lack of evidence for an association between persistently high BMI and depression, indicates that BMI might be a risk indicator for depression rather than a risk factor because individuals in the persistently high BMI trajectory would likely have been exposed to the "largest dose" of BMI. Therefore, if BMI were the risk factor, one would have expected the largest effect size for depression in that trajectory. Consequently, environmental and/or genetic factors influencing BMI during puberty are likely to be important risk factors for depression. For instance, social stressors, such as bullying, may predispose to altered eating behaviors and an increased risk of depression in adolescents.[39] In addition, deviating childhood BMI trajectories have been associated with a greater risk of adolescent and adult eating disorders,[40] which are commonly comorbid with depression.[41] Also, the effects of estrogen may be relevant, since the associations of puberty-onset BMI increases and depression appeared to be stronger in females than males. Changes in estrogen levels are associated with depressive symptoms throughout life in women, including pregnancy,[42] menopause,[43] and puberty.[44] Estrogen is associated with obesity[45] and may explain the genetic correlation of age at menarche with adult obesity[46] and depression.[47] Further research is needed to identify factors influencing pubertal BMI increases, as they may represent important preventive targets for depression.

We did not find consistent evidence for associations of FI level trajectories with depression or of BMI trajectories with psychosis. Previous research has reported mixed findings regarding the association between insulin resistance and depression in young adults.[48,49] However, some estimates for the associations of BMI trajectories with psychosis outcomes in our analyses had wide 95% CIs, possibly owing to sample size. These particular findings require replication in larger samples of people with psychosis.

Strengths and Limitations

Strengths of the study include a longitudinal design with repeated measurements of BMI and FI levels between ages 1 and 24 years in a relatively large sample enabling a detailed examination of dynamic cardiometabolic changes from childhood to early adulthood. We included several relevant depression and psychosis outcomes, which allowed us to examine for specificity and for a biological gradient of evidence.

Limitations of the study include missing data. Although we used a robust method to handle missing data, full-information maximum likelihood may be biased in instances in which data were not missing at random.[50] However, the risk of bias in full-information maximum likelihood is no greater than the bias associated with traditional complete-case methods,[51] and full-information maximum likelihood permitted a larger sample size and therefore increased statistical power. Nevertheless, missing psychiatric outcome data may have affected our results. Furthermore, although we adjusted for a number of relevant potential confounders,

residual confounding could still be an issue. For example, we could not account for psychological stress since data on cortisol levels were available only at age 9 years in a small subsection of the cohort. Therefore, further research is needed, such as mendelian randomization analysis, to examine for potentially unconfounded associations. In addition, the 95% CIs were relatively wide for the sex-stratified analysis, likely owing to reduced statistical power. Therefore, replication of our work in larger samples is required. In addition, the ALSPAC data set does not include *International Statistical Classification of Diseases* and *DSM* diagnoses of schizophrenia. However, our psychotic disorder outcome would likely meet the threshold for clinical intervention, and all our psychosis outcomes lie on the schizophrenia continuum.

Conclusions

We report that the cardiometabolic comorbidity of psychosis and depression may have distinct early-life origins. Disrupted insulin sensitivity from mid-childhood appeared to be associated with adult psychosis, and BMI increases starting around the time of puberty onset were associated with adult depression. Although residual confounding may be an issue, our results suggest that these cardiometabolic markers could be among shared risk factors and indicators for adult cardiometabolic and psychiatric disorders and may represent novel targets for prevention and treatment of cardiometabolic disorders in people with psychosis and depression.

ARTICLE INFORMATION

Accepted for Publication: November 10, 2020.

Published Online: January 13, 2021. doi:10.1001/jamapsychiatry.2020.4180

Correction: This article was corrected on April 7, 2021, to fix errors in Results, Figures, and Tables.

Author Affiliations: Department of Psychiatry, University of Cambridge School of Clinical Medicine, Cambridge, United Kingdom (Perry, Stochl, Jones, Khandaker); Cambridgeshire and Peterborough NHS Foundation Trust, Cambridge, United Kingdom (Perry, Jones, Khandaker); Department of Kinanthropology, Charles University, Prague, Czechia (Stochl); Institute for Mental Health, University of Birmingham, Birmingham, United Kingdom (Upthegrove); Centre for Academic Mental Health, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom (Zammit, Khandaker); MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, United Kingdom (Zammit); MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge, United Kingdom (Wareham, Langenberg, Winpenny); Department of Paediatrics, University of Cambridge School of Clinical Medicine, Cambridge, United Kingdom (Dunger); MRC Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom (Khandaker).

Author Contributions: Dr Perry had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Concept and design: Perry, Upthegrove, Jones, Khandaker.
Acquisition, analysis, or interpretation of data: All authors.
Drafting of the manuscript: Perry, Jones, Khandaker.
Critical revision of the manuscript for important intellectual content: Stochl, Upthegrove, Zammit, Wareham, Langenberg, Winpenny, Dunger, Jones, Khandaker.
Statistical analysis: Perry, Stochl.
Obtained funding: Perry, Upthegrove, Zammit, Jones, Khandaker.
Administrative, technical, or material support: Jones, Khandaker.
Supervision: Upthegrove, Wareham, Langenberg, Dunger, Jones, Khandaker.

Conflict of Interest Disclosures: Dr Perry reported receiving grants from the National Institute for Health Research (NIHR) during the conduct of the study. Dr Upthegrove reported receiving grants from the NIHR, UK Research and Innovation, EU Seventh Framework Programme, and personal fees from Sunovion outside the submitted work. Dr Winpenny reported receiving grants from the Medical Research Council during the conduct of the study. Dr Jones reported paid consulting for Recordati and Janssen outside the submitted work. No other disclosures were reported.

Funding/Support: Dr Perry received funding support from the NIHR (doctoral research fellowship DRF-2018-11-ST2-018). Dr Khandaker received funding support from the Wellcome Trust (intermediate clinical fellowship, grant 201486/Z/16/Z), the MQ: Transforming Mental Health (data science award, grant MQDS17/40), the Medical Research Council UK (MICA: Mental Health Data Pathfinder grant MC_PC_17213 and Therapeutic Target Validation in Mental Health grant MR/S037675/1), and the BMA Foundation (J Moulton grant 2019). Dr Jones received funding from the Medical Research Council and MQ: Transforming Mental Health (data science award, grant MQDS17/40), programmatic funding from NIHR (grant RP-PG- 0616-20003), and support from the Applied Research Collaboration East of England. Dr Upthegrove received funding support from the NIHR (HTA grant 127700) and Medical Research Council (Therapeutic Target Validation in Mental Health grant MR/S037675/1). Dr Zammit is supported by the NIHR Biomedical Research Centre at University Hospitals Bristol National Health Service (NHS) Foundation Trust and the University of Bristol. The UK Medical Research Council and Wellcome Trust (grant 102215/2/13/2) and the University of Bristol provide core support for the Avon Longitudinal Study of Parents and Children (ALSPAC). A comprehensive list of grants funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf); this research was specifically funded by the Wellcome Trust (grant 08426812/Z/07/Z), Wellcome Trust & MRC (grant 217065/Z/19/Z), and the MRC (grant MR/M006727/1).

Role of the Funder/Sponsor: The funding organizations had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: The views expressed in this publication are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care.

Additional Contributions: We are grateful to all the families who took part in this study, the midwives for their help in recruiting the families, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses.

Additional Information: Information about the data available in ALSPAC is available at http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/.

REFERENCES

1. Firth J, Siddiqi N, Koyanagi A, et al. *The Lancet Psychiatry* Commission: a blueprint for protecting physical health in people with mental illness. *Lancet Psychiatry*. 2019;6(8):675-712. doi:10.1016/S2215-0366(19)30132-4

2. Naylor C, Parsonage M, McDaid D, Knapp M, Fossey M, Galea A. *Long-term Conditions and Mental Health: the Cost of Co-morbidities*. The King's Fund; 2012.

3. Laursen TM, Plana-Ripoll O, Andersen PK, et al. Cause-specific life years lost among persons diagnosed with schizophrenia: is it getting better or worse? *Schizophr Res*. 2019;206:284-290. doi:10.1016/j.schres.2018.11.003

4. Plana-Ripoll O, Pedersen CB, Agerbo E, et al. A comprehensive analysis of mortality-related health metrics associated with mental disorders: a nationwide, register-based cohort study. *Lancet*. 2019;394(10211):1827-1835. doi:10.1016/S0140-6736(19)32316-5

5. Leucht S, Cipriani A, Spineli L, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet*. 2013;382(9896):951-962. doi:10.1016/S0140-6736(13)60733-3

6. Perry BI, McIntosh G, Weich S, Singh S, Rees K. The association between first-episode psychosis and abnormal glycaemic control: systematic review and meta-analysis. *Lancet Psychiatry*. 2016;3(11):1049-1058. doi:10.1016/S2215-0366(16)30262-0

7. Pillinger T, Beck K, Gobjila C, Donocik JG, Jauhar S, Howes OD. Impaired glucose homeostasis in first-episode schizophrenia: a systematic review

and meta-analysis. *JAMA Psychiatry*. 2017;74(3):261-269. doi:10.1001/jamapsychiatry.2016.3803

8. Penninx BW, Beekman AT, Honig A, et al. Depression and cardiac mortality: results from a community-based longitudinal study. *Arch Gen Psychiatry*. 2001;58(3):221-227. doi:10.1001/archpsyc.58.3.221

9. van Melle JP, de Jonge P, Spijkerman TA, et al. Prognostic association of depression following myocardial infarction with mortality and cardiovascular events: a meta-analysis. *Psychosom Med*. 2004;66(6):814-822. doi:10.1097/01.psy.0000146294.82810.9c

10. Kucukgoncu S, Kosir U, Zhou E, Sullivan E, Srihari VH, Tek C. Glucose metabolism dysregulation at the onset of mental illness is not limited to first episode psychosis: A systematic review and meta-analysis. *Early Interv Psychiatry*. 2019;13(5):1021-1031. doi:10.1111/eip.12749

11. Perry BI, Upthegrove R, Thompson A, et al. Dysglycaemia, inflammation and psychosis: findings from the UK ALSPAC Birth Cohort. *Schizophr Bull*. 2019;45(2):330-338. doi:10.1093/schbul/sby040

12. Mannan M, Mamun A, Doi S, Clavarino A. Prospective associations between depression and obesity for adolescent males and females—a systematic review and meta-analysis of longitudinal studies. *PLoS One*. 2016;11(6):e0157240. doi:10.1371/journal.pone.0157240

13. Buscot MJ, Thomson RJ, Juonala M, et al. Distinct child-to-adult body mass index trajectories are associated with different levels of adult cardiometabolic risk. *Eur Heart J*. 2018;39(24):2263-2270. doi:10.1093/eurheartj/ehy161

14. Tyrrell J, Mulugeta A, Wood AR, et al. Using genetics to understand the causal influence of higher BMI on depression. *Int J Epidemiol*. 2019;48(3):834-848. doi:10.1093/ije/dyy223

15. Li Z, Chen P, Chen J, et al. Glucose and insulin-related traits, type 2 diabetes and risk of schizophrenia: a mendelian randomization study. *EBioMedicine*. 2018;34:182-188. doi:10.1016/j.ebiom.2018.07.037

16. Boyd A, Golding J, Macleod J, et al. Cohort profile: the "children of the 90s"—the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol*. 2013;42(1):111-127. doi:10.1093/ije/dys064

17. Fraser A, Macdonald-Wallis C, Tilling K, et al. Cohort profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol*. 2013;42(1):97-110. doi:10.1093/ije/dys066

18. Northstone K, Lewcock M, Groom A, et al. The Avon Longitudinal Study of Parents and Children (ALSPAC): an update on the enrolled sample of index children in 2019. *Wellcome Open Res*. 2019;4:51. doi:10.12688/wellcomeopenres.15132.1

19. Harris PA, Taylor R, Minor BL, et al; REDCap Consortium. The REDCap consortium: building an international community of software platform partners. *J Biomed Inform*. 2019;95:103208. doi:10.1016/j.jbi.2019.103208

20. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42(2):377-381. doi:10.1016/j.jbi.2008.08.010

21. Sullivan SA, Kounali D, Cannon M, et al. A Population-based cohort study examining the incidence and impact of psychotic experiences from childhood to adulthood, and prediction of psychotic disorder. *Am J Psychiatry*. 2020;177(4):308-317. doi:10.1176/appi.ajp.2019.19060654

22. World Health Organization. Division of Mental Health. Schedules for clinical assessment in neuropsychiatry: version 2. World Health Organization. Published 1994. Accessed December 7, 2020. https://apps.who.int/iris/handle/10665/40356

23. Yung AR, Yuen HP, McGorry PD, et al. Mapping the onset of psychosis: the Comprehensive Assessment of At-Risk Mental States. *Aust N Z J Psychiatry*. 2005;39(11-12):964-971. doi:10.1080/j.1440-1614.2005.01714.x

24. Konings M, Bak M, Hanssen M, van Os J, Krabbendam L. Validity and reliability of the CAPE: a self-report instrument for the measurement of psychotic experiences in the general population. *Acta Psychiatr Scand*. 2006;114(1):55-61. doi:10.1111/j.1600-0447.2005.00741.x

25. Lewis G, Pelosi AJ, Araya R, Dunn G. Measuring psychiatric disorder in the community: a standardized assessment for use by lay interviewers. *Psychol Med*. 1992;22(2):465-486. doi:10.1017/S0033291700030415

26. Goodman R. Psychometric properties of the strengths and difficulties questionnaire. *J Am Acad Child Adolesc Psychiatry*. 2001;40(11):1337-1345. doi:10.1097/00004583-200111000-00015

27. Ram N, Grimm KJ. Growth mixture modeling: a method for identifying differences in longitudinal change among unobserved groups. *Int J Behav Dev*. 2009;33(6):565-576. doi:10.1177/0165025409343765

28. Holm S. A simple sequentially rejective multiple test procedure. *Scand J Stat*. 1979;6(2):65-70.

29. Asparouhov T, Muthén B. Auxiliary variables in mixture modeling: three-step approaches using Mplus. *Struct Equ Modeling*. 2014;21(3):329-341. doi:10.1080/10705511.2014.915181

30. Hackinger S, Prins B, Mamakou V, et al. Evidence for genetic contribution to the increased risk of type 2 diabetes in schizophrenia. *Transl Psychiatry*. 2018;8(1):252. doi:10.1038/s41398-018-0304-6

31. Tomasik J, Lago SG, Vázquez-Bourgon J, et al. Association of insulin resistance with schizophrenia polygenic risk score and response to antipsychotic treatment. *JAMA Psychiatry*. 2019;76(8):864-867. doi:10.1001/jamapsychiatry.2019.0304

32. Barker DJ, Gluckman PD, Godfrey KM, Harding JE, Owens JA, Robinson JS. Fetal nutrition and cardiovascular disease in adult life. *Lancet*. 1993;341(8850):938-941. doi:10.1016/0140-6736(93)91224-A

33. Gariepy G, Nitka D, Schmitz N. The association between obesity and anxiety disorders in the population: a systematic review and meta-analysis. *Int J Obes (Lond)*. 2010;34(3):407-419. doi:10.1038/ijo.2009.252

34. Luppino FS, de Wit LM, Bouvy PF, et al. Overweight, obesity, and depression: a systematic review and meta-analysis of longitudinal studies. *Arch Gen Psychiatry*. 2010;67(3):220-229. doi:10.1001/archgenpsychiatry.2010.2

35. Zhang T, Xu J, Li S, et al. Trajectories of childhood BMI and adult diabetes: the Bogalusa Heart Study. *Diabetologia*. 2019;62(1):70-77. doi:10.1007/s00125-018-4753-5

36. Rolland-Cachera MF, Péneau S. Growth trajectories associated with adult obesity. *World Rev Nutr Diet*. 2013;106:127-134.

37. Barker DJ, Osmond C, Forsén TJ, Kajantie E, Eriksson JG. Trajectories of growth among children who have coronary events as adults. *N Engl J Med*. 2005;353(17):1802-1809. doi:10.1056/NEJMoa044160

38. Mattsson M, Maher GM, Boland F, Fitzgerald AP, Murray DM, Biesma R. Group-based trajectory modelling for BMI trajectories in childhood: a systematic review. *Obes Rev*. 2019;20(7):998-1015. doi:10.1111/obr.12842

39. Lee KS, Vaillancourt T. Longitudinal associations among bullying by peers, disordered eating behavior, and symptoms of depression during adolescence. *JAMA Psychiatry*. 2018;75(6):605-612. doi:10.1001/jamapsychiatry.2018.0284

40. Yilmaz Z, Gottfredson NC, Zerwas SC, Bulik CM, Micali N. Developmental premorbid body mass index trajectories of adolescents with eating disorders in a longitudinal population cohort. *J Am Acad Child Adolesc Psychiatry*. 2019;58(2):191-199. doi:10.1016/j.jaac.2018.11.008

41. Welch E, Jangmo A, Thornton LM, et al. Treatment-seeking patients with binge-eating disorder in the Swedish national registers: clinical course and psychiatric comorbidity. *BMC Psychiatry*. 2016;16:163. doi:10.1186/s12888-016-0840-7

42. Schiller CE, Meltzer-Brody S, Rubinow DR. The role of reproductive hormones in postpartum depression. *CNS Spectr*. 2015;20(1):48-59. doi:10.1017/S1092852914000480

43. Dalal PK, Agarwal M. Postmenopausal syndrome. *Indian J Psychiatry*. 2015;57(suppl 2):S222-S232. doi:10.4103/0019-5545.161483

44. Soares CN, Zitek B. Reproductive hormone sensitivity and risk for depression across the female life cycle: a continuum of vulnerability? *J Psychiatry Neurosci*. 2008;33(4):331-343.

45. Li W, Liu Q, Deng X, Chen Y, Liu S, Story M. Association between obesity and puberty timing: a systematic review and meta-analysis. *Int J Environ Res Public Health*. 2017;14(10):E1266. doi:10.3390/ijerph14101266

46. Bell JA, Carslake D, Wade KH, et al. Influence of puberty timing on adiposity and cardiometabolic traits: a mendelian randomisation study. *PLoS Med*. 2018;15(8):e1002641. doi:10.1371/journal.pmed.1002641

47. Lewis G, Ioannidis K, van Harmelen AL, et al. The association between pubertal status and depressive symptoms and diagnoses in adolescent females: a population-based cohort study. *PLoS One*. 2018;13(6):e0198804. doi:10.1371/journal.pone.0198804

48. Timonen M, Rajala U, Jokelainen J, Keinänen-Kiukaanniemi S, Meyer-Rochow VB, Räsänen P. Depressive symptoms and insulin resistance in young adult males: results from the Northern Finland 1966 birth cohort. *Mol Psychiatry*. 2006;11(10):929-933. doi:10.1038/sj.mp.4001838

49. Perry BI, Khandaker GM, Marwaha S, et al. Insulin resistance and obesity, and their association with depression in relatively young people: findings from a large UK birth cohort. *Psychol Med*. 2020;50(4):556-565. doi:10.1017/S0033291719000308

50. Cham H, Reshetnyak E, Rosenfeld B, Breitbart W. Full information maximum likelihood estimation for latent variable interactions with incomplete indicators. *Multivariate Behav Res*. 2017;52(1):12-30. doi:10.1080/00273171.2016.1245600

51. Little TD, Jorgensen TD, Lang KM, Moore EW. On the joys of missing data. *J Pediatr Psychol*. 2014;39(2):151-162. doi:10.1093/jpepsy/jst048

Appendix C

Appendix C Tables

Appendix C Table 1: Full LDSC Results for Schizophrenia, Cardiometabolic and Inflammatory Traits

Trait 1	Trait 2	RG	SE	Z-Score	H_2	H_2 SE	*p*-value
Schizophrenia	Fasting Insulin	-0.0285	0.038	-0.733	0.079	0.008	0.463
Schizophrenia	HDL	-0.074	0.032	1.412	0.127	0.024	0.158
Schizophrenia	Triglycerides	0.035	0.029	-0.858	0.168	0.030	0.391
Schizophrenia	LDL	-0.023	0.022	-0.977	0.201	0.049	0.329
Schizophrenia	FPG	-0.048	0.037	-1.311	0.075	0.015	0.190
Schizophrenia	BMI	-0.091	0.015	-5.924	0.205	0.007	<0.001
Schizophrenia	T2D	-0.073	0.023	-3.14	0.043	0.002	0.002
Schizophrenia	Two Hour Glucose	-0.020	0.062	-0.33	0.030	0.010	0.743
Schizophrenia	HOMA-IR	-0.028	0.050	-0.56	0.038	0.007	0.578
Schizophrenia	HbA1C	-0.009	0.029	-0.29	0.074	0.009	0.771
Schizophrenia	CAD	0.029	0.019	1.52	0.108	0.008	0.129
Schizophrenia	CRP	-0.017	0.003	-0.635	0.232	0.009	0.205
Two Hour Glucose	T2DM	0.400	0.098	4.06	0.041	0.003	0.002
Two Hour Glucose	HOMA-IR	-0.094	0.179	-0.527	0.039	0.007	0.598
Two Hour Glucose	HbA1C	0.371	0.115	3.24	0.077	0.009	0.001
Two Hour Glucose	CAD	0.258	0.088	2.95	0.064	0.005	0.003
Two Hour Glucose	LDL	-0.001	0.078	-0.013	0.207	0.048	0.986
Two Hour Glucose	HDL	0.101	0.104	0.968	0.112	0.020	0.333
Two Hour Glucose	Fasting Insulin	0.057	0.119	0.476	0.083	0.009	0.640
Two Hour Glucose	FPG	0.027	0.156	0.172	0.075	0.016	0.863
Two Hour Glucose	BMI	-0.038	0.054	-0.711	0.203	0.006	0.477
Two Hour Glucose	CRP	0.026	0.096	0.271	0.087	0.019	0.786
Two Hour Glucose	Triglycerides	0.081	0.098	0.823	0.144	0.025	0.410
HOMA-IR	T2D	0.624	0.088	7.11	0.042	0.003	<0.001
HOMA-IR	HbA1C	0.195	0.081	2.40	0.078	0.009	0.016
HOMA-IR	CAD	0.222	0.071	3.12	0.065	0.005	0.002
HOMA-IR	LDL	0,021	0.056	0.357	0.209	0.045	0.722
HOMA-IR	HDL	-0.597	0.098	-6.122	0.112	0.019	<0.001
HOMA-IR	Fasting Insulin	1.000	0.079	15.078	0.037	0.008	<0.001
HOMA-IR	FPG	0.354	0.093	3.810	0.075	0.016	<0.001
HOMA-IR	BMI	0.667	0.075	8.844	0.304	0.063	<0.001
HOMA-IR	CRP	0.447	0.124	3.603	0.088	0.018	<0.001
HOMA-IR	Triglycerides	0.527	0.100	5.266	0.141	0.024	<0.001
T2D	HbA1C	0.466	0.044	10.66	0.075	0.009	<0.001
T2D	CAD	0.389	0.034	11.58	0.065	0.005	<0.001
T2D	LDL	0.063	0.041	1.531	0.201	0.049	0.126
T2D	HDL	-0.415	0.038	-11.05	0.109	0.014	<0.001
T2D	Fasting Insulin	0.556	0.056	10.00	0.041	0.003	<0.001
T2D	FPG	0.471	0.074	6.398	0.076	0.016	<0.001
T2D	BMI	0.554	0.028	19.735	0.205	0.007	<0.001
T2D	CRP	0.349	0.056	5.899	0.018	1.022	<0.001
T2D	Triglycerides	0.392	0.057	6.82	0.131	0.022	<0.001
CAD	HbA1C	0.253	0.035	7.234	0.075	0.009	<0.001
CAD	LDL	0.189	0.045	4.174	0.201	0.045	<0.001
CAD	HDL	-0.316	0.033	-0.696	0.108	0.016	<0.001

CAD	Fasting Insulin	0.305	0.050	6.145	0.107	0.008	<0.001
CAD	FPG	0.121	0.042	2.892	0.076	0.016	0.004
CAD	BMI	0.307	0.021	14.675	0.205	0.007	<0.001
CAD	CRP	0.242	0.060	4.050	0.089	0.018	<0.001
CAD	Triglycerides	0.283	0.032	6.925	0.131	0.023	<0.001
HbA1C	LDL	0.152	0.057	2.682	0.199	0.044	0.0073
HbA1C	HDL	-0.131	0.049	-2.670	0.105	0.018	0.0076
HbA1C	Fasting Insulin	0.182	0.068	2.647	0.076	0.009	0.008
HbA1C	FPG	0.522	0.084	6.235	0.076	0.016	<0.001
HbA1C	BMI	0.237	0.028	8.571	0.204	0.007	<0.001
HbA1C	CRP	0.225	0.056	4.064	0.087	0.019	<0.001
HbA1C	Triglycerides	0.167	0.054	3.082	0.133	0.024	0.0021
CRP	Fasting Insulin	0.353	0.087	4.085	0.088	0.019	<0.001
CRP	HDL	-0.325	0.062	-5.289	0.127	0.023	<0.001
CRP	Triglycerides	0.298	0.093	3.202	0.169	0.029	0.001
CRP	LDL	0.157	0.102	1.544	0.202	0.050	0.123
CRP	FPG	0.230	0.080	2.888	0.076	0.015	0.004
CRP	BMI	0.480	0.056	8.660	0.204	0.006	<0.001
Triglycerides	Fasting Insulin	0.415	0.099	4.211	0.138	0.025	<0.001
Triglycerides	HDL	-0.568	0.063	-9.007	0.131	0.023	<0.001
Triglycerides	LDL	0.381	0.053	7.15	0.201	0.047	<0.001
Triglycerides	FPG	0.144	0.092	1.490	0.075	0.016	0.139
Triglycerides	BMI	0.281	0.037	7.700	0.204	0.007	<0.001
HDL	LDL	-0.018	0.065	-0.272	0.209	0.047	0.786
HDL	FPG	-0.209	0.076	-2.759	0.075	0.016	0.006
HDL	BMI	-0.396	0.036	-11.034	0.204	0.007	<0.001
HDL	Fasting Insulin	-0.537	0.078	-6.908	0.104	0.021	<0.001
Fasting Insulin	LDL	0.079	0.054	1.480	0.212	0.047	0.139
Fasting Insulin	FPG	0.327	0.104	3.148	0.074	0.016	0.001
Fasting Insulin	BMI	0.587	0.043	13.690	0.204	0.007	<0.001

HDL=high-density lipoprotein; LDL=low-density lipoprotein; FPG=fasting plasma glucose; BMI=body mass index; T2D=type 2 diabetes mellitus; HOMA=homeostatic model assessment of insulin resistance; HbA1C=glycated haemoglobin; CAD=coronary artery disease; CRP=C-reactive protein.

Appendix C Table 2: Regions of Local Genetic Correlation Surpassing Bonferroni Evidential Threshold between Schizophrenia and Cardiometabolic and Inflammatory Traits

Chromosome	Start	End	Local Genetic Correlation (95% CI)	*p-value*
Fasting Insulin				
5	165642395	166847740	0.92 (0.66-0.97)	2.01E-12
7	11299198	12635461	1.00 (0.65-1.00)	5.99E-11
3	197075987	197946622	0.73 (0.42-0.98)	5.24E-09
7	28360309	31137289	0.70 (0.42-0.96)	1.47E-07
20	42680176	44839056	0.66 (0.41-0.87)	2.71E-07
1	186810023	188759945	0.62 (0.31-0.87)	4.87E-07
1	102898745	103914211	0.66 (0.30-0.92)	5.60E-07
12	43984474	46024229	0.62 (0.42-0.87)	5.92E-07
7	128778386	130422414	-0.90 (-1.00 - -0.54)	5.96E-07
4	180122043	182066807	1.00 (0.67-1.00)	8.76E-07
15	50008043	51677560	0.63 (0.37-0.77)	1.28E-06
17	41772087	43056905	-0.82 (-1.00 - -0.49)	1.71E-06
6	145319810	146665424	-0.81 (-1.00 - -0.38)	2.31E-06
2	147277162	150210292	-0.78 (-1.00 - -0.51)	2.43E-06
9	28224283	28811584	-0.59 (-0.85 - -0.24)	4.33E-06
8	116096495	119685457	0.55 (0.31-1.00)	6.45E-06
9	85440801	86938196	0.77 (0.43-1.00)	7.78E-06
12	117087471	118135375	0.61 (0.42-0.88)	8.42E-06
4	27965868	29762208	0.83 (0.36-1.00)	9.24E-06
11	6322869	7436701	0.58 (0.32-1.00)	1.20E-05
13	54682864	55817131	0.58 (0.32-0.81)	1.37E-05
17	1172399	1928731	0.57 (0.31-0.76)	1.54E-05
2	58297315	60292000	0.40 (0.24-0.58)	1.95E-05
3	190226607	192343814	0.69 (0.38-1.00)	2.08E-05
1	74326907	76728135	0.77 (0.41-1.00)	2.15E-05
3	181511166	183769683	0.56 (0.30-0.68)	2.34E-05
17	45876022	47517400	-0.66 (-0.96 - -0.47)	2.60E-05
5	43983499	50163398	-0.51 (-0.75 - -0.37)	2.63E-05
18	39892648	42922106	-0.72 (-1.00 - -0.47)	2.75E-05
10	89127064	91013381	0.51 (0.27-0.58)	2.86E-05
Type 2 Diabetes				
4	154477641	155056126	0.49 (0.29-0.76)	3.2670e-06
5	178413464	179401244	0.45 (0.25-0.76)	5.9955e-06
7	68234074	69085364	0.66 (0.37-0.96)	7.1611e-06
14	72889615	76444767	0.42 (0.42-0.64)	7.8567e-06
7	1353067	2062398	-0.42 (-0.61- -0.33)	1.3872e-05
7	126869221	128778386	0.40 (0.22-0.72)	1.8553e-05
18	57630483	59020751	0.51 (0.27-0.48)	2.3705e-05
4	105305294	107501305	0.50 (0.27-0.50)	2.7390e-05
High-Density Lipoprotein				
6	30798168	31571218	0.56 (0.37-0.75)	9.3891e-09
12	122007651	124977980	-0.79 (-1.00 - -0.50)	1.4944e-07
7	71874885	73334602	0.53 (0.32-0.75)	1.3009e-06
6	31571218	32682664	0.39 (0.21-0.56)	1.4672e-05
Low-Density Lipoprotein				
6	31571218	32682664	0.52 (0.32-0.72)	4.47e-07
19	44744108	46102697	0.38 (0.20-0.56)	2.29e-05

Triglycerides				
12	122007651	124977980	0.57 (0.37-0.78)	2.9534e-08
15	58441366	59694116	0.350 (0.22-0.49)	8.1201e-07
10	100668400	102949239	-0.77 (-1.00 - -0.58)	4.4472e-06
4	103221356	105305294	-0.86 (-1.00 - -0.48)	1.1815e-05
3	135456906	137371083	0.77 (0.42-1.12)	1.4880e-05
Body Mass Index				
4	100678360	103221356	0.74 (0.58-1.00)	1.8800e-18
8	116096495	119685457	-0.71 (-0.87- -0.64)	1.5271e-17
16	29036613	31382943	-0.69 (-0.85- -0.42)	3.9790e-17
11	134205993	134946452	-0.85 (-1.00 - -0.44)	7.6067e-17
10	33707968	35109355	-0.76 (-0.95 - -0.60)	1.6559e-16
11	27020461	28481593	0.65 (0.49-0.70)	1.3360e-15
12	122007651	124977980	0.54 (0.41-1.00)	3.0322e-15
1	154770403	156336133	-0.64 (-0.81 - -0.44)	2.8648e-13
14	29972145	32383265	0.65 (0.45-1.00)	1.0969e-10
2	40281483	43309590	-0.69 (-0.90 - -0.62)	1.5763e-10
14	103012102	105001723	0.55 (0.38-1.00)	2.6666e-10
14	61680424	63790015	-0.74 (-0.97- -0.68)	4.7168e-10
22	19912358	22357325	-0.70 (-0.93 - -0.59)	7.4182e-10
18	51554175	55213838	-0.52 (-0.68 - -0.33)	9.0991e-10
2	144519484	146445570	-0.55 (-0.73 - -0.43)	2.0740e-09
11	130342575	131074612	-0.68 (-0.90 - -0.59)	2.6699e-09
4	103221356	105305294	0.48 (0.32-0.94)	8.0355e-09
9	121321537	122260297	-0.63 (-0.85- -0.51)	2.5874e-08
14	93132299	94325285	-0.61 (-0.83 - -0.54)	2.7201e-08
17	1172399	1928731	-0.62 (-0.84 - -0.49)	2.8907e-08
4	18841874	20544557	0.58 (0.38-1.00)	3.4517e-08
2	14335308	16329735	0.80 (0.51-1.00)	4.4638e-08
9	76973081	78900183	-1.00 (-1.00- -1.00)	4.7885e-08
2	147277162	150210292	-0.55 (-0.75-0.43)	4.8358e-08
5	58524622	60935907	-0.53 (-0.72 - -0.38)	5.0203e-08
5	152867774	153773088	-0.50 (-0.68 - -0.43)	5.8424e-08
17	72672203	74375560	0.71 (0.45-1.00)	6.5898e-08
5	155373505	156628700	0.85 (0.54-1.00)	7.4102e-08
6	108464380	110304247	-0.54 (-0.73 - -0.43)	9.8576e-08
1	38731847	40200567	-0.59 (-0.81 - -0.43)	1.2270e-07
11	12564229	13373124	-0.73 (-0.99 - -0.52)	1.2379e-07
18	45939732	47730584	-0.78 (-1.00 - -0.60)	1.7993e-07
4	43965045	45189157	-0.49 (-0.69 - -0.22)	2.0291e-07
1	1892607	3582736	0.58 (0.36-1.00)	2.0822e-07
8	143044914	144236881	-0.46 (-0.64 - -0.72)	2.1199e-07
14	32383265	34846251	0.56 (0.34-1.00)	2.8869e-07
10	102949239	104380410	-0.46 (-0.64- -0.40)	4.2282e-07
19	13471127	14486347	-0.62 (-0.85 - -0.73)	4.2562e-07
13	55817131	57554217	-0.70 (-0.97 - -0.45)	4.6785e-07
3	30717955	32351715	-0.73 (-1.00 - -0.42)	5.0177e-07
13	58410626	59302271	-0.45 (-0.63- -0.38)	6.9698e-07
8	143044914	144236881	-0.46 (-0.64 - -0.32)	2.1199e-07
14	32383265	34846251	0.56 (0.35-1.00)	2.8869e-07
10	102949239	104380410	-0.46 (-0.64- -0.50)	4.2282e-07
19	13471127	14486347	-0.62 (-0.85 - -0.73)	4.2562e-07
13	55817131	57554217	-0.70 (-0.98 - -0.55)	4.6785e-07
3	30717955	32351715	-0.73 (-1.00 - -0.54)	5.0177e-07
13	58410626	59302271	-0.44 (-0.65 - -0.38)	6.9698e-07
2	229370787	231843389	-0.66 (-0.93 - -0.48)	1.0193e-06
8	9640787	10463197	0.49 (0.29-0.94)	1.3453e-06
6	97842284	100630146	-0.54 (-0.76 - -0.31)	1.3989e-06

13	63971559	65200602	0.72 (0.43-1.00)	1.4903e-06
15	76398624	78516053	-0.52 (-0.73 - -0.42)	1.5391e-06
11	30141357	32276901	-0.44 (-0.62 - -0.38)	1.5419e-06
17	1928731	3702312	-0.57 (-0.81 - -0.39)	1.8773e-06
19	30727954	32746520	-0.45 (-0.64 - -0.35)	1.9589e-06
8	4480476	5146927	0.62 (0.36-1.71)	1.9798e-06
11	17578402	19569535	-0.68 (-0.96- -0.81)	2.8624e-06
6	13209388	14802924	-0.59 (-0.84 – 0.43)	3.0978e-06
1	106087842	108409665	-0.55 (-0.78 - -0.33)	3.2491e-06
8	2573279	3392926	-0.74 (-1.00 - -0.44)	3.3313e-06
1	97885249	99800604	-0.42 (-0.59 - -0.37)	3.4856e-06
15	73628714	76398624	-0.47 (-0.66- -0.29)	4.1370e-06
2	191973357	195861164	-0.59 (-0.84 - -0.44)	4.6666e-06
18	22996651	24026191	0.69 (0.40-1.00)	4.7871e-06
1	224938520	226810860	-0.70 (-1.00 - -0.55)	4.7961e-06
7	128778386	130422414	0.69 (0.39-1.00)	4.8353e-06
9	107581749	109298754	0.59 (0.34-1.00)	5.1291e-06
11	55082657	58457495	0.55 (0.32-1.00)	5.5468e-06
6	24852275	25684587	-0.56 (-0.81 - -0.36)	7.0368e-06
20	13689864	15958359	-0.64 (-0.92 - -0.43)	7.3961e-06
16	13154437	14464002	-0.60 (-0.86 - -0.48)	8.3274e-06
9	84211233	85440801	-0.61 (-0.88 - -0.50)	8.4905e-06
16	20150571	22448904	0.44 (0.24-0.86)	8.5474e-06
4	58935008	60741087	0.65 (0.36-1.00)	8.8613e-06
18	5834180	7090485	-0.60 (-0.86 - -0.40)	8.9776e-06
9	31310383	32019368	0.66 (0.37-1.00)	9.5428e-06
17	27334244	29786491	-0.48 (-0.69 - -0.37)	1.0430e-05
3	131836516	133252173	-0.58 (-0.84 - -0.44)	1.0541e-05
6	25684587	26791233	-0.33 (-0.48 - -0.22)	1.2825e-05
8	144236881	146303867	-0.49 (-0.71 - -0.31)	1.3140e-05
9	126971887	129059665	-0.54 (-0.78 - -0.42)	1.6506e-05
2	209941529	212379518	-0.47 (-0.68 - -0.30)	1.6960e-05
15	48136048	50008043	-0.71 (-1.00 - -0.59)	1.7653e-05
7	49212278	51675322	-0.48 (-0.70 - -0.30)	1.7793e-05
14	59448336	61680424	-0.56 (-0.82 - -0.41)	1.8802e-05
5	165642395	166847740	-0.61 (-0.89 - -0.52)	1.9529e-05
13	38878163	41069263	-0.52 (-0.76 - -0.43)	2.0126e-05
4	6773043	7539692	-0.76 (-1.00 - -0.51)	2.0157e-05
6	42038721	43756169	-0.56 (-0.81 - -0.40)	2.1629e-05
Coronary Artery Disease				
10	104380410	106695048	0.47 (0.33-0.77)	1.8131e-12
15	90475551	92164392	-0.55 (-0.72 - -0.29)	3.2537e-10
7	128778386	130422414	0.64 (0.39-0.66)	8.4244e-07
22	19912358	22357325	-0.39 (-0.58 - -0.25)	2.7840e-05
C-Reactive Protein				
7	87825004	90661784	0.99 (0.88-1.00)	2.3160e-08
1	153180829	154770403	-0.40 (-0.56- -0.11)	7.0758e-07
9	130055510	132165470	-0.96 (-1.00 - -0.90)	3.8599e-06
2	60292000	62429044	-0.87 (-1.00 - -0.66)	1.2429e-05
4	43965045	45189157	-0.78 (-1.00 - -0.51)	2.8390e-05

Appendix C Table 3: Sensitivity Analysis Results for Schizophrenia, Cardiometabolic and Inflammatory Traits: Modifying Prior Configurations and Regional/Alignment Thresholds

Candidate SNP	Colocalized Traits	PP_{coloc}[1]	$PP_{explained}$[2]	N SNPs[3]	Prior Prob 2	Reg/Align Threshold
rs8192675	SCZ, T2D, CRP, BMI	0.9299	0.5033	919	0.95	0.5
rs8192675	SCZ, T2D, CRP, BMI	0.9299	0.5033	919	0.95	0.6
rs8192675	SCZ, T2D, CRP, BMI	0.9299	0.5033	919	0.95	0.7
rs8192675	SCZ, T2D, CRP, BMI	0.9299	0.5033	919	0.95	0.8
rs8192675	SCZ, T2D, CRP, BMI	0.9299	0.5033	919	0.95	0.9
rs8192675	SCZ, T2D, CRP, BMI	0.8434	0.5033	919	0.98	0.5
rs8192675	SCZ, T2D, CRP, BMI	0.8434	0.5033	919	0.98	0.6
rs8192675	SCZ, T2D, CRP, BMI	0.8434	0.5033	919	0.98	0.7
rs8192675	SCZ, T2D, CRP, BMI	0.8434	0.5033	919	0.98	0.8
rs8192675	SCZ, CRP, BMI	0.9093	0.351	919	0.98	0.9
rs8192675	SCZ, T2D, CRP, BMI	0.7261	0.5033	919	0.99	0.5
rs8192675	SCZ, T2D, CRP, BMI	0.7261	0.5033	919	0.99	0.6
rs8192675	SCZ, T2D, CRP, BMI	0.7261	0.5033	919	0.99	0.7
rs8192675	SCZ, CRP, BMI	0.8326	0.351	919	0.99	0.8
rs8192675	SCZ, BMI	0.8957	0.6095	919	0.99	0.9
rs8192675	SCZ, BMI	0.4097	0.6095	919	0.999	0.5
rs340874	SCZ, T2D	0.5437	0.6637	1324	0.95	0.5
rs340874	SCZ, T2D	0.5437	0.6637	1324	0.95	0.6
rs340874	SCZ, T2D	0.5437	0.6637	1324	0.95	0.7
rs2108349	SCZ, FI	0.6005	0.8826	1272	0.95	0.5
rs2108349	SCZ, FI	0.6005	0.8826	1272	0.95	0.6
rs2108349	SCZ, FI	0.6005	0.8826	1272	0.95	0.7
rs2108349	SCZ, FI	0.6005	0.8826	1272	0.98	0.5
rs2108349	SCZ, FI	0.6005	0.8826	1272	0.98	0.6
rs17514846	SCZ, CAD	0.9989	1	1071	0.95	0.5
rs17514846	SCZ, CAD	0.9989	1	1071	0.95	0.6
rs17514846	SCZ, CAD	0.9989	1	1071	0.95	0.7
rs17514846	SCZ, CAD	0.9989	1	1071	0.95	0.8
rs17514846	SCZ, CAD	0.9989	1	1071	0.95	0.9
rs17514846	SCZ, CAD	0.9971	1	1071	0.98	0.5
rs17514846	SCZ, CAD	0.9971	1	1071	0.98	0.6
rs17514846	SCZ, CAD	0.9971	1	1071	0.98	0.7
rs17514846	SCZ, CAD	0.9971	1	1071	0.98	0.8
rs17514846	SCZ, CAD	0.9971	1	1071	0.98	0.9
rs17514846	SCZ, CAD	0.9943	1	1071	0.99	0.5
rs17514846	SCZ, CAD	0.9943	1	1071	0.99	0.6
rs17514846	SCZ, CAD	0.9943	1	1071	0.99	0.7
rs17514846	SCZ, CAD	0.9943	1	1071	0.99	0.8
rs17514846	SCZ, CAD	0.9943	1	1071	0.99	0.9
rs17514846	SCZ, CAD	0.9457	1	1071	0.999	0.5
rs17514846	SCZ, CAD	0.9457	1	1071	0.999	0.6
rs17514846	SCZ, CAD	0.9457	1	1071	0.999	0.7
rs17514846	SCZ, CAD	0.9457	1	1071	0.999	0.8
rs17514846	SCZ, CAD	0.9457	1	1071	0.999	0.9
rs13107325	SCZ, TG, T2D, HDL, BMI	0.8569	1	936	0.95	0.5
rs13107325	SCZ, TG, T2D, HDL, BMI	0.8569	1	936	0.95	0.6
rs13107325	SCZ, TG, T2D, HDL, BMI	0.8569	1	936	0.95	0.7
rs13107325	SCZ, TG, T2D, HDL, BMI	0.8569	1	936	0.95	0.8
rs13107325	SCZ, TG, HDL, BMI	0.9477	1	936	0.95	0.9
rs13107325	SCZ, TG, T2D, HDL, BMI	0.7092	1	936	0.98	0.5
rs13107325	SCZ, TG, T2D, HDL, BMI	0.7092	1	936	0.98	0.6

rs13107325	SCZ, TG, T2D, HDL, BMI	0.7092	1	936	0.98	0.7
rs13107325	SCZ, TG, HDL, BMI	0.8807	1	936	0.98	0.8
rs13107325	SCZ, TG, HDL, BMI	0.8807	1	936	0.98	0.9
rs13107325	SCZ, TG, T2D, HDL, BMI	0.5423	1	936	0.99	0.5
rs13107325	SCZ, TG, T2D, HDL, BMI	0.5423	1	936	0.99	0.6
rs13107325	SCZ, TG, HDL, BMI	0.7853	1	936	0.99	0.7
rs13107325	SCZ, TG, HDL, BMI	0.7853	1	936	0.99	0.8
rs13107325	SCZ, HDL, BMI	1	1	936	0.99	0.9
rs13107325	SCZ, HDL, BMI	1	1	936	0.999	0.5
rs13107325	SCZ, HDL, BMI	1	1	936	0.999	0.6
rs13107325	SCZ, HDL, BMI	1	1	936	0.999	0.7
rs13107325	SCZ, HDL, BMI	1	1	936	0.999	0.8
rs13107325	SCZ, HDL, BMI	1	1	936	0.999	0.9
rs3814883	SCZ, BMI	0.9885	0.9964	193	0.95	0.5
rs3814883	SCZ, BMI	0.9885	0.9964	193	0.95	0.6
rs3814883	SCZ, BMI	0.9885	0.9964	193	0.95	0.7
rs3814883	SCZ, BMI	0.9885	0.9964	193	0.95	0.8
rs3814883	SCZ, BMI	0.9885	0.9964	193	0.95	0.9
rs3814883	SCZ, BMI	0.9717	0.9964	193	0.98	0.5
rs3814883	SCZ, BMI	0.9717	0.9964	193	0.98	0.6
rs3814883	SCZ, BMI	0.9717	0.9964	193	0.98	0.7
rs3814883	SCZ, BMI	0.9717	0.9964	193	0.98	0.8
rs3814883	SCZ, BMI	0.9717	0.9964	193	0.98	0.9
rs3814883	SCZ, BMI	0.9449	0.9964	193	0.99	0.5
rs3814883	SCZ, BMI	0.9449	0.9964	193	0.99	0.6
rs3814883	SCZ, BMI	0.9449	0.9964	193	0.99	0.7
rs3814883	SCZ, BMI	0.9449	0.9964	193	0.99	0.8
rs3814883	SCZ, BMI	0.9449	0.9964	193	0.99	0.9
rs3814883	SCZ, BMI	0.6316	0.9964	193	0.999	0.5
rs3814883	SCZ, BMI	0.6316	0.9964	193	0.999	0.6
rs12782894	SCZ, BMI	0.8834	0.6847	1255	0.95	0.5
rs12782894	SCZ, BMI	0.8834	0.6847	1255	0.95	0.6
rs12782894	SCZ, BMI	0.8834	0.6847	1255	0.95	0.7
rs12782894	SCZ, BMI	0.8834	0.6847	1255	0.95	0.8
rs12782894	SCZ, BMI	0.8834	0.6847	1255	0.95	0.9
rs12782894	SCZ, BMI	0.7447	0.6847	1255	0.98	0.5
rs12782894	SCZ, BMI	0.7447	0.6847	1255	0.98	0.6
rs12782894	SCZ, BMI	0.7447	0.6847	1255	0.98	0.7
rs12782894	SCZ, BMI	0.7447	0.6847	1255	0.98	0.8
rs12782894	SCZ, BMI	0.5787	0.6847	1255	0.99	0.5
rs12782894	SCZ, BMI	0.5787	0.6847	1255	0.99	0.6
rs6265	SCZ, TG, CAD, CRP, BMI	0.4796	0.8429	925	0.95	0.5
rs6265	SCZ, TG, CAD, CRP, BMI	0.4796	0.8429	925	0.95	0.6
rs6265	SCZ, CAD, CRP, BMI	0.8607	0.7491	925	0.95	0.7
rs6265	SCZ, CAD, CRP, BMI	0.8607	0.7491	925	0.95	0.8
rs6265	SCZ, CAD, BMI	0.9552	0.8067	925	0.95	0.9
rs6265	SCZ, CAD, CRP, BMI	0.7111	0.7491	925	0.98	0.5
rs6265	SCZ, CAD, CRP, BMI	0.7111	0.7491	925	0.98	0.6
rs6265	SCZ, CAD, CRP, BMI	0.7111	0.7491	925	0.98	0.7
rs6265	SCZ, CAD, BMI	0.8951	0.8067	925	0.98	0.8
rs6265	SCZ, CAD, BMI	0.8951	0.8067	925	0.98	0.9
rs6265	SCZ, CAD, CRP, BMI	0.5406	0.7491	925	0.99	0.5
rs6265	SCZ, CAD, CRP, BMI	0.5406	0.7491	925	0.99	0.6
rs6265	SCZ, CAD, BMI	0.8071	0.8067	925	0.99	0.7
rs6265	SCZ, CAD, BMI	0.8071	0.8067	925	0.99	0.8
rs3800229	SCZ, BMI	0.8889	0.9519	872	0.95	0.5

rs3800229	SCZ, BMI	0.8889	0.9519	872	0.95	0.6
rs3800229	SCZ, BMI	0.8889	0.9519	872	0.95	0.7
rs3800229	SCZ, BMI	0.8889	0.9519	872	0.95	0.8
rs3800229	SCZ, BMI	0.8889	0.9519	872	0.95	0.9
rs3800229	SCZ, BMI	0.7546	0.9519	872	0.98	0.5
rs3800229	SCZ, BMI	0.7546	0.9519	872	0.98	0.6
rs3800229	SCZ, BMI	0.7546	0.9519	872	0.98	0.7
rs3800229	SCZ, BMI	0.7546	0.9519	872	0.98	0.8
rs3800229	SCZ, BMI	0.5909	0.9519	872	0.99	0.5
rs3800229	SCZ, BMI	0.5909	0.9519	872	0.99	0.6
rs3800229	SCZ, BMI	0.5909	0.9519	872	0.99	0.7
rs2239647	SCZ, T2D, BMI	0.7872	0.6625	1584	0.95	0.6
rs2239647	SCZ, T2D, BMI	0.7872	0.6625	1584	0.95	0.7
rs2239647	SCZ, T2D, BMI	0.7872	0.6625	1584	0.95	0.8
rs2239647	SCZ, BMI	0.9883	0.6317	1584	0.95	0.9
rs2239647	SCZ, T2D, BMI	0.5916	0.6625	1584	0.98	0.5
rs2239647	SCZ, T2D, BMI	0.5916	0.6625	1584	0.98	0.6
rs2239647	SCZ, BMI	0.9712	0.6317	1584	0.98	0.7
rs2239647	SCZ, BMI	0.9712	0.6317	1584	0.98	0.8
rs2239647	SCZ, BMI	0.9712	0.6317	1584	0.98	0.9
rs2239647	SCZ, BMI	0.944	0.6317	1584	0.99	0.5
rs2239647	SCZ, BMI	0.944	0.6317	1584	0.99	0.6
rs2239647	SCZ, BMI	0.944	0.6317	1584	0.99	0.7
rs2239647	SCZ, BMI	0.944	0.6317	1584	0.99	0.8
rs2239647	SCZ, BMI	0.944	0.6317	1584	0.99	0.9
rs2239647	SCZ, BMI	0.6234	0.6317	1584	0.999	0.5
rs2239647	SCZ, BMI	0.6234	0.6317	1584	0.999	0.6
rs11191514	SCZ, CAD, BMI	0.7651	0.297	710	0.95	0.5
rs11191514	SCZ, CAD, BMI	0.7651	0.297	710	0.95	0.6
rs11191514	SCZ, CAD, BMI	0.7651	0.297	710	0.95	0.7
rs11191514	SCZ, CAD, BMI	0.5695	0.297	710	0.98	0.5
rs6031855	SCZ, BMI	0.5877	0.2771	990	0.95	0.5
rs6031855	SCZ, BMI	0.5877	0.2771	990	0.95	0.6
rs6031855	SCZ, BMI	0.5877	0.2771	990	0.95	0.7
rs6031855	SCZ, BMI	0.3224	0.2771	990	0.98	0.5

Appendix C Table 4: SNPs used as instruments for fasting insulin, triglycerides and high-density lipoprotein in MR Analysis

Fasting Insulin
rs10195252
rs2126259
rs2943645
rs308971
rs3822072
rs459193
rs4846565
rs4865796
rs731839

Triglycerides
rs1011685
rs10195252
rs132985
rs2699429
rs3861397
rs7973683
rs731839
rs4804311
rs2943645

HDL
rs1011685
rs10195252
rs2126259
rs2745353
rs2943645
rs3822072
rs3861397
rs459193
rs4804311
rs7973683
rs4976033
rs683135
rs731839
rs972283

Appendix C Table 5: SNPs used as instruments for fasting plasma glucose in MR Analysis

rs10276674	rs11603334	rs2191349	rs4869272	rs7644261
rs10830963	rs11605924	rs2524299	rs560887	rs780093
rs10974438	rs11708067	rs2908282	rs6113722	rs882020
rs11039138	rs17747324	rs4148804	rs7173964	rs983309
rs11195502	rs11558471			

Appendix C Table 6: SNPs used as instruments for type 2 diabetes mellitus in MR Analysis

rs1060105	rs1801212	rs328	rs60980157	rs7607980
rs1127787	rs1801282	rs35169799	rs665268	rs781831
rs1169288	rs2032844	rs35658696	rs6762208	rs9379084
rs1260326	rs2073721	rs35720761	rs72928978	rs9891146
rs13266634	rs2276853	rs3764002	rs738409	rs2307111
rs140386498	rs2296172	rs5219	rs7572857	rs28265
rs1800437	rs1800961	rs56200889	rs58542926	

Appendix C Table 7: SNPs used as instruments for body mass index in MR Analysis

rs1000940	rs13107325	rs205262	rs6091540	rs9540493
rs10132280	rs13130484	rs2060604	rs6465468	rs9579083
rs1016287	rs13191362	rs2112347	rs6477694	rs977747
rs10182181	rs13201877	rs2176040	rs6567160	rs9926784
rs10733682	rs13329567	rs2176598	rs657452	rs2033732
rs10840100	rs1421085	rs2183825	rs6804842	rs943005
rs11030104	rs1441264	rs2245368	rs7138803	
rs11165643	rs1460676	rs2365389	rs7144011	
rs11663558	rs14810	rs2820292	rs7239883	
rs11672660	rs1516725	rs2890652	rs7531118	
rs1167827	rs1528435	rs3736485	rs7550711	
rs11727676	rs16851483	rs3817334	rs7599312	
rs12286929	rs17001654	rs3849570	rs7715256	
rs12429545	rs17066856	rs3888190	rs7899106	
rs12448257	rs17094222	rs4740619	rs7903146	
rs12940622	rs17203016	rs4889606	rs879620	
rs12986742	rs17381664	rs492400	rs9304665	
rs13021737	rs17724992	rs4981693	rs9374842	
rs13078960	rs1928295	rs543874	rs9400239	

Appendix C Table 8: SNPs used as instruments for glucose tolerance in MR Analysis

rs1019503	rs11717195	rs12255372	rs6975024
rs11672660	rs11782386	rs6547829	

Appendix C Table 9: SNPs used as instruments for low density lipoprotein in MR Analysis

rs10195252	rs13277801	rs2328223	rs364585	rs676388
rs10490626	rs1367117	rs2390536	rs3757354	rs6818397
rs10832962	rs1408272	rs2419604	rs3780181	rs6882076
rs10893499	rs1564348	rs247616	rs4253776	rs6909746
rs10903129	rs16831243	rs2495495	rs4530754	rs7254892
rs112201728	rs16891156	rs2587534	rs4722551	rs72902576
rs11563251	rs17404153	rs2642438	rs4942486	rs7534572
rs11591147	rs174583	rs267733	rs4970712	rs7551981
rs1169288	rs1800961	rs2710642	rs5763662	rs75687619
rs12066643	rs1801689	rs2737252	rs579459	rs7640978
rs1250229	rs1883025	rs2738459	rs6016373	rs7832643
rs12721109	rs2000999	rs2886232	rs6065311	rs8017377
rs12748152	rs2030746	rs2954029	rs646776	rs964184
rs12916	rs2073547	rs2965157	rs6504872	rs9875338
rs13206249	rs2228603	rs314253	rs6511720	rs9987289
rs6709904	rs2315065	rs3184504	rs6544713	

Appendix C Table 10: SNPs used as instruments for glycated haemoglobin in MR Analysis

rs1046896	rs12621844	rs17747324	rs423117	rs6474359
rs10774625	rs13134327	rs1800562	rs4607517	rs7040409
rs11248914	rs13266634	rs2246434	rs4737009	rs7616006
rs11603334	rs1387153	rs2383208	rs4745982	rs8192675
rs11708067	rs1547247	rs267738	rs560887	rs855791
rs11964178	rs17509001	rs2979422	rs579459	rs9818758
rs12368284	rs17533903	rs3782123	rs592423	rs9914988
rs9935401				

Appendix C Table 11: SNPs used as instruments for leptin in MR Analysis

rs900400	rs6738627
rs6071166	rs780093

Appendix C Table 12: Inflammation-related SNPs for fasting insulin in MR Analysis

SNP	Inflammation-Related Pleiotropy	Effect Allele
rs2126259[a]	CRP, Neutrophil Count, Granulocyte Count, Basophil Count, Myeloid White Cell Count	T
rs731839	Lymphocyte Count, White Blood Cell Count, Neutrophil Count	G
rs10195252	Lymphocyte Count, Neutrophil % of White Cells,	T
rs308971	T-Cell Surface Protein CD3 Epsilon Chain	G
rs3822072	CD32, Lymphocyte Count, Neutrophil Count	A

[a]Genome-Wide Significance Inflammation-Related SNP; CRP=C-reactive protein

Appendix C Table 13: Inflammation-related SNPs for triglycerides in MR Analysis

SNP	Inflammation-Related Pleiotropy	Effect Allele
rs10195252	Lymphocyte Count, Neutrophil % of White cells,	T
rs731839	Lymphocyte Count, White Blood Cell Count, Neutrophil Count	G
rs1011685	Eosinophil Count, Granulocyte Count	C
rs3861397	Eosinophil % Granulocytes, C-X-C Motif Chemokine 14	G

Appendix C Table 14: Inflammation-related SNPs for high-density lipoprotein in MR Analysis

SNP	Inflammation-Related Pleiotropy	Effect Allele
rs2126259[a]	CRP, Neutrophil Count, Granulocyte Count, Basophil Count, Myeloid White Cell Count	T
rs731839	Lymphocyte Count, White Blood Cell Count, Neutrophil Count	G
rs1011685	Eosinophil % of White Cells, Neutrophil % of Granulocytes, Eosinophil Count	C
rs3822072	CD32, Lymphocyte Count, Neutrophil Count	A
rs10195252	Lymphocyte Count, Neutrophil % of White cells,	T
rs3861397	Eosinophil % Granulocytes, C-X-C Motif Chemokine 14	G
rs2745353	Lymphocyte Count	T

[a]Genome-Wide Significance Inflammation-Related SNP

Appendix C Table 15: Inflammation-related SNPs for low-density lipoprotein in MR Analysis

SNP	Inflammation-Related Pleiotropy	Effect Allele
rs1169288[a]	CRP	C
rs17404153[a]	C-C Motif Chemokine 21	G
rs174583[a]	Neutrophil Count, Basophil Count, White Cell Count, Eosinophil Count	C
rs1800961[a]	CRP, Neutrophil Count, Basophil Count	C
rs2642438[a]	Granulocyte% White Cells Monocyte% White Cells	G
rs2886232[a]	C-C Motif Chemokine 22	T
rs2954029[a]	Neutrophil Count, Eosinophil Count	A
rs3184504[a]	Eosinophil Count, Basophil Count, Lymphocyte Count, WCC, Neutrophil Count, IL-2b	C
rs4970712[a]	Monocyte Count, Granulocyte Count, Lymphocyte Count	C
rs579459[a]	IL-3a, Neutrophil Count, Eosinophil Count, WCC, IL-6	C
rs646776[a]	CRP	T
rs9875338[a]	Monocyte Count, Neutrophil Count	G
rs9987289[a]	CRP, Neutrophil Count, Basophil Count	G
rs964184	Eosinophil Count, Neutrophil% Granulocytes	G
rs75687619	CRP	T
rs4253776	Eosinophil% White Cells, Eosinophil% Granulocytes	G
rs2228603	CRP, Lymphocyte Count	C
rs1408272	Monocyte Count	T
rs10195252	Lymphocyte Count, Neutrophil Count	T
rs12748152	Basophil Count, Eosinophil Count, Eosinophil% Granulocytes, Neutrophil% Granulocytes	T
rs2000999	Granulocyte% White Cells	A
rs2737252	Monocyte Count	G
rs314253	CD4:%Act(DR+38+)	T
rs676388	Basophil Count	C
rs6882076	Lymphocyte Count, WCC	C

[a]Genome-Wide Significant Inflammation-Related SNP; CRP=C-reactive protein; WCC=White Cell Count; IL-=interleukin-.

Appendix C Table 16: Inflammation-related SNPs for fasting plasma glucose in MR Analysis

SNP	Inflammation-Related Pleiotropy	Effect Allele
rs780093[a]	Monocyte Count, Basophils, Neutrophil%	T
rs983309[a]	Neutrophil%, Granulocytes	T
rs1130391138	Lymphocyte% White Cells	A
rs2524299	Basophil Count, Neutrophil Count, Granulocyte Count % White Cells, White Cell Count, Eosinophil Count, Monocyte Count	A

[a]Genome-Wide Significant Inflammation-Related SNPs

Appendix C Table 17: Inflammation-related SNPs for glycated haemoglobin in MR Analysis

SNP	Inflammation-Related Pleiotropy	Effect Allele
rs10774625[a]	Neutrophil Count, Eosinophil Count, Monocyte Count, Lymphocyte Count, Basophil Count	A
rs11964178[a]	Granulocyte% White Cells, Basophil Count, Neutrophil Count, White Cell Count, Lymphocyte Count	A
rs1547247 [a]	White Cell Count, Monocyte Count, Neutrophil Count, Granulocyte% White Cells	A
rs17509001[a]	Lymphocyte Count, Monocyte Count, Neutrophil Count	C
rs4737009[a]	Lymphocyte Count, Neutrophil Count	A
rs579459[a]	IL-6, Neutrophil Count, Basophil Count, CRP	C
rs6474359[a]	Lymphocyte Count	C
rs7616006	White Cell Count, Monocyte Count, Lymphocyte Count, Neutrophil Count	A
rs1800562	Monocyte Count	A
rs2246434	Lymphocyte Count, Neutrophil Count, Lymphocyte% White Cells, Monocyte Count	A

[a]Genome-Wide Significant Inflammation-Related SNP; CRP=C-reactive protein; IL-=interleukin

Appendix C Table 18: Inflammation-related SNPs for type 2 diabetes in MR Analysis

SNP	Inflammation-Related Pleiotropy	Effect Allele
rs1260326[a]	CRP, Basophil Count, Neutrophil Count, Lymphocyte Count	C
rs2276853[a]	Lymphocyte Count, Neutrophil Count, Granulocyte Count	A
rs2073721[a]	Neutrophil Count, Monocyte %, Granulocyte %	G
rs1169288[a]	CRP	C
rs1060105[a]	Neutrophil Count, Lymphocyte Count	C
rs9891146[a]	Neutrophil Count, Eosinophil Count, Granulocyte Count	T
rs1800961[a]	CRP, Neutrophil Count, Granulocyte Count, Basophil Count, Myeloid Count	T
rs665268	Monocyte Count	G
rs1801282	Eosinophil% White Cells, Eosinophil Count	C
rs60980157	Basophil Count, Neutrophil Count; Myeloid White Cell Count, White Cell Count	C

[a]Genome-Wide Significant Inflammation-Related SNPs; CRP=C-reactive protein

Appendix C Table 19: Inflammation-related SNPs for body mass index in MR Analysis

SNP	Inflammation-Related Pleiotropy	Effect Allele
rs16851483[a]	Monocyte count	G
rs891389[a]	Neutrophil %, Eosinophils, Basophils	C
rs1558902[a]	CRP	A
rs571312[a]	CRP	A
rs1000940	Lymphocyte Count, White Cell Count	G
rs11663558	Neutrophil Count, Granulocyte Count, Basophil Count, Myeloid White Cell Count	A
rs12448257	Basophil Count, Neutrophil Count, Granulocyte Count, Myeloid White Cell Count	G
rs13107325	Eosinophil Count, Monocyte Count	C
rs205262	Lymphocyte Count	A
rs3817334	Granulocyte% Myeloid White Cells	C
rs4889606	Lymphocyte Count	G
rs6567160	Neutrophil Count, Myeloid Count, Granulocyte Count, White Cell Count	C

[a]Genome-Wide Significant Inflammation-Related SNPs; CRP=C-reactive protein

Appendix C Table 20: Inflammation-related SNPs for schizophrenia in MR Analysis

SNP	Inflammation-Related Pleiotropy	Effect Allele
rs2851447	Lymphocyte%, Neutrophil%	G
rs3130820	Lymphocyte Count, Basophil Count, Monocyte count	T
rs4925114	Lymphocyte Count	T
rs12416331	Monocyte Count	A
rs7216638	White cell count, Basophil count	T

Appendix C Table 21: SNPs used for CRP in MVMR Analysis

rs1205*
rs3093077
rs1130864*
rs1800947

*SNPs pruned during clumping procedure

Appendix C Table 22: Cochran Q Tests for Heterogeneity and MR Egger Intercept Tests for Horizontal Pleiotropy for the Association between all Cardiometabolic SNPs and Schizophrenia

Cardiometabolic Risk Factor	IVW		MR Egger			
	Cochran's Q (df)	*p*-value	Cochran's Q (df)	*p*-value	MR Egger Intercept (SE)	Intercept *p*-value
Fasting Insulin	19.37 (8)	0.013	11.03 (7)	0.137	-0.05 (0.02)	0.055
Triglycerides	46.66 (9)	<0.001	40.88 (8)	<0.001	-0.01 (0.01)	0.319
HDL	65.05 (14)	<0.001	46.80 (13)	<0.001	0.02 (0.01)	0.032
Fasting Plasma Glucose	50.11 (21)	<0.001	49.90 (20)	<0.001	<0.01 (0.01)	0.773
Type 2 Diabetes	119.81 (26)	<0.001	118.68 (25)	<0.001	-0.01 (0.01	0.646
Body Mass Index	320.28 (81)	<0.001	304.74 (80)	<0.001	-0.01 (0.01)	0.047
HbA1C	66.04 (34)	0.008	63.05 (33)	0.001	-0.01 (0.00)	0.219
Glucose Tolerance	14.47 (6)	0.024	7.47 (5)	0.188	-0.06 (0.03)	0.083
Leptin	10.45 (3)	0.015	6.75 (2)	0.034	-0.12 (0.11)	0.405
LDL	141.44 (75)	<0.001	141.39 (74)	<0.001	<0.01 (0.00)	0.873

IVW=inverse variance weighted regression; df=degrees of freedom; SE=standard error; HDL=high-density lipoprotein; HbA1C=glycated haemoglobin; LDL=low-density lipoprotein.

Appendix C Table 23: Cochran Q Tests for Heterogeneity and MR Egger Intercept Tests for Horizontal Pleiotropy for the Association between Inflammation-Related Cardiometabolic SNPs and Schizophrenia

Cardiometabolic Risk Factor	IVW		MR Egger			
	Cochran's Q (df)	*p*-value	Cochran's Q (df)	*p*-value	Regression Intercept (SE)	Direction *p*-value
Fasting Insulin	*	*	*	*	*	*
HDL	*	*	*	*	*	*
Type 2 Diabetes	34.89 (6)	<0.001	32.71 (5)	<0.001	-0.02 (0.03)	0.589
Fasting Plasma Glucose	7.09 (1)	0.008	*	*	*	*
HbA1C	34.89 (6)	<0.001	32.71 (5)	<0.001	-0.01 (0.01)	0.628
Body Mass Index	1.51 (4)	0.471	1.00 (3)	0.752	0.02 (0.02)	0.446
LDL	37.29 (11)	<0.001	32.65 (10)	0.001	0.01 (0.01)	0.261

IVW=inverse variance weighted regression; df=degrees of freedom; SE=standard error; HDL=high-density lipoprotein; HbA1C=glycated haemoglobin; LDL=low-density lipoprotein.
*insufficient *n* SNPs

Appendix C Table 24: Cochran Q Tests for Heterogeneity and MR Egger Intercept Tests for Horizontal Pleiotropy for the Association between Schizophrenia SNPs and Cardiometabolic Outcomes

Cardiometabolic Outcome	IVW		MR Egger			
	Cochran's Q (df)	*p*-value	Cochran's Q (df)	*p*-value	Regression Intercept (SE)	Direction *p*-value
Fasting Insulin	129.93 (100)	0.024	128.96 (99)	0.023	0.00 (0.00)	0.442
Triglycerides	205.86 (100)	<0.001	205.37 (99)	<0.001	0.00 (0.00)	0.628
HDL	373.38 (100)	<0.001	353.54 (99)	<0.001	0.01 (0.00)	0.020
Fasting Plasma Glucose	125.03 (104)	0.078	124.98 (103)	0.069	0.00 (0.00)	0.843
Type 2 Diabetes	139.83 (108)	0.021	139.50 (107)	0.019	0.00 (0.01)	0.612
Body Mass Index	264.27 (100)	<0.001	264.80 (99)	<0.001	-0.01 (0.00)	0.041
HbA1C	131.35 (103)	0.031	131.34 (102)	0.027	0.00 (0.00)	0.966
Glucose Tolerance	110.78 (100)	0.217	110.54 (99)	0.201	-0.01 (0.01)	0.642
Leptin	6.76 (2)	0.034	0.388 (1)	0.533	-0.16 (0.06)	0.240
LDL	183.06 (100)	<0.001	175.38 (99)	<0.001	0.01 (0.00)	0.040

IVW=inverse variance weighted regression; df=degrees of freedom; SE=standard error; HDL=high-density lipoprotein; HbA1C=glycated haemoglobin; LDL=low-density lipoprotein.

Appendix C Table 25: Cochran Q Tests for Heterogeneity and MR Egger Intercept Tests for Horizontal Pleiotropy for the Association between Inflammation-Related Schizophrenia SNPs and Cardiometabolic Outcomes

Cardiometabolic Outcome	IVW		MR Egger			
	Cochran's Q (df)	*p*-value	Cochran's Q (df)	*p*-value	Regression Intercept (SE)	Direction *p*-value
Fasting Insulin	6.83 (5)	0.233	6.50 (4)	0.165	0.00 (0.00)	0.676
Triglycerides	73.73 (5)	<0.001	56.79 (4)	<0.001	-0.02 (0.02)	0.336
HDL	4.00 (1)	0.050	*	*	*	*
Type 2 Diabetes	13.30 (5)	0.021	11.10 (4)	0.026	0.02 (0.03)	0.421
Fasting Plasma Glucose	6.91 (3)	0.075	2.97 (2)	0.226	0.01 (0.01)	0.245
Body Mass Index	36.18 (5)	<0.001	25.96 (4)	<0.001	-0.01 (0.01)	0.278
LDL	4.36 (2)	0.113	0.16 (1)	0.687	0.01 (0.00)	0.289
HbA1C	6.24 (3)	0.100	4.21 (2)	0.121	0.00 (0.00)	0.430
Glucose Tolerance	0.80 (2)	0.671	0.787 (1)	0.375	0.00 (0.03)	0.941
Leptin	0.42 (2)	0.812	0.351 (1)	0.554	0.00 (0.00)	0.841

IVW=inverse variance weighted regression; df=degrees of freedom; SE=standard error; HDL=high-density lipoprotein; HbA1C=glycated haemoglobin; LDL=low-density lipoprotein.
*insufficient *n* SNPs

Appendix C Table 26: MR-PRESSO Tests of Cardiometabolic All-SNP Analysis to Examine For and Correct Horizontal Pleiotropy

Risk Factor	MR-PRESSO Global Test		Outlier-Corrected IVW		Distortion Test	
	RSS	*p*-value	β (SE)	*p*-value	Coefficient	*p*-value
Fasting Insulin	24.35	0.018	0.08 (0.18)	0.669	171.67	0.156
Triglycerides	71.43	<0.001	0.23 (0.06)	0.008	-64.56	0.531
HDL	85.02	<0.001	-0.12 (0.08)	0.169	92.64	0.666
Fasting Plasma Glucose	53.78	<0.001	0.03 (0.05)	0.594	120.88	0.300
Type 2 Diabetes	148.58	<0.001	-0.06 (0.06)	0.390	-54.32	0.353
Body Mass Index	328.03	<0.001	0.02 (0.07)	0.815	187.40	0.255
HbA1C	69.33	0.002	0.06 (0.12)	0.651	-93.64	0.828
Glucose Tolerance	20.77	0.020	*	*	*	*
LDL	148.15	<0.001	-0.01 (0.03)	0.581	46.44	0.840
Leptin	32.73	0.002	0.27 (0.25)	0.382	22.42	0.338

MR PRESSO= Mendelian Randomization Pleiotropy Residual Sum and Outlier; β=beta coefficient; S.E=standard error. IVW=inverse variance weighted regression; df=degrees of freedom; RSS=residual sum of squares; SE=standard error; HDL=high-density lipoprotein; HbA1C=glycated haemoglobin; LDL=low-density lipoprotein.
*no evidence of horizontal pleiotropy

Appendix C Table 27: MR-PRESSO Tests of Inflammation-Related Cardiometabolic SNPs to Examine For and Correct Horizontal Pleiotropy

Risk Factor	MR-PRESSO Global Test		Outlier-Corrected IVW		Distortion Test	
	RSS	*p*-value	β (SE)	*p*-value	Coefficient	*p*-value
Fasting Insulin	†	†	†	†	†	†
HDL	†	†	†	†	†	†
Fasting Plasma Glucose	†	†	†	†	†	†
Type 2 Diabetes	47.76	0.001	0.11 (0.13)	0.436	-162.05	0.210
Body Mass Index	12.93	0.124	†	†	†	†
HbA1C	10.24	0.409	*	*	*	*
LDL	50.16	0.001	0.00 (0.07)	0.968	-261.84	0.230

MR PRESSO= Mendelian Randomization Pleiotropy Residual Sum and Outlier; β=beta coefficient; S.E=standard error. IVW=inverse variance weighted regression; df=degrees of freedom; RSS=residual sum of squares; SE=standard error; HDL=high-density lipoprotein; HbA1C=glycated haemoglobin; LDL=low-density lipoprotein.
*no evidence of horizontal pleiotropy; †no identified outliers

Appendix C Table 28: MR-PRESSO Tests of Schizophrenia All-SNP Analysis to Examine For and Correct Horizontal Pleiotropy

Outcome	MR-PRESSO Global Test		Outlier-Corrected IVW		Distortion Test	
	RSS	*p*-value	β (SE)	*p*-value	Coefficient	*p*-value
Fasting Insulin	161.53	0.020	†	†	†	†
Triglycerides	249.82	<0.001	0.00 (0.02)	0.210	590.84	0.064
HDL	434.93	<0.001	-0.01 (-0.02)	0.567	117.13	0.251
Fasting Plasma Glucose	155.12	0.067	*	*	*	*
Type 2 Diabetes	174.18	0.012	†	†	†	†
Body Mass Index	372.10	<0.001	-0.04 (0.02)	0.014	1.89	0.966
HbA1C	149.23	0.107	*	*	*	*
Glucose Tolerance	137.63	0.235	*	*	*	*
LDL	216.08	<0.001	0.00 (0.02)	0.866	-501.07	0.100
Leptin	113.87	0.772	*	*	*	*

MR PRESSO= Mendelian Randomization Pleiotropy Residual Sum and Outlier; β=beta coefficient; S.E=standard error. IVW=inverse variance weighted regression; df=degrees of freedom; RSS=residual sum of squares; SE=standard error; HDL=high-density lipoprotein; HbA1C=glycated haemoglobin; LDL=low-density lipoprotein.
*no evidence of horizontal pleiotropy; †no identified outliers

Appendix C Table 29: MR-PRESSO Tests of Inflammation-Related Schizophrenia SNP Analysis to Examine For and Correct Horizontal Pleiotropy

Outcome	MR-PRESSO Global Test		Outlier-Corrected IVW		Distortion Test	
	RSS	*p*-value	β (SE)	*p*-value	Coefficient	*p*-value
Fasting Insulin	1.08	0.883	*	*	*	*
Triglycerides	23.51	0.058	*	*	*	*
HDL	9.56	0.276	*	*	*	*
Fasting Plasma Glucose	15.34	0.095	*	*	*	*
Type 2 Diabetes	18.41	0.048	0.22 (0.09)	0.144	-182.77	<0.001
Body Mass Index	15.11	0.128	*	*	*	*
HbA1C	10.81	0.182	*	*	*	*
Glucose Tolerance	2.54	0.729	*	*	*	*
LDL	14.13	0.165	*	*	*	*
Leptin	3.29	0.647	*	*	*	*

MR PRESSO= Mendelian Randomization Pleiotropy Residual Sum and Outlier; β=beta coefficient; S.E=standard error. IVW=inverse variance weighted regression; df=degrees of freedom; RSS=residual sum of squares; SE=standard error; HDL=high-density lipoprotein; HbA1C=glycated haemoglobin; LDL=low-density lipoprotein.
*no evidence of horizontal pleiotropy

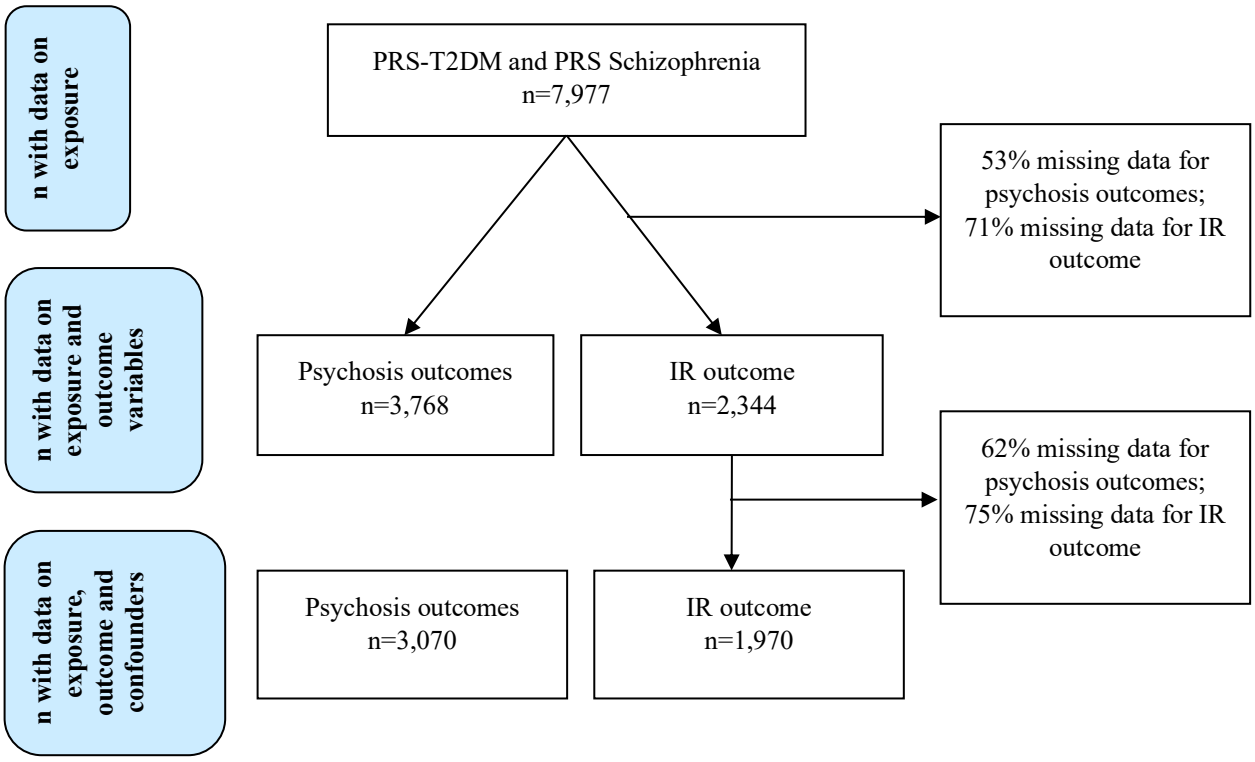
Appendix C Table 30: I_{2GX} Statistics to Examine for Potential Violation of the 'No Measurement Error' Assumption for MR Egger Analyses

Exposure	I^2_{GX} of SNP-Exposure Associations	
	All-SNP Analyses	Inflammation-Related SNP Analyses
Fasting Insulin	0.95	*
Triglycerides	0.99	*
HDL	0.99	*
Fasting Plasma Glucose	0.95	0.90
Type 2 Diabetes	0.84	0.62
Body Mass Index	0.93	0.99
HbA1C	0.94	0.91
Glucose Tolerance	0.51	*
Leptin	0.00	*
LDL	0.99	0.98
Schizophrenia	0.00	0.83

HDL=high-density lipoprotein; HbA1C=glycated haemoglobin; LDL=low-density lipoprotein.
*insufficient *n* SNPs

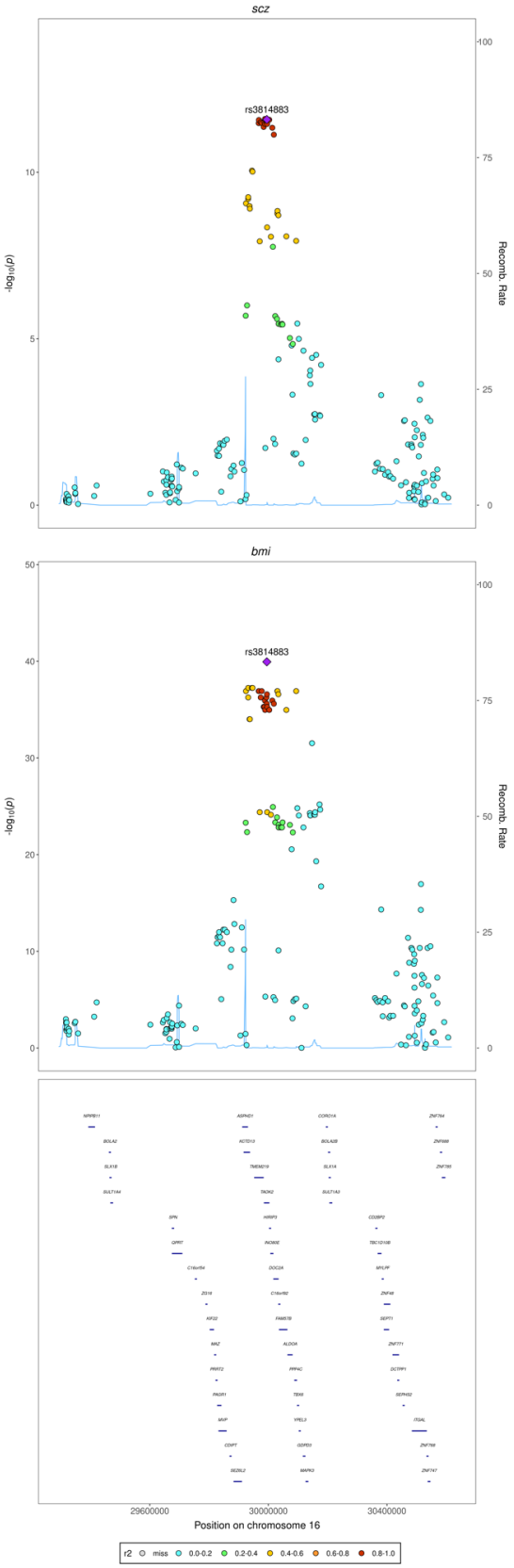
Appendix C Figures

Appendix C Figure 1: Flowchart of Included Participants in ALSPAC Analysis

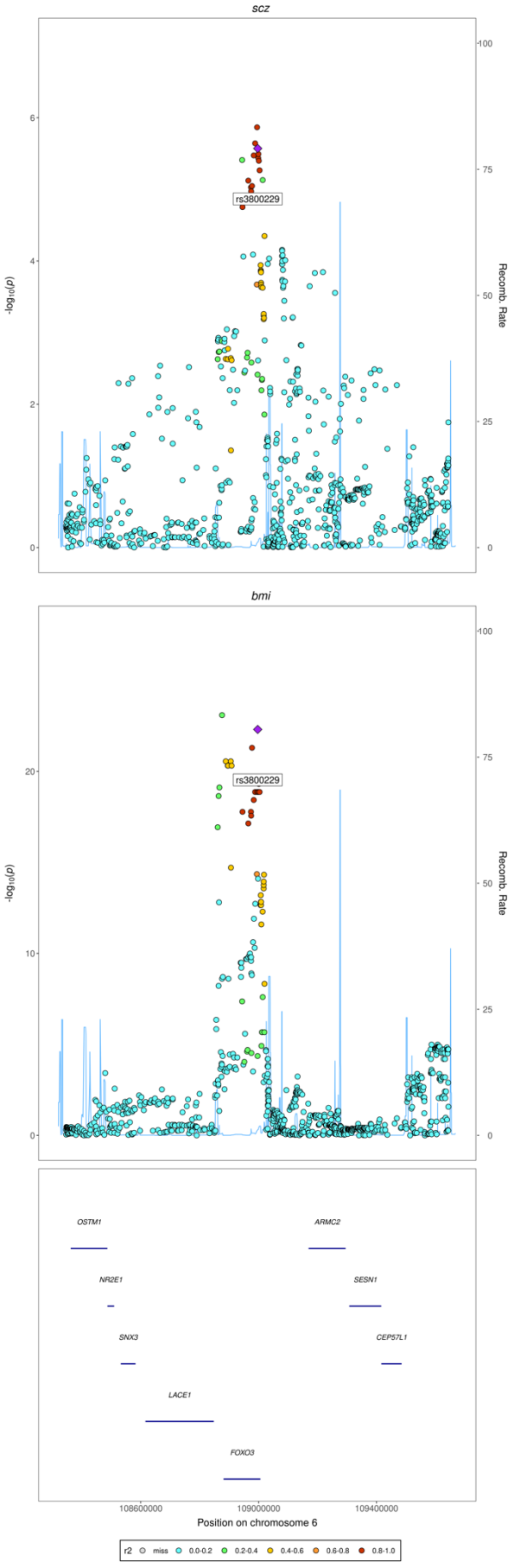


Appendix C Figure 2: Regional Genetic Association Plots for Additional Loci Indicating Evidence of Colocalization

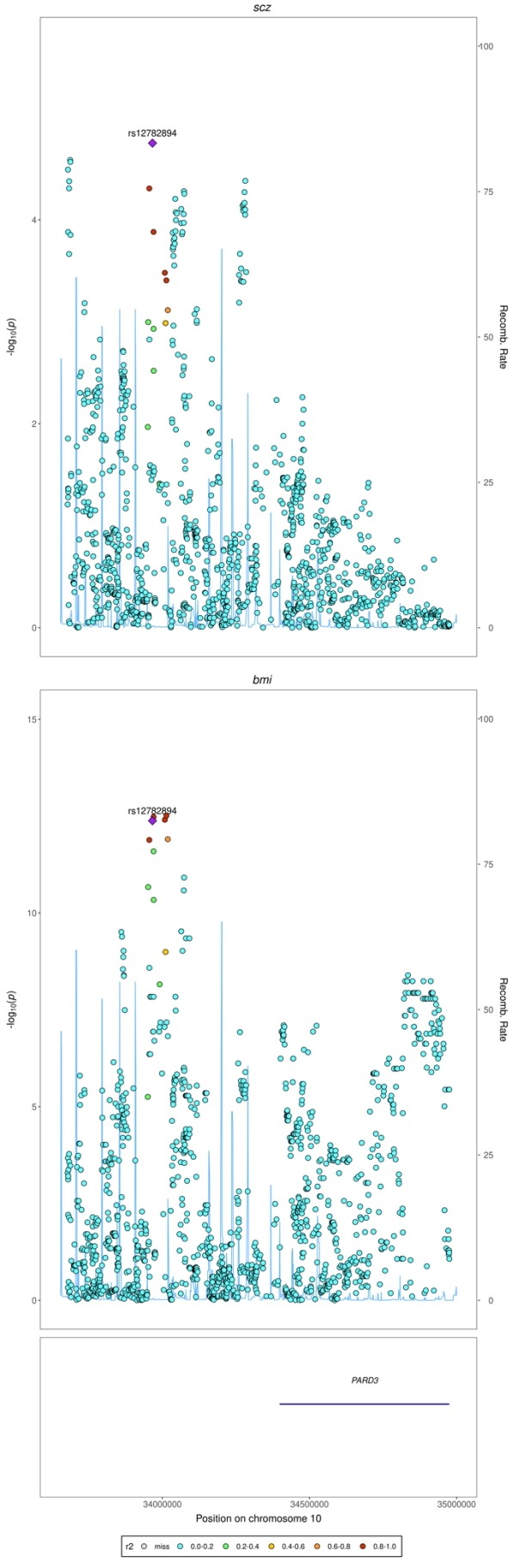
A. rs3814883 – *TAOK2*



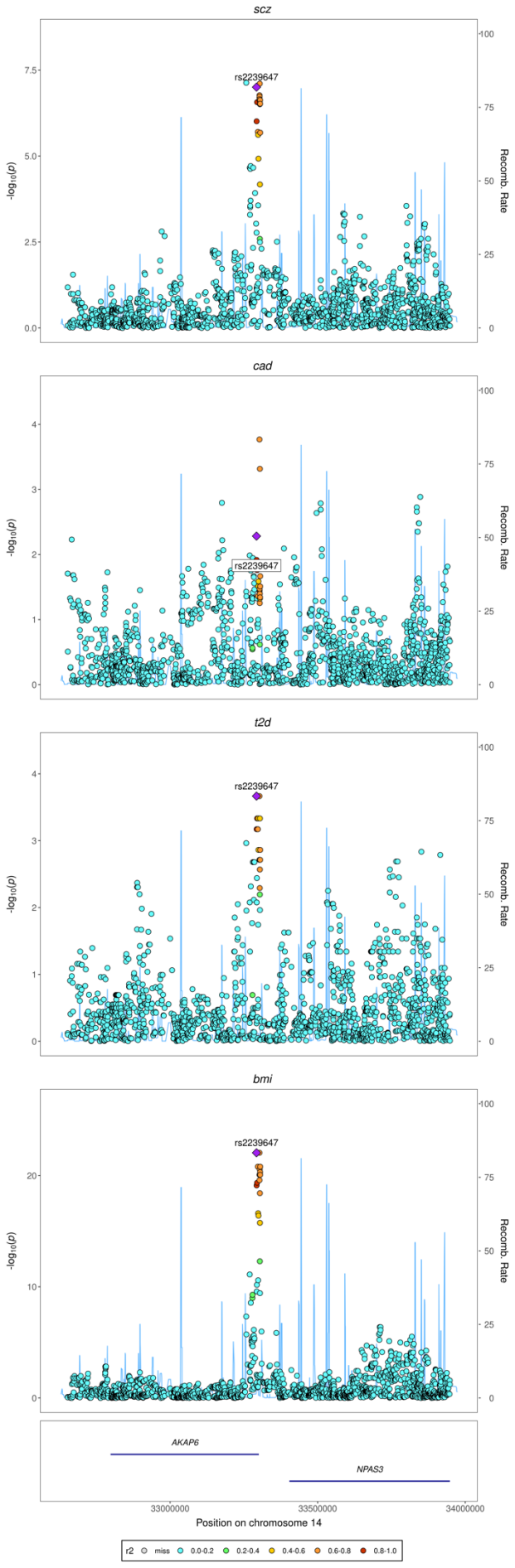
B. rs3800229 – *FOXO3*



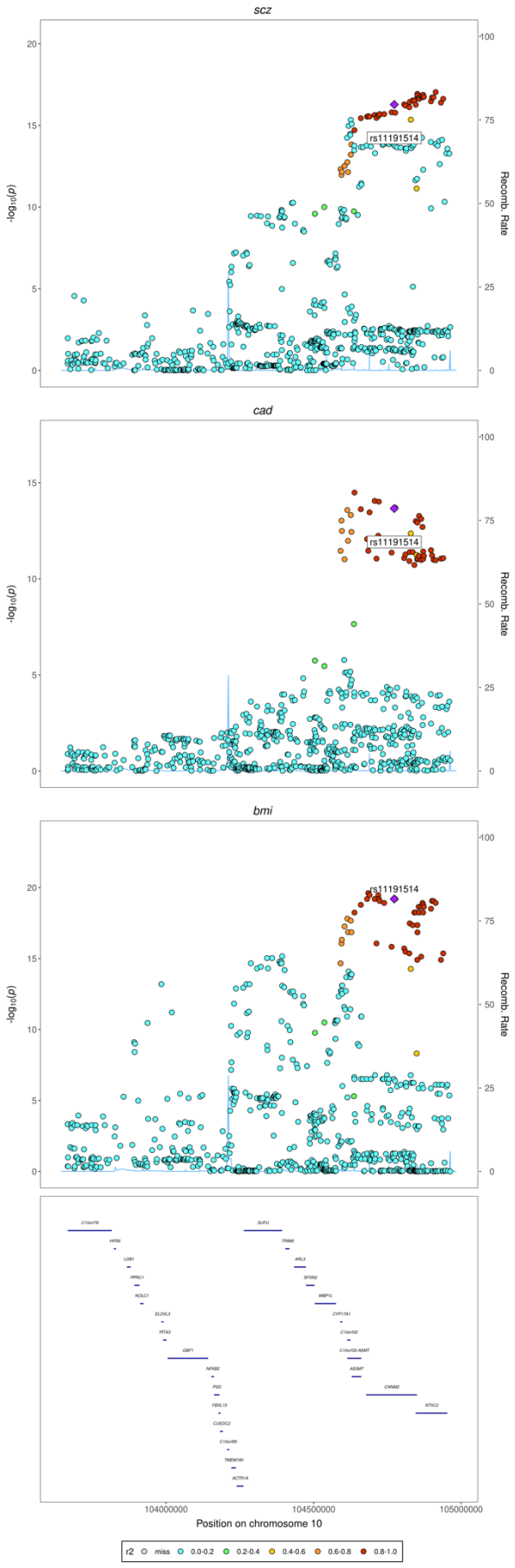
C. rs12782894



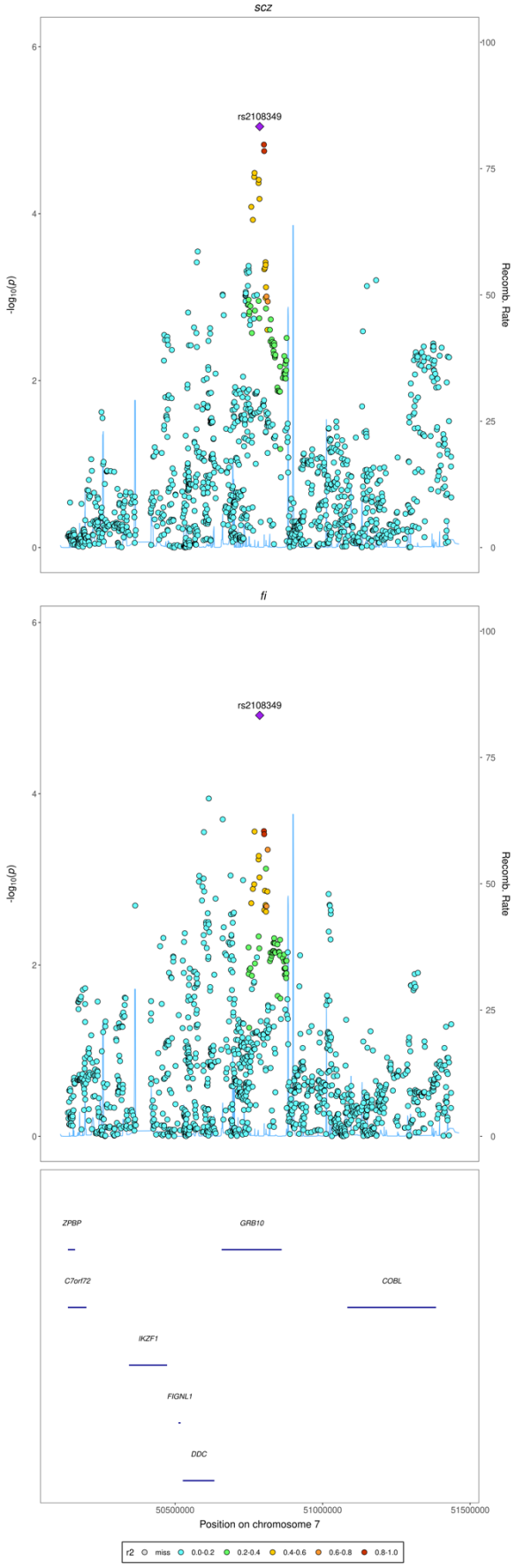
D. rs2239647 – *AKAP6*



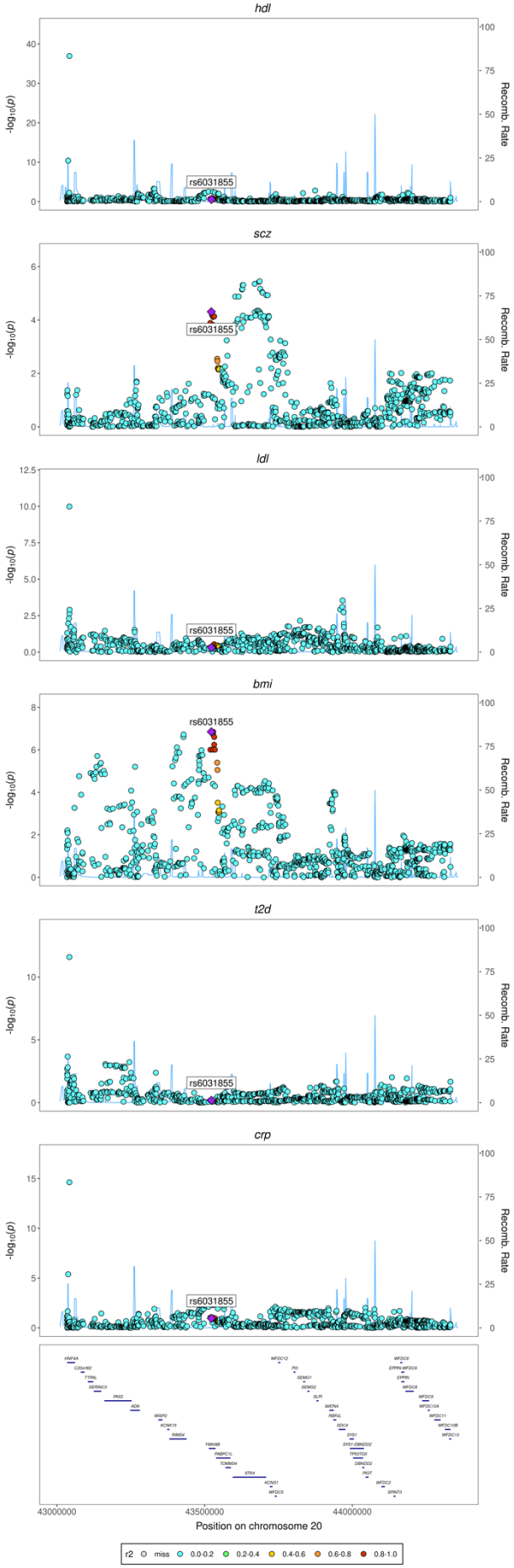
E. rs1191514 – *CNNM2*



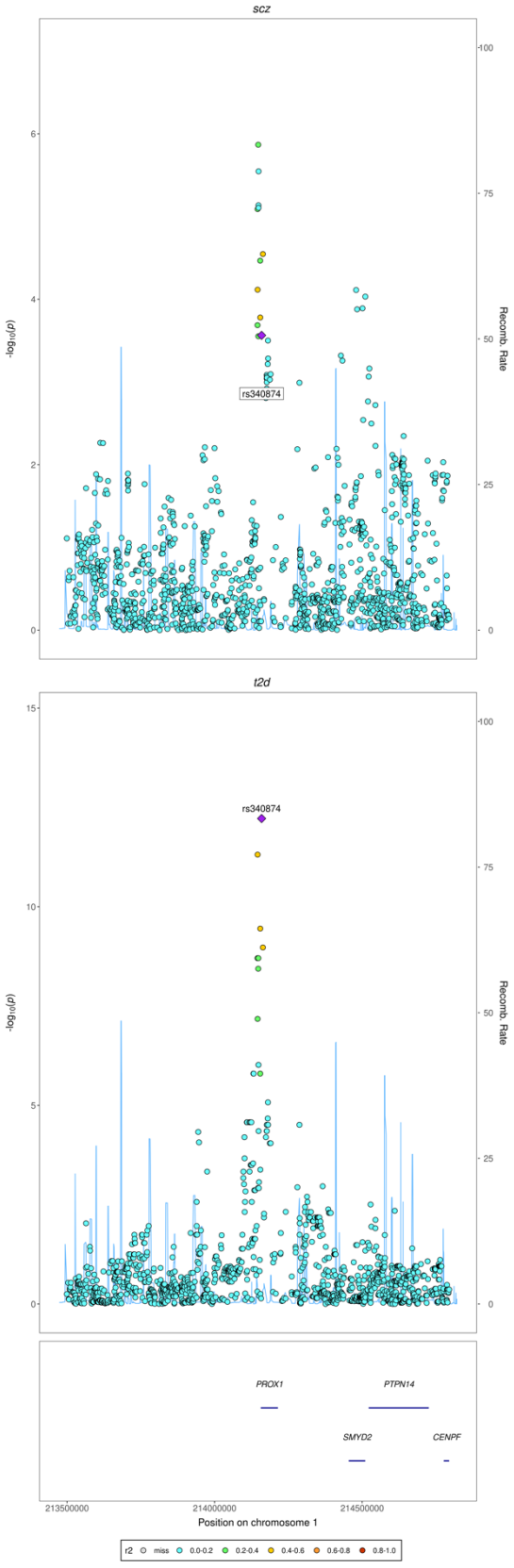
F. rs2108349 – *GRB10*



G. rs6031855 - *YWHAB*



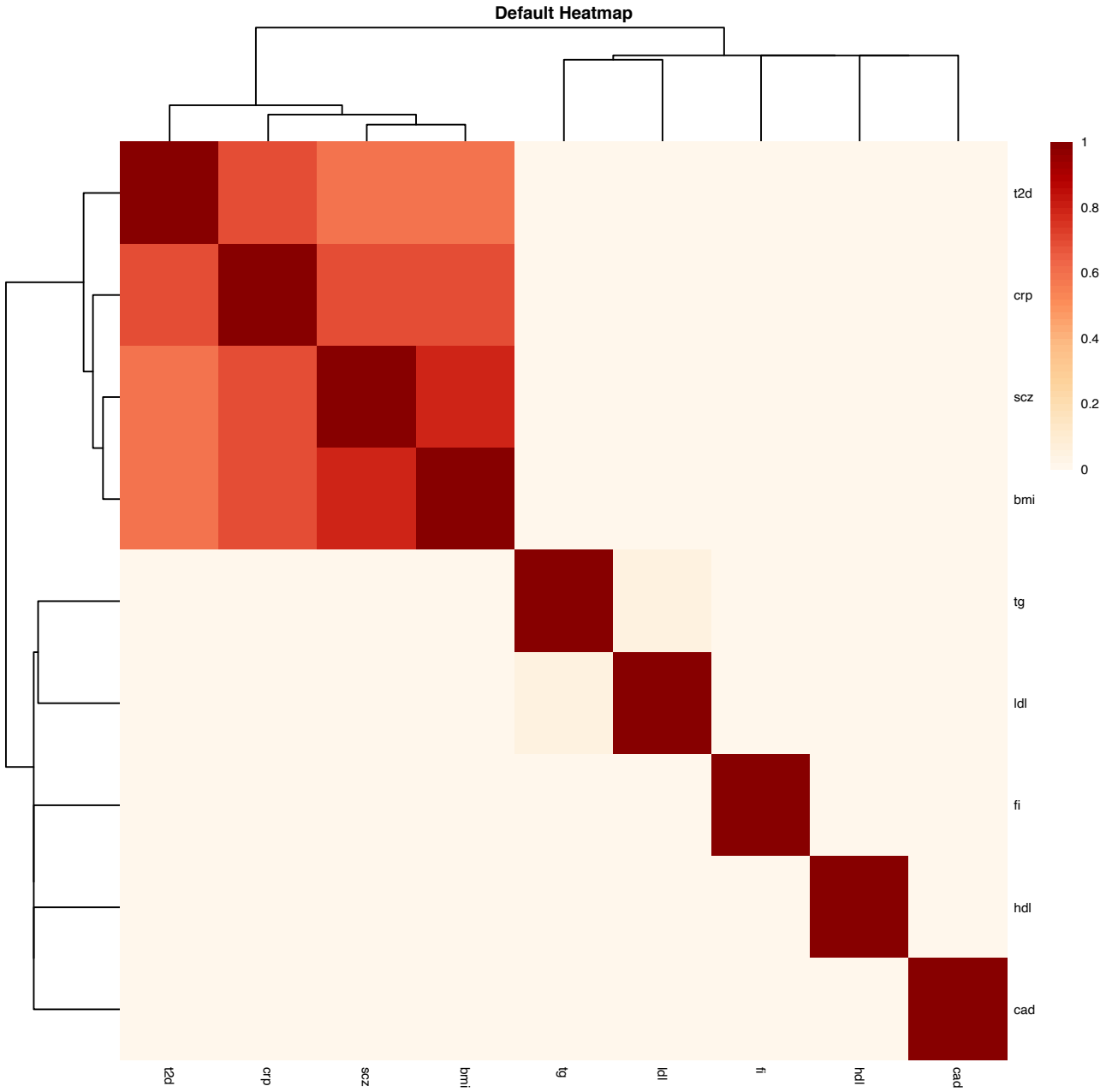
H. rs340874 – *PROX1*



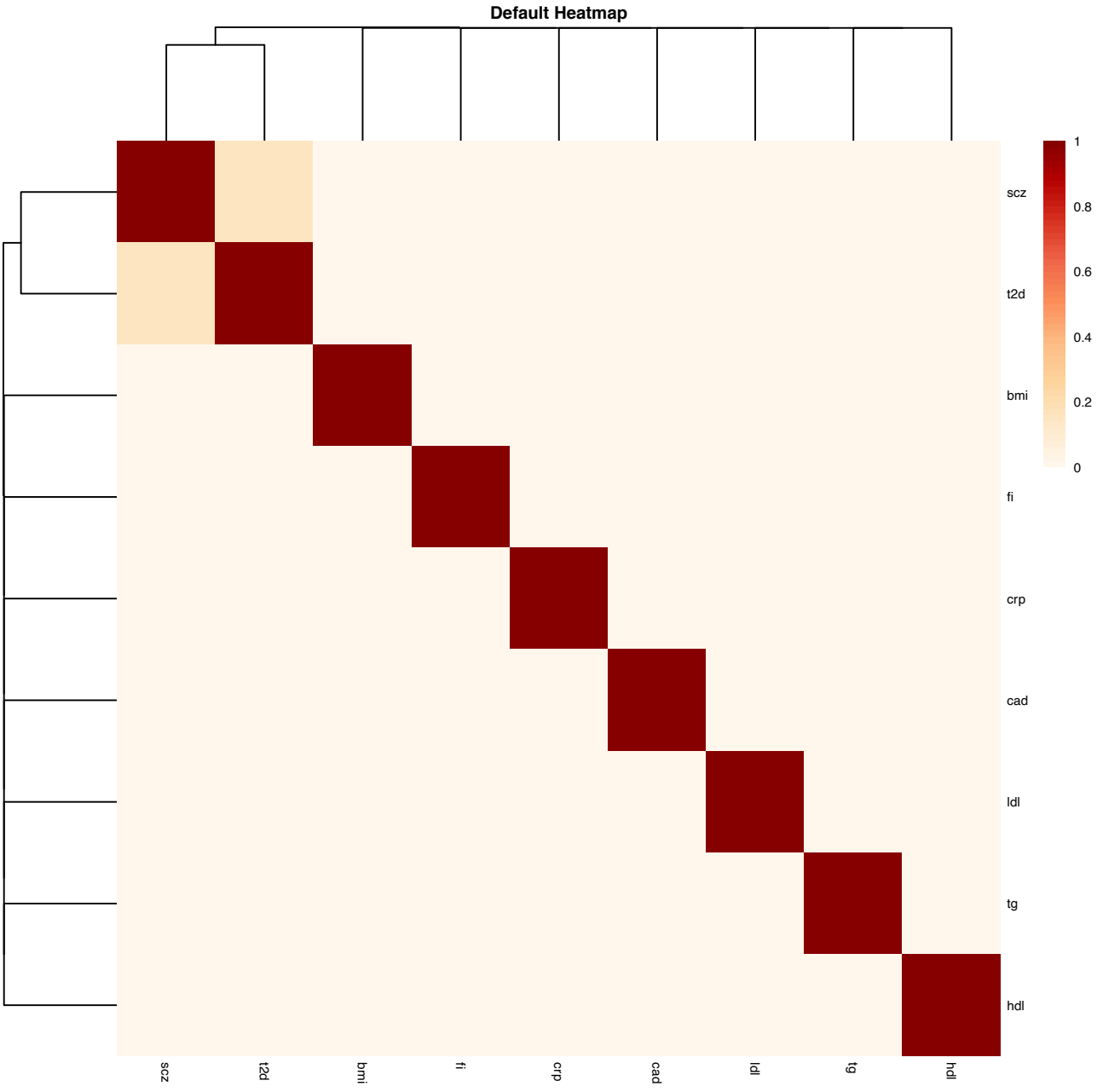
Regional association plots denoting chromosomal location (x axis) and strength of association with listed trait ($-\log_{10(p)}$) (y axis). SNP r^2 estimated from the EPIC-Norfolk cohort.

Appendix C Figure 3: Heatmap Sensitivity Plots for SNPs with Evidence of Colocalization Between Schizophrenia and Cardiometabolic and Inflammatory Traits

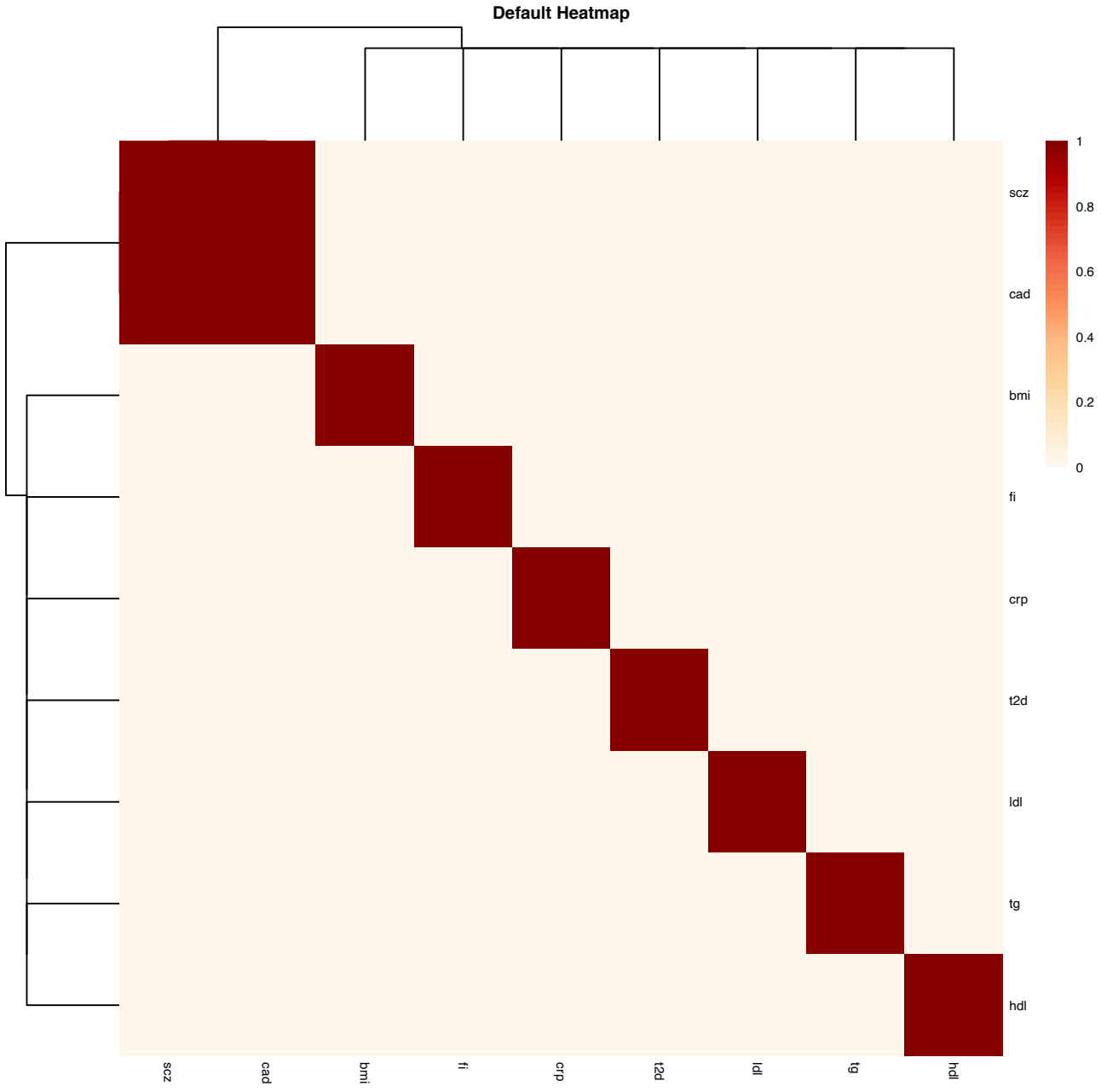
A. rs8192675 – *SLC2A2*



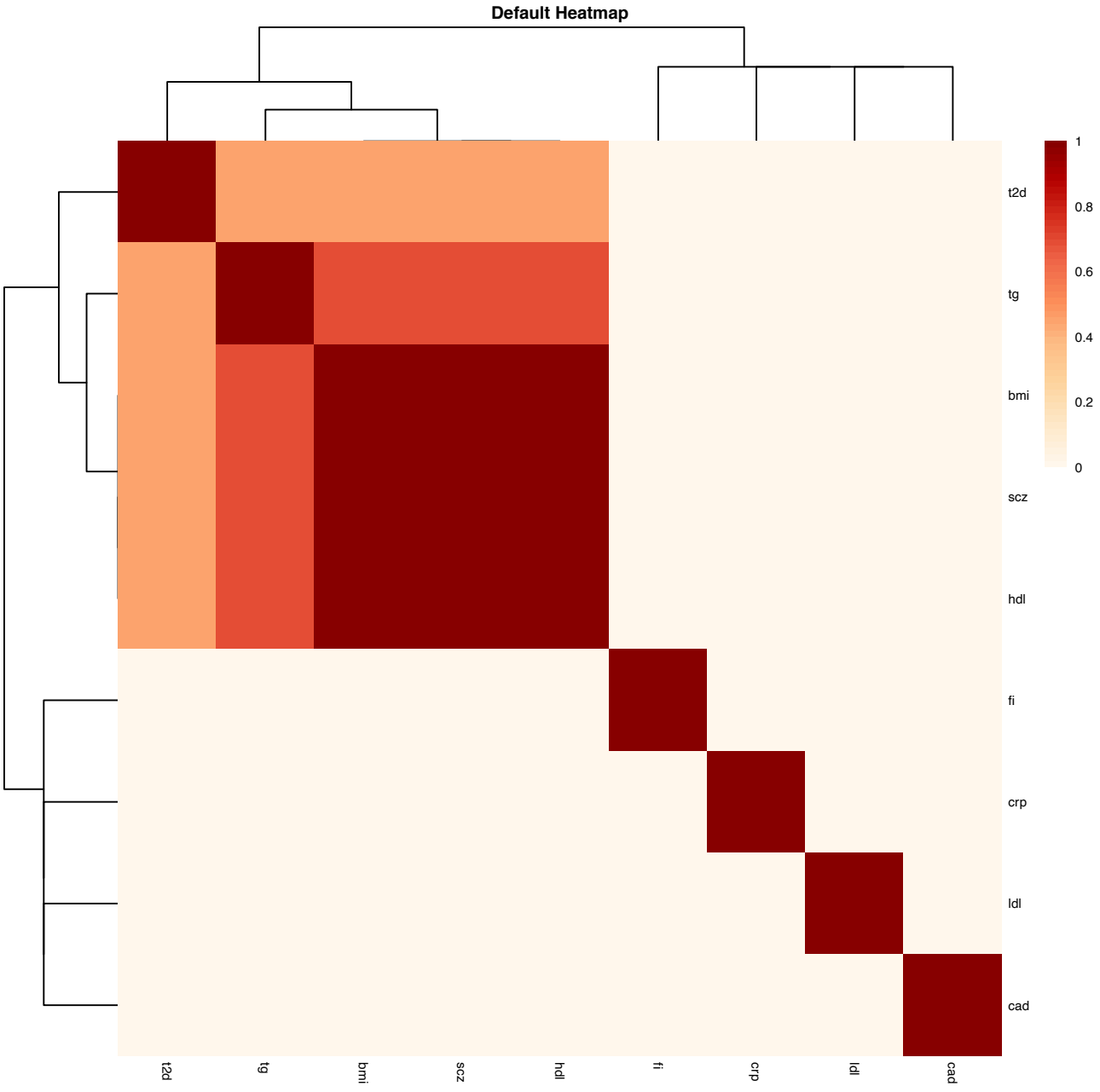
B. rs340874 – *PROX1*



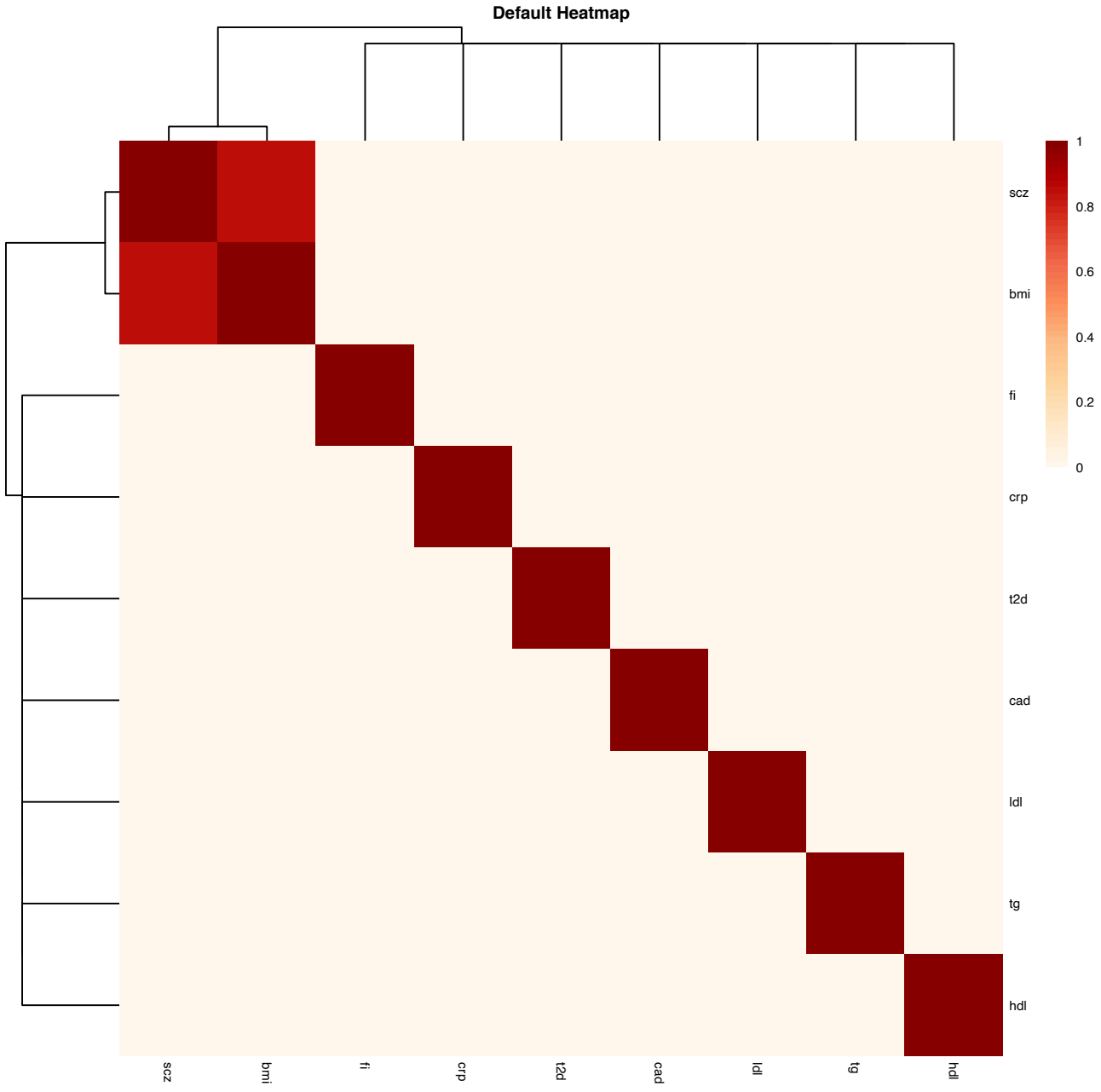
C. rs17514846 - *FURIN*



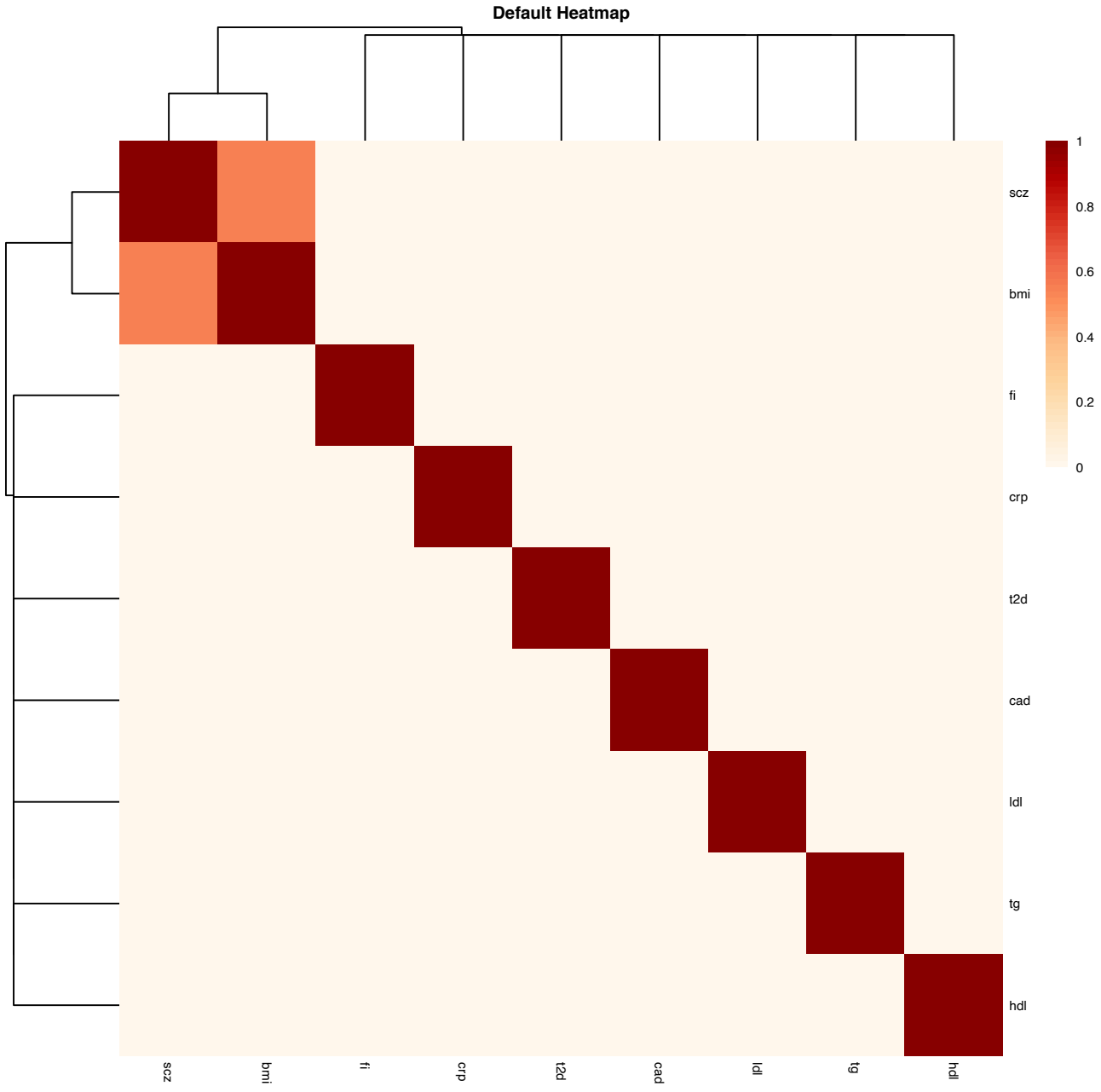
D. rs13107325 – *SLC39A8*



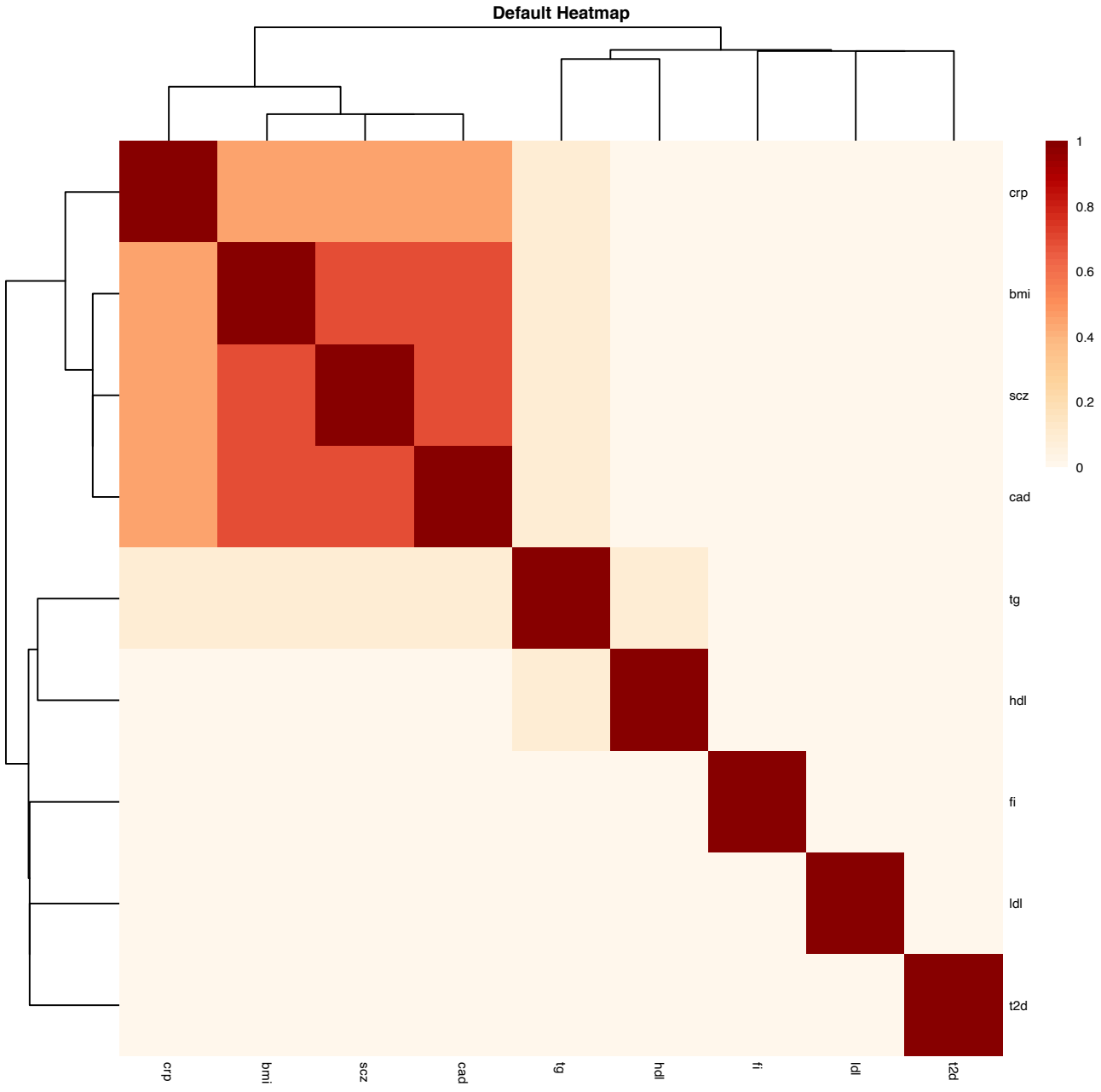
E. rs3814883 – *TAOK2*



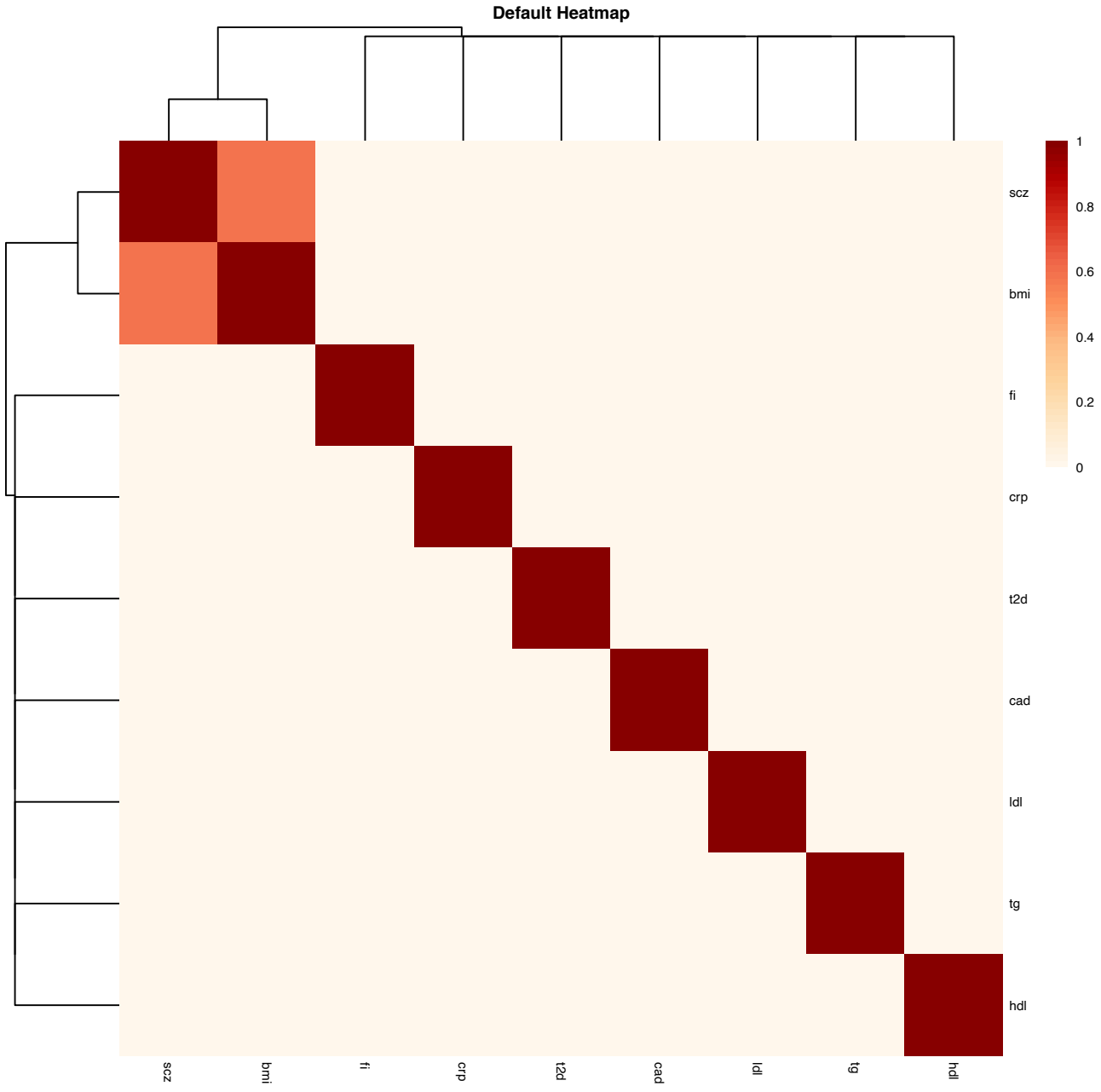
F. rs12782894



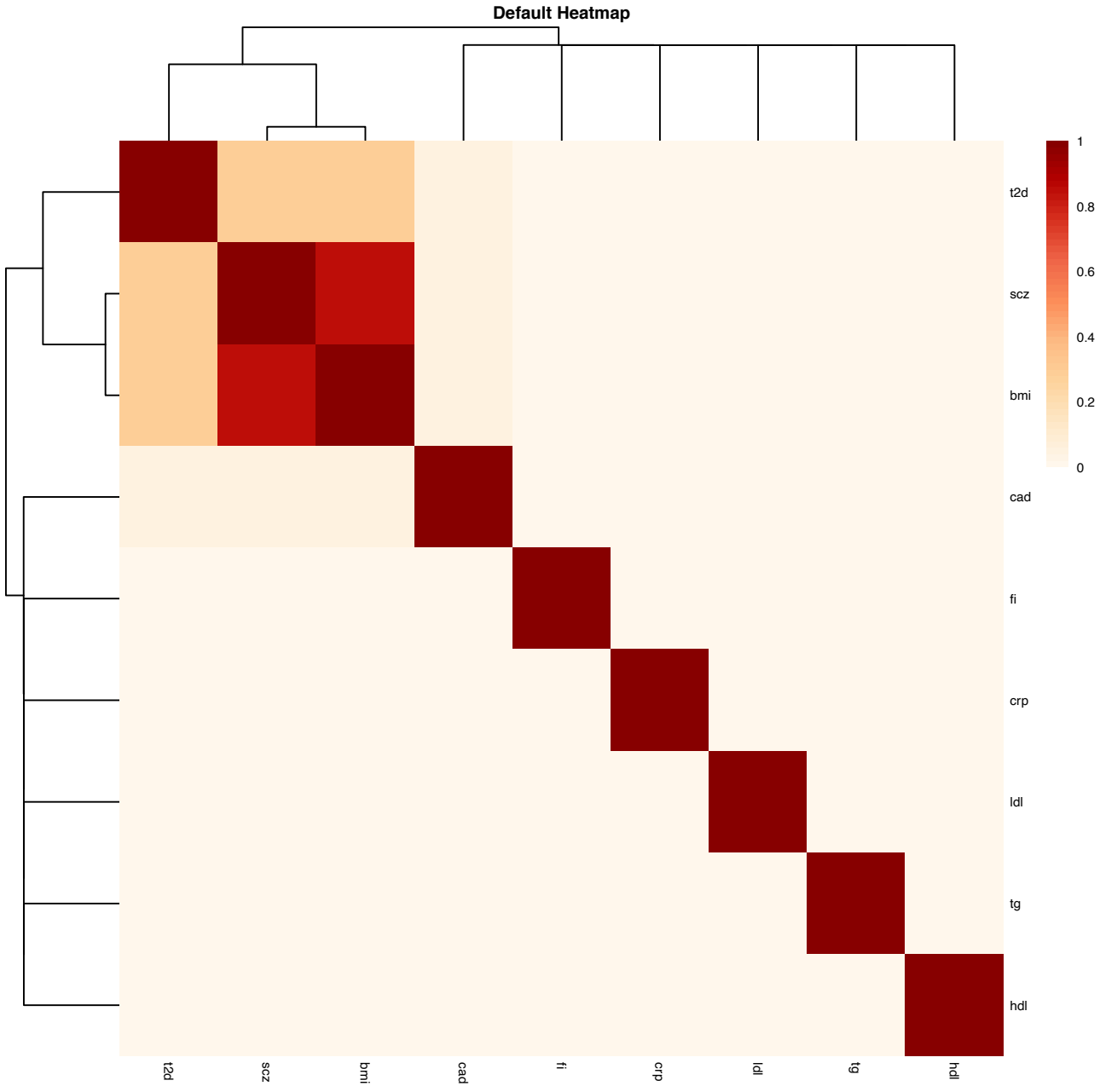
G. rs6265 - *BDNF*



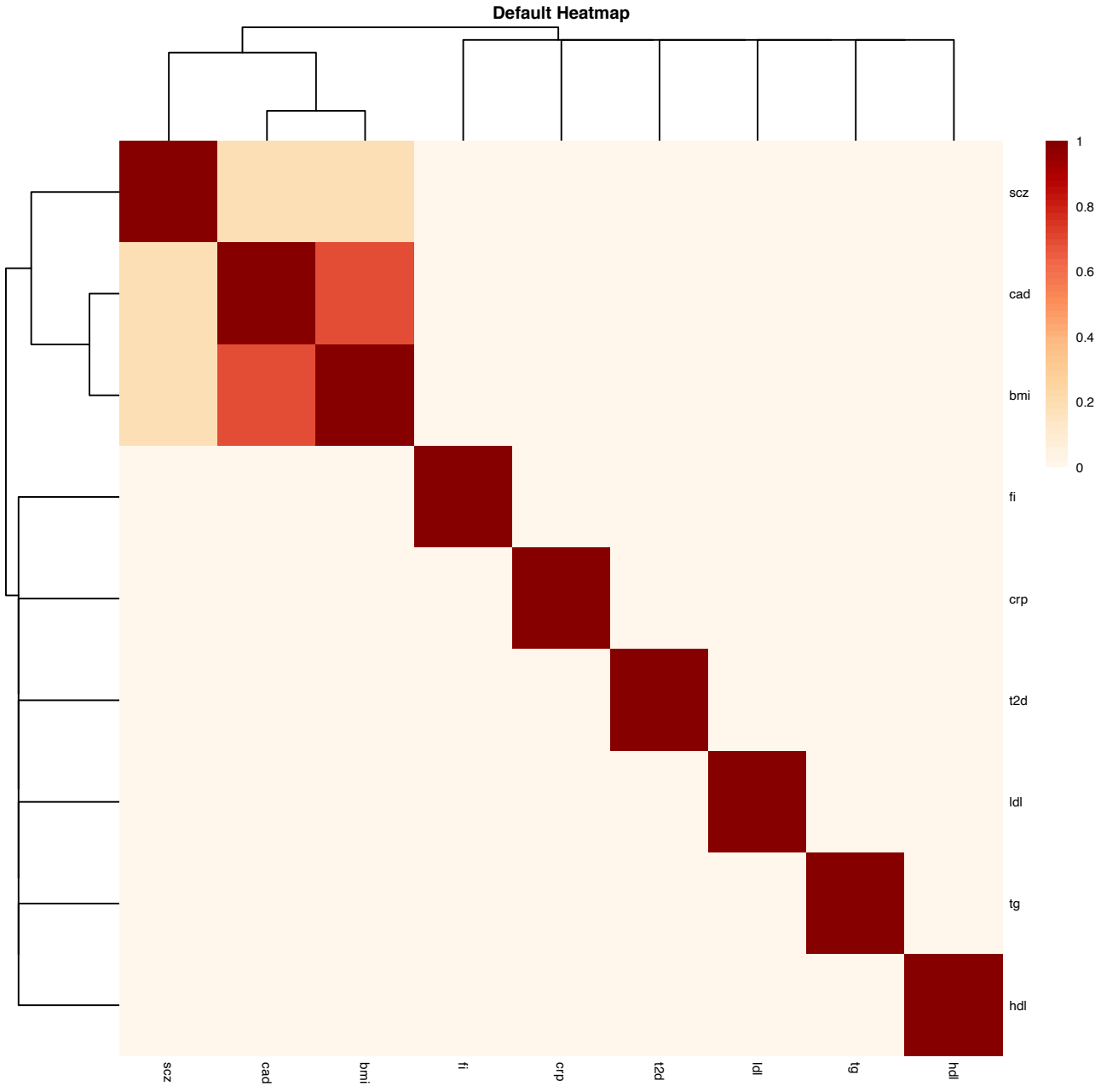
H. rs3800229 – *FOXO3*



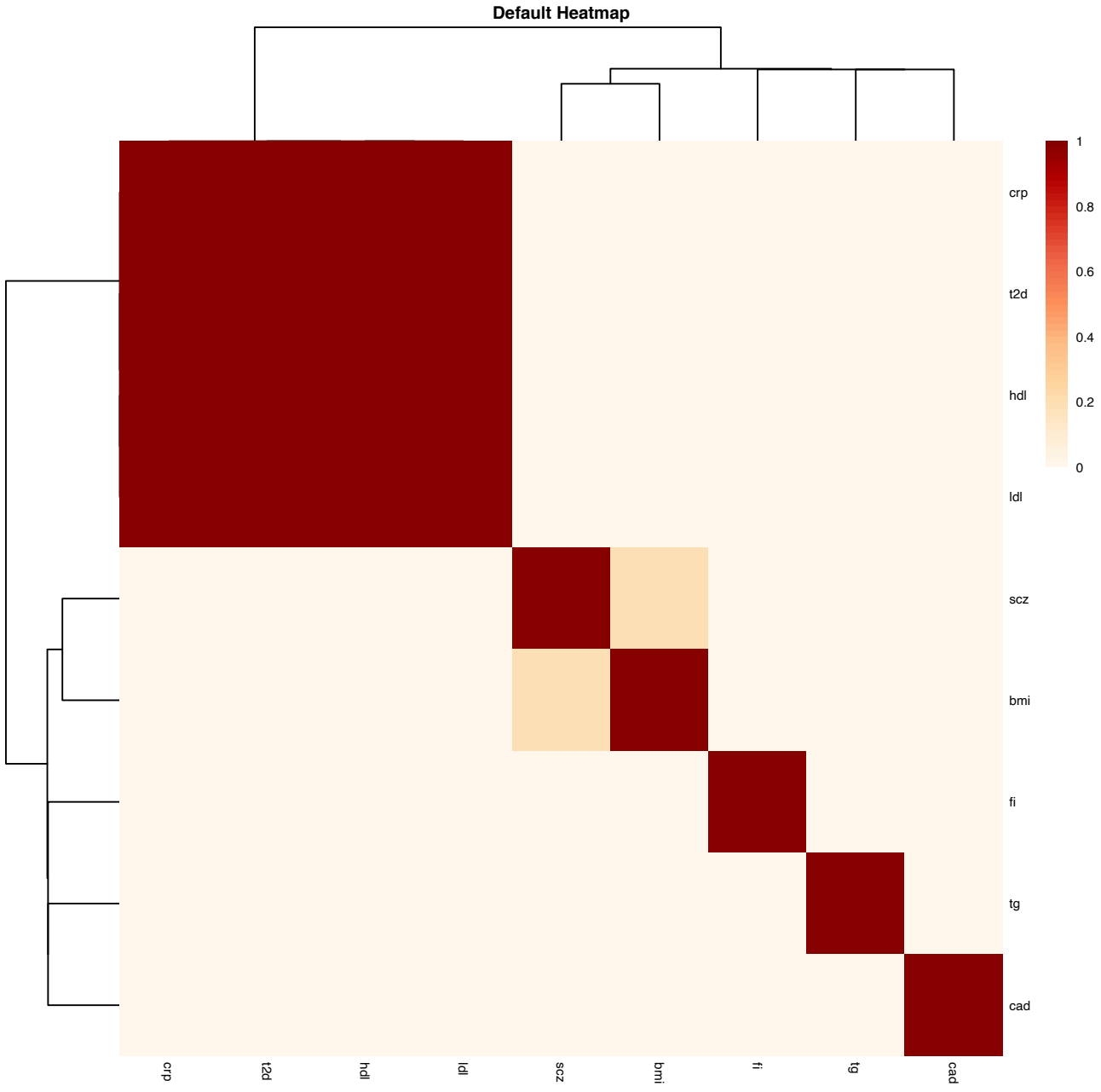
I. rs2239647 – *AKAP6*



J. rs11191514 – *CNNM2*



K. rs6031855 - *YWHAB*



Heatmaps drawn based on a similarity matrix across increasingly stringent prior and threshold permutations, from prior1 = $1x10^{-4}$; prior2 = 0.05; regional/alignment thresholds = 0.5, to prior2 = 0.001; regional/alignment thresholds = 0.9).
1 = evidence of colocalization across all permutations (dark red) and 0 = no evidence of colocalization at any permutation (beige). bmi=body mass index; hdl = high-density lipoprotein; tg=triglycerides; scz=schizophrenia; ldl=low-density lipoprotein; t2d=type 2 diabetes mellitus; crp=C-reactive protein; fi=fasting insulin; cad=coronary artery disease.

Appendix C Published Manuscripts

Schizophrenia Research 223 (2020) 227–235

Contents lists available at ScienceDirect

Schizophrenia Research

journal homepage: www.elsevier.com/locate/schres

Common mechanisms for type 2 diabetes and psychosis: Findings from a prospective birth cohort

Benjamin I. Perry [a,b,*], Hannah J. Jones [c,d], Tom G. Richardson [e], Stan Zammit [c,d,f], Nicholas J. Wareham [g], Glyn Lewis [h], Peter B. Jones [a,b], Golam M. Khandaker [a,b]

[a] Department of Psychiatry, University of Cambridge School of Clinical Medicine, Cambridge, England, UK
[b] Cambridgeshire and Peterborough NHS Foundation Trust, Cambridge, England, UK
[c] Centre for Academic Mental Health, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, England, UK
[d] NIHR Biomedical Research Centre, University Hospitals Bristol NHS Foundation Trust and University of Bristol, Bristol, UK
[e] MRC Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, England, UK
[f] MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, Wales, UK
[g] MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge, England, UK
[h] Division of Psychiatry, University College London, London, England, UK

ARTICLE INFO

Article history:
Received 5 November 2019
Received in revised form 15 April 2020
Accepted 11 August 2020
Available online 19 August 2020

Keywords:
Psychosis
Schizophrenia
Diabetes mellitus
Polygenic risk
ALSPAC

ABSTRACT

Background: Psychosis and type 2 diabetes mellitus (T2DM) are commonly comorbid and may share pathophysiologic mechanisms. To investigate shared genetic variation and inflammation as potential common mechanisms, we tested: (i) associations between genetic predisposition for T2DM and psychotic experiences and psychotic disorder in young adults; (ii) the association between genetic predisposition for schizophrenia and insulin resistance (IR), a precursor of T2DM; and (iii) whether these associations are mediated by childhood inflammation.
Methods: Psychotic experiences (PEs), psychotic disorder and IR were assessed at age 18. Polygenic risk scores (PRS) for T2DM and schizophrenia were derived based on large genome-wide association studies. Associations between PRS and psychotic/IR outcomes were assessed using regression analysis based on 3768 ALSPAC birth cohort participants with complete data. Inflammatory markers C-reactive protein (CRP) and interleukin 6 (IL-6) measured at age 9 were used in regression and mediation analyses.
Results: Genetic predisposition for T2DM was associated with PEs (adjusted OR = 1.21; 95% CI, 1.01–1.45) and psychotic disorder (adjusted OR = 1.51; 95% CI, 1.04–2.03) at age 18 in a linear dose-response fashion. Genetic predisposition for schizophrenia was weakly associated with IR (adjusted OR = 1.10; 95% C·I, 0.99–1.22) at age 18. The association between genetic risk for T2DM and PEs was partly mediated by childhood CRP ($p = .040$).
Conclusions: Comorbidity between psychosis and T2DM may be partly underpinned by shared genes and inflammation. A summation of minor genetic variation representing lifetime risk for T2DM at conception may predispose individuals to psychosis in adulthood by influencing physiologic changes, such as low-grade inflammation, detectable as early as childhood.

1. Introduction

Reduced life-expectancy in schizophrenia is largely attributable to physical comorbidity including cardiometabolic disorders, which are up to 30% more prevalent in people with schizophrenia than in the general population (Holt et al., 2004) (Lappin et al., 2018). Compared with controls, markers of abnormal glucose-insulin homeostasis are two to three times higher in young people with psychotic experiences (PEs) (Perry et al., 2018), and in medication-naive first-episode psychosis (FEP) (Perry et al., 2016; Pillinger et al., 2017) after controlling for anthropometric and sociodemographic factors. This suggests that increased T2DM in patients with psychosis may not be fully explained by common lifestyle factors or side-effects of antipsychotic drugs, though may be exacerbated by them (Rajkumar et al., 2017).

One contributor to comorbidity between cardiometabolic disorders and schizophrenia could be shared genetic susceptibility (Lin and Shuldiner, 2010). Risk of insulin resistance (IR) (Chouinard et al., 2019) and impaired glucose tolerance (Ferentinos and Dikeos, 2012), two key precursors of T2DM, are higher in unaffected relatives of patients with psychosis compared with controls. People with comorbid schizophrenia and T2DM have a higher genetic predisposition for both disorders compared to controls (Hackinger et al., 2018), and an association between genetic predisposition for schizophrenia and IR has been

* Corresponding author at: Inflammation and Psychiatry Research Group, Department of Psychiatry, University of Cambridge, Herchel Smith Building, Robinson Way, Cambridge, CB2 0SZ, UK.
E-mail address: bip20@medschl.cam.ac.uk (B.I. Perry).

https://doi.org/10.1016/j.schres.2020.08.006

reported in a clinical sample (Tomasik et al., 2019). Conversely, a relatively small study found no evidence of an association between genetic risk for T2DM and psychosis (Padmanabhan et al., 2016), and previous research using linkage-disequilibrium (LD) score regression found limited evidence for a genetic correlation between schizophrenia and T2DM (Bulik-Sullivan et al., 2015). However, a key feature of existing studies is that they are based on adult cases of established schizophrenia or T2DM or rely on blood measurements taken in adulthood, so confounding by cumulative effects of lifestyle and other factors is possible (Reinikainen et al., 2015). Population-based prospective studies have identified early markers of disease risk associated with T2DM and schizophrenia. For instance, PEs in adolescence or young adulthood are associated with risk of schizophrenia in adulthood (Poulton et al., 2000; Zammit et al., 2013), and IR is a precursor of T2DM (Martin et al., 1992). To our knowledge, no studies have examined whether genetic predispositions for T2DM or schizophrenia are associated with, respectively, PEs or IR, in young adulthood. Demonstrating such associations with early markers of illness in young adults with lessened effects of cumulative lifestyle confounding would be consistent with the idea that shared genetic variation is a common mechanism for comorbid T2DM and schizophrenia.

Although existing studies provide some evidence for a shared genetic basis for T2DM and schizophrenia, underlying pathophysiologic mechanisms remain unclear. Low-grade inflammation may be one such mechanism, which has been reported to be associated with IR (Festa et al., 2000), T2DM (Pradhan et al., 2001) and psychosis (Upthegrove et al., 2014). Population-based longitudinal studies report that higher levels of circulating inflammatory markers at baseline are associated with risks of psychosis and abnormal glucose-insulin homeostasis subsequently at follow-up (Khandaker et al., 2014; Perry et al., 2018). Mendelian randomisation (MR) studies have reported associations of genetic variants regulating inflammatory biomarkers such as interleukin-6 (IL-6) with schizophrenia (Hartwig et al., 2017), suggesting that inflammation may be associated with schizophrenia beyond any effects of confounding. Inflammation has also been implicated in the pathogenesis of IR and T2DM (Pradhan et al., 2001).

We examined whether shared genetic variation and inflammation could be common mechanisms for T2DM and psychosis using prospective, population-based data from the ALSPAC birth cohort. We tested whether: (i) genetic predisposition for T2DM is associated with risk of PEs and psychotic disorder at age 18; (ii) genetic predisposition for schizophrenia is associated with IR at age 18; (iii) whether these associations are mediated by CRP or IL-6 levels measured in childhood at age 9.

2. Methods

2.1. Description of cohort and sample selection

The ALSPAC birth cohort (Boyd et al., 2013; Fraser et al., 2013) comprises 14,062 live births from mothers residing in former County Avon in Southwest England, with expected dates of delivery between April 1991 and December 1992 (http://www.bristol.ac.uk/alspac/researchers/our-data/). The study received ethics approval from the ALSPAC Ethics and Law Committee and local research ethics committees. All participants provided written or implied informed consent. In total, 7977 participants had genotyping data, 3768 participants had data on both genotyping and psychosis outcomes, and 2344 participants had data on genotyping and IR (Supplementary Fig. 1). Our analysis was conducted on participants without missing data for the covariates or outcomes of interest.

2.2. Assessment of psychotic outcomes at age 18

2.2.1. Psychotic experiences (PEs)

PEs were identified through the face-to-face, semi-structured Psychosis-Like Symptom Interview (PLIKSi) conducted by trained psychology graduates. The PLIKSi comprised of an introductory set of questions on unusual experiences, and then 12 'core' questions eliciting key symptoms covering the three main domains of positive psychotic symptoms: hallucinations (visual and auditory); delusions (delusions of being spied on, persecution, thoughts being read, reference, control, grandiose ability and other unspecified delusions); and symptoms of thought interference (thought broadcasting, insertion and withdrawal). For these 12 core items, 7 stem questions were derived from the Diagnostic Interview Schedule for Children–IV (DISC–IV) and 5 stems from section 17–19 of the Schedules for Clinical Assessment in Neuropsychiatry version 2.0 (SCAN 2.0). After cross-questioning, interviewers rated PEs as not present, suspected, or definitely present. Interviewers rated down (i.e. suspected rather than definite, or none rather than suspected) if unsure. For suspected or definite PEs, interviewers also recorded the frequency; effects on social/educational/ occupational function; help seeking; and attributions including fever, hypnopompic/hypnogogic state, or illicit drugs. For interrater reliability, the interviewers recorded audio interviews at three time points, approximately 6 months apart, across the clinic duration (75 interviews in total). The average kappa value of PEs was 0.83, with no evidence of differences across time. Test-retest reliability was assessed using 162 individuals reinterviewed after approximately 47 days (kappa = 0.76, SE = 0.078), 46 of whom were reinterviewed by the same interviewer (kappa = 0.86, SE = 0.136). Our primary outcome was presence of *definite* PEs, referring to at least one definite PE since age 12; the comparator group was suspected/no PEs. Our outcome is reflective of 6-year period prevalence of definite PEs. From the total number of participants with definite PEs at 18y (230, 4.9%), 80 participants (45.3%) had suffered definite PEs at least once in the month preceding assessment. From the total sample of participants reporting definite PEs, 146 participants (63.5%) reported auditory hallucinations, 63 participants (28.2%) reported any delusion, and 22 participants (9.9%) reported thought disturbance. See Supplementary Table 1 for full frequency data, Supplementary Table 2 for information on timing of onset of PEs, and the main reporting study for further information (Zammit et al., 2013).

2.2.2. Psychotic disorder

Psychotic disorder was defined (Zammit et al., 2013) as the presence of PEs when symptoms were not attributable to fever/sleep/drugs, had occurred at least once per month over the previous 6 months, and caused significant distress resulting in either help-seeking from a professional source (general practitioner, counsellor, mental health team), or significantly disrupted social/occupational function. From the total ALSPAC sample who underwent the PLIKSi, 46 participants (1.0%) met criteria for psychotic disorder. We included psychotic disorder as a secondary outcome due to its lower prevalence in the study sample.

2.3. Assessment for a T2DM-risk outcome at age 18

2.3.1. Insulin resistance

IR was calculated as a binary variable based on fasting plasma glucose and insulin levels at age 18, using the well-validated homeostasis model assessment (HOMA) method (Matthews et al., 1985). There is no consensus-agreed cut-off for clinical IR in the literature since levels can vary between populations (Wallace et al., 2004). Therefore, we used the 75th centiles of the study population to define IR. The 75th centile cut-off has been used in previous research (Cediel et al., 2016; Geloneze et al., 2006; Hedblad et al., 2000; Marques-Vidal et al., 2002). The 75th centile in our study population was 2.15.

2.4. Assessment for polygenic risk scores for T2DM and schizophrenia

From the ALSPAC cohort, 8812 participants were genotyped using the Illumina HumanHap550 quad genome-wide SNP genotyping platform by *23andMe* subcontracted to the Wellcome Trust Sanger Institute, Cambridge, UK and the Laboratory Corporation of America, Burlington,

NC, USA. Following quality control assessment and imputation, and restricting to 1 young person per family, genetic data was available for 7977 ALSPAC individuals. See Supplementary Methods for further information.

Polygenic risk scores (PRS) for schizophrenia and T2DM were constructed for all 7977 participants with genotype data, using training sets based on the second Psychiatric Genomics Consortium (PGC) Schizophrenia GWAS (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014) and a large T2DM GWAS (Mahajan et al., 2014), respectively. Both GWAS analyses adjusted for principal components to reduce the impact of population stratification (Price et al., 2006). PRS were calculated using the PLINK (v1.9) (Chang et al., 2015; Purcell et al., 2007) 'score' command following the methodology described by the International Schizophrenia Consortium (ISC) (Purcell et al., 2009). Prior to construction of scores, single nucleotide polymorphisms (SNPs) were removed from the analysis if they had a minor allele frequency less than 0.01, an imputation quality less than 0.8 or if there was allelic mismatch between samples (see Supplementary methods for details). Due to the presence of strand differences between ALSPAC and the T2DM GWAS, and lack of allele frequency information in the T2DM summary statistics, palindromic SNPs were also removed prior to construction of the T2DM PRS. Because of the high linkage disequilibrium (LD) within the extended major histocompatibility complex (MHC; chromosome 6: 25-34 Mb) only a single SNP was included to represent this region. SNPs were pruned for LD using the PLINK 'clump' command to remove SNPs in LD ($r^2 > 0.25$) with a more significant SNP in the training set. Windows of 500 kb were used to assess inter-SNP LD for pruning.

For the primary analysis, PRS were constructed using a list of SNPs with the optimal *p*-value thresholds to capture phenotypic variance defined by both GWAS individually ($p \leq 10^{-5}$ for T2DM (Mahajan et al., 2014) and $p \leq .05$ for schizophrenia (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014)). Scores were weighted by the logarithm of the odds ratio (OR) for schizophrenia or T2DM reported by the GWAS training sets, for the schizophrenia and T2DM PRS, respectively. 10 Principal components (PCs) were generated using unrelated individuals (IBS < 0.05) and independent SNPs (with long range LD regions removed) using the `– pca` command in PLINK1.90. All PRS analyses were adjusted for the 10 PCs to reduce the risk of population stratification. Two PRS measures were calculated for T2DM; the first including all SNPs associated with T2DM, and the second after excluding a SNP located in the *FTO* gene region, which is widely understood to be associated with T2DM only through its influence on body mass index (BMI) variation (Frayling et al., 2007); the latter was used in sensitivity analysis. Additionally, since the optimal *p*-value thresholds of both PRS scores differed, we conducted sensitivity analyses to examine PRS-outcome associations using a range of *p*-value thresholds from $p = .5$ to genome-wide significance ($p < 5 \times 10^{-8}$).

2.5. Assessment of inflammatory markers at age 9

Data on two inflammatory markers at age 9 years (IL-6 and CRP) were available in ALSPAC, for 5076 and 5086 participants respectively. Blood samples were collected at non-fasting state. Please see supplementary methods for further information.

2.6. Assessment of potential confounders

We included sex (categorical), ethnicity (binary caucasian/non-caucasian due to the predominantly caucasian sample), social class (categorical) and BMI at age 18 years (continuous). We excluded participants with hsCRP levels >10 mg/L to minimize potential bias from recent/ongoing infection or chronic inflammatory disease.

2.7. Statistical analysis

We examined the distribution of PRS-T2DM and PRS-schizophrenia using the Shapiro-Wilk test for normality, and from visual inspection of Q-Q plots. The distributions were $p > .05$ and appeared normally distributed. Both PRS variables were standardized (*Z*-transformed).

2.7.1. Association between PRS and outcomes at age 18

We conducted logistic regression analyses to examine the association between PRS-T2DM and risks for PEs and psychotic disorder, and PRS-schizophrenia and IR at age 18. The odds ratios (OR) and 95% confidence intervals (95% C.I.) indicate increase in risk per standard deviation (SD) increase in PRS. Regression models were adjusted for sex, ethnicity, social class, and BMI. *p*-values for adjusted regression models in our primary analysis were corrected for multiple testing per the three outcomes we included (definite PEs, psychotic disorder and IR) using the Holm-Bonferroni method (Holland and Copenhaver, 1987). We used the p.adjust() command in R (R Core Team, 2017) to perform adjustments. In results tables, we present the original unadjusted *p*-values alongside Holm-Bonferroni adjusted *p*-values. To test for linearity of associations, we included a quadratic term (PRS^2) in the logistic regression models.

2.7.2. Association between PRS scores and childhood inflammatory markers at age 9 years

We used linear regression analyses to test associations of PRS for T2DM or schizophrenia, separately, with IL-6 and CRP levels at age 9 years (Z-transformed values), before and after adjustments for potential confounders listed above.

2.7.3. Mediation by childhood CRP

We performed mediation analyses to examine whether any evident associations may be mediated by childhood CRP levels. We calculated direct and indirect effects between exposure (PRS-T2DM or PRS-schizophrenia) and outcome (e.g., PEs or IR) taking into account the mediator variable (e.g., CRP). Evidence of an indirect effect is consistent with mediation. The indirect effect was bootstrapped using 5000 iterations to determine the 95% CIs. Mediation analysis was performed using the PROCESS macro V3.1 for IBM SPSS 24.0 (http://www.afhayes.com).

2.8. Missing data

We assessed the potential impact of missing data by comparing mean PRS score between the analytic sample and participants with missing data for psychosis and IR outcomes, using separate variance *t*-tests. We also performed logistic regression analysis to determine sociodemographic and other predictors (sex, ethnicity, BMI and social class) of missing data.

3. Results

3.1. Baseline characteristics of sample

Of the 3768 participants with data on PRS-T2DM and psychotic outcomes, 283 met the criteria for suspected/definite PEs (7.5%), 183 for definite PEs (5.1%), 29 (0.7%) for psychotic disorder at age 18 (Table 1). Of the 2344 participants with data on PRS-schizophrenia and IR, 173 met the criteria for IR at age 18 (7.3%).

3.2. Association between genetic predisposition for T2DM and psychotic outcomes at age 18

The prevalence of psychotic outcomes at age 18 years was higher for participants in the top third of PRS-T2DM distribution compared with those in the bottom third (Fig. 1). PRS-T2DM was associated with

Table 1
Baseline characteristics of sample.

Characteristic, *n* (%) unless otherwise stated	All sample	Definite PEs	Psychotic disorder	No/suspected PEs
Male sex	1846 (49)	71 (38)	7 (15)	1775 (49)
White British ethnicity	3692 (98)	179 (98)	39 (95)	3513 (98)
Social class				
I & II	1582 (42)	62 (35)	5 (16)	1456 (40)
III - non manual & manual	1616 (43)	75 (43)	15 (48)	1630 (44)
IV & V	565 (15)	38 (22)	11 (36)	583 (16)
BMI (kg/m^2) at 18 years, mean (SD)	22.71 (3.76)	23.37 (4.49)	22.73 (4.26)	22.60 (3.71)
HOMA at 18 years, mean (SD)	0.92 (0.73)	1.03 (0.75)	1.28 (1.00)	0.92 (0.73)
Insulin resistance	251 (8)	25 (17)	7 (20)	209 (7)
Current smoking	220 (7)	22 (15)	5 (18)	188 (7)
CRP (mg/L) at 9 years, mean (SD)	0.68 (2.52)	0.72 (2.61)	0.75 (1.33)	0.67 (2.49)
PEs attributed to sleep/fever/drugs[a]	N/A	31 (0.7)	7 (0.1)	27 (0.6)
Help-seeking from professional source[a]	N/A	55 (24)	41 (51.9)	6 (3)

Information based on total ALSPAC sample.
[a] Recorded from Zammit et al. (2013).

definite PEs (adjusted OR = 1.21; 95% CI, 1.01–1.45 per SD increase in PRS-T2DM) and psychotic disorder (adjusted OR = 1.51; 95% CI, 1.04–2.05 per SD increase in PRS-T2DM) at age 18 years after controlling for sex, ethnicity, social class and BMI (Table 2). Quadratic terms for PRS-T2DM in these regression models were non-significant suggesting no evidence for departure from linearity (all $p > .05$). The results for sensitivity analyses using PRS-T2DM score excluding a SNP in the *FTO* gene region were similar (Supplementary Table 3).

3.3. Association between genetic predisposition for schizophrenia and IR at age 18

There was weaker evidence for an association between PRS-schizophrenia and IR at age 18 (adjusted OR = 1.10; 95% CI, 0.99–1.22 per SD increase in PRS-schizophrenia) after controlling for sex, ethnicity, social class and BMI. The quadratic term for PRS-schizophrenia was non-significant suggesting no evidence for departure from linearity ($p > .05$).

3.4. Associations between PRS scores and inflammatory markers at age 9

Data on both PRS scores and serum IL-6 and CRP levels were available for 2180 and 2176 participants respectively. After adjustments for sex, ethnicity, social class and BMI, PRS-T2DM was associated with CRP (β = 0.03; 95% CI, 0.01–0.08, $p = .040$), but not with IL-6 (β = 0.01; 95% CI, −0.02–0.05, $p = .082$). There was also trend level evidence for an association between PRS-schizophrenia and CRP (β = 0.05; 95% CI, −0.01–0.10, $p = .061$) but not with IL-6 (β = 0.01; 95% CI, −0.04-0.09, $p = .670$).

3.5. Mediating effect of childhood CRP levels on the associations of PRS scores with psychotic outcomes or IR

Based on 1955 participants with data on PRS-T2DM, CRP levels at age 9 and PEs at age 18, CRP at age 9 partially mediated the association between PRS-T2DM and definite PEs at age 18. There was evidence of an indirect effect indicative of mediation; the coefficients were 0.28; 95% CI, 0.07–0.45, $p = .044$ for direct effect; co-efficient = 0.05; 95% CI 0.02–0.12, $p = .040$ for indirect effect. Since IL-6 levels at age 9 years were not associated with PRS-T2DM, we did not perform mediation analysis using IL-6. There was no evidence for a mediating effect of CRP on the association between PRS-schizophrenia and IR at age 18; the coefficients were 0.14; 95% CI, −0.06–0.34, $p = .756$ for direct effect; co-efficient = 0.01; 95% CI, −0.01–0.03, $p = .180$ for indirect effect.

3.6. Results for sensitivity analysis using different P-value thresholds for PRS

Fig. 2 presents the associations between PRS-T2DM and PEs alongside the associations between PRS-schizophrenia and IR, at different PRS *p*-value thresholds. The point estimates for the PRS-T2DM-PEs associations were >1 for all *p*-value thresholds, though the strength of

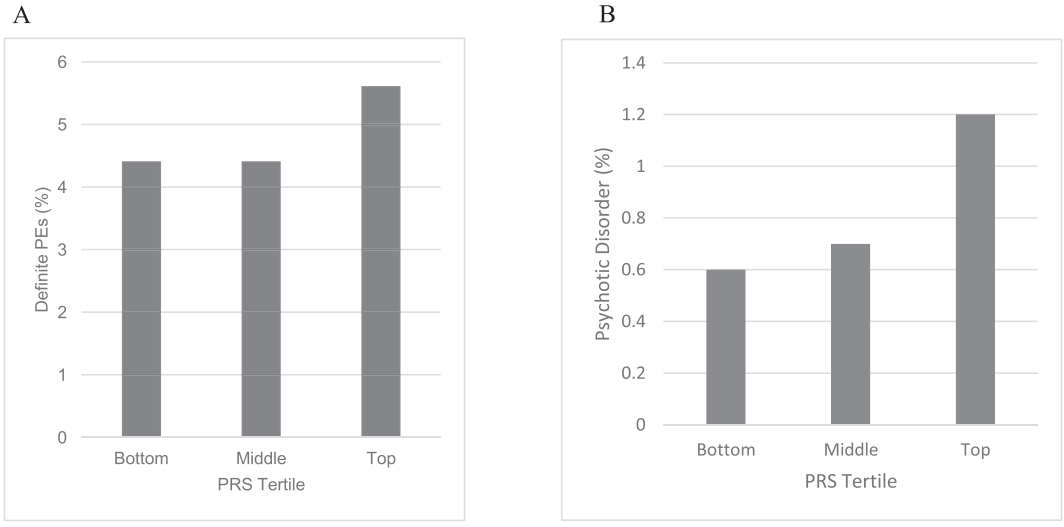


Fig. 1. Prevalence of psychotic experiences and psychotic disorder at age 18 per tertile of PRS-T2DM.

Table 2
Odds ratios (95% CI) for outcomes at age 18 per SD increase in polygenic risk score for T2DM or schizophrenia.

Outcome/risk factor	Sample	OR (95% C.I.)		*p*-value	Corrected *p*-value[c]
		Unadjusted[a]	Adjusted for sex, ethnicity, social class and BMI[b]		
Definite PEs					
PRS-T2DM	3768	1.15 (0.99–1.34)	1.21 (1.01–1.45)	0.027	0.054
Psychotic disorder					
PRS-T2DM	3768	1.42 (1.00–1.96)	1.51 (1.04–2.05)	0.016	0.048*
Insulin resistance					
PRS-SCZ	2344	1.16 (1.04–1.32)	1.10 (0.99–1.22)	0.089	0.089

[a] Unadjusted analysis adjusted for 10 principal components only.
[b] Samples for adjusted analysis included 3070 participants for psychotic outcomes and 1970 participants for insulin resistance outcome.
[c] *p*-value corrected from adjusted analysis using Holm-Bonferroni method.
* Evidence surpasses Holm-Bonferroni threshold.

association weakened at more stringent *p*-value thresholds. A similar pattern was observed for the PRS-schizophrenia-IR association, where the evidence for a positive association attenuated at *p*-value thresholds more stringent than 1.00×10^{-4}.

3.7. Missing data

Fifty-three percent of participants with data on PRS-T2DM had psychotic outcomes data missing, and 71% of participants with PRS-schizophrenia had IR outcome data missing (Supplementary Fig. 1). Compared with the analytic sample, the missing sample had higher mean PRS-schizophrenia but lower PRS-T2DM scores (Supplementary Table 4). Male sex, lower social class and higher BMI predicted missing data for psychotic outcomes, and non-white ethnicity was associated with having missing data for IR (Supplementary Table 5).

4. Discussion

4.1. Main findings and comparisons with the literature

Using prospective birth cohort data, we report that genetic predisposition for T2DM is associated with psychotic outcomes at age 18 in a linear fashion. The PRS-T2DM findings were consistent using two genetic scores; one with and one without a SNP at the *FTO* locus, which is understood to be related to BMI (Frayling et al., 2007). Additionally, there was evidence for a dose-response pattern in the association between PRS-T2DM and psychotic outcomes; the effect size was strongest for psychotic disorder, which is a more clinically relevant outcome than PEs. We also report some evidence, albeit slightly weaker, for an association between genetic predisposition for schizophrenia and IR at age 18. However, the sample of participants with missing data had higher mean PRS-schizophrenia scores than included participants, thus missing data may help to at least partly explain the weaker evidence. Nonetheless, our findings provide some evidence that the comorbidity between T2DM and schizophrenia arises partly due to shared genetic factors.

The point estimates across various *p*-value thresholds for T2DM and schizophrenia were similar in both combinations of genotype-phenotype analysis, though in both cases at more stringent *p*-value thresholds, the evidence of association weakened. This weakening effect is consistent with a previous study examining the association between PRS-schizophrenia and adolescent psychopathology (Jones et al., 2016), which also reported that PRS-schizophrenia was associated with attrition. Therefore, type II statistical error may be one explanation for the weaker associations between PRS-schizophrenia and IR.

Our results are in line with one previous study in a relatively large sample, which found that people with comorbid schizophrenia and T2DM have a higher genetic predisposition to both disorders compared to controls (Hackinger et al., 2018), and another recent report of an association between PRS for schizophrenia and IR in a clinical sample of people with schizophrenia (Tomasik et al., 2019). Another study found evidence for a genetic overlap between schizophrenia and both triglycerides and HDL (Andreassen et al., 2013), which are cardiometabolic indices known to be tightly linked with an insulin resistance phenotype (Laws and Reaven, 1992), alongside other cardiometabolic factors including systolic blood pressure, BMI and waist: hip ratio. One previous study however found no evidence for an association between PRS-T2DM and schizophrenia (Padmanabhan et al., 2016), though the latter

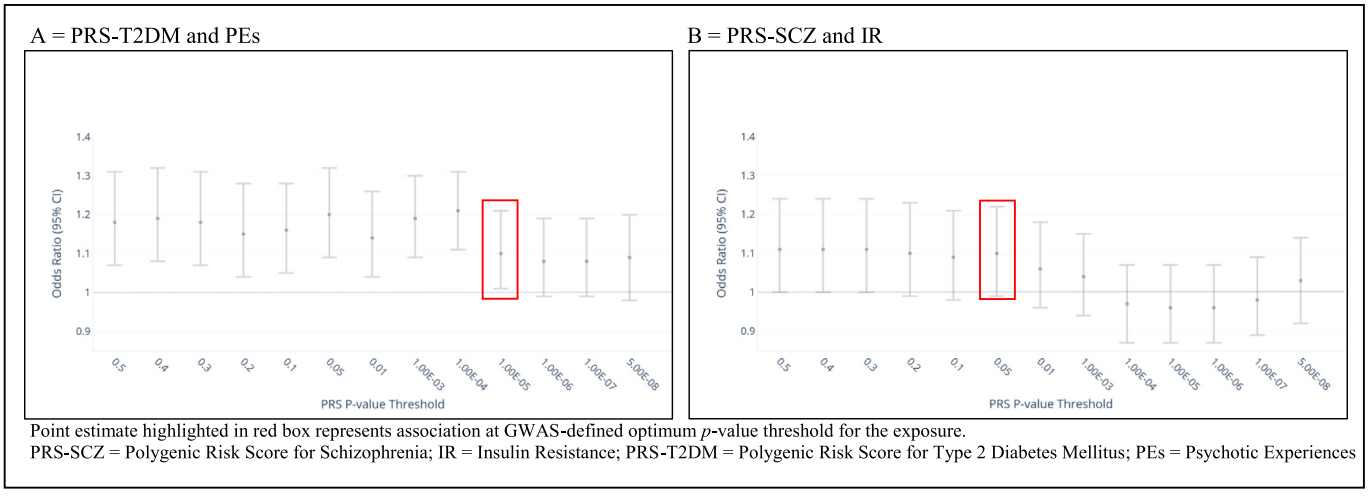


Fig. 2. Association between PRS score and outcome at age 18 years at different PRS *P*-value thresholds.

study featured a much smaller sample size than in our study and may therefore have been underpowered to detect a difference. Another study using LD-score regression (Bulik-Sullivan et al., 2015) found limited evidence for a genetic correlation between schizophrenia and T2DM, though the latter study was based on older and less-powered GWAS for both disorders. However, the same study did find some evidence for genetic correlation between schizophrenia and BMI, and, another recent study provides some evidence for shared genetic loci between BMI and mental disorders including schizophrenia (Bahrami et al., 2020). In future, genetic studies may seek to examine the association between PRS scores for other cardiometabolic traits in their association with schizophrenia and other mental disorders.

It is also possible that genetic-risk for T2DM or schizophrenia may increase the risk of both disorders via pleiotropic mechanisms. This may help to explain the differences in our results compared with genetic correlation analyses (Bulik-Sullivan et al., 2015). For example, it is possible that genetic-risk for schizophrenia may predispose to adverse experiences in childhood, which could in t urn influence inflammation (Slopen et al., 2013). We found some evidence for the association of childhood CRP levels with both PRS-T2DM and PRS-schizophrenia. However, we did not find an association with IL-6. This is perhaps unexpected since IL-6 stimulates the production of CRP (Calabro et al., 2003), and is associated with both psychotic outcomes (Khandaker et al., 2014) and IR (Kim et al., 2009). However, it is also possible that genetic predisposition for T2DM or schizophrenia influences CRP via mechanisms other than IL-6. CRP has been shown to play an active role in hepatic insulin resistance, at least partly through impairment in insulin signalling, independent of IL-6 (Xi et al., 2011). Interestingly, CRP has shown to be protective of schizophrenia in MR studies (Hartwig et al., 2017), however, the GWAS studies included in previous MR research measured phenotypic markers in adults. We used CRP measured in childhood, which may be reflective of a distinct biological environment.

We report some evidence that genetic predisposition for T2DM may influence risk of psychosis in early-adulthood by increasing inflammation in childhood, but the magnitude of this mediating effect was small, suggesting that other mechanisms are likely to be involved. On the other hand, we found no evidence that childhood IL-6/CRP mediated the association between genetic predisposition for schizophrenia and IR. The mediating effect of inflammation for the outcome of PEs is consistent with previous research reporting an association between genetic risk for schizophrenia and immune-related disorders (Stringer et al., 2014; Tylee et al., 2018). However, due to the relatively small number of participants with psychotic disorder in our sample and associated lack of power, we were unable to consider testing psychotic disorder in mediation analyses. Future longitudinal research conducted on larger samples of participants may seek to perform a mediation analysis of CRP between PRS-T2DM and more clinically relevant psychotic outcomes.

Other mediators for PRS-T2DM and psychotic outcomes may include non-immune mechanisms such as pleotropic genes affecting distinct biological pathways relevant for each condition. For example, a study examining the genetic overlap between T2DM and schizophrenia highlighted, among others, *PROX1* as a potentially pleiotropic locus (Hackinger et al., 2018). *PROX1* acts both as a transcriptional activator and repressor. It has been implicated in murine beta-cell development as well as in neurogenesis in humans (Holzmann et al., 2015). Due to the relatively small number of participants with psychotic disorder in our sample and associated lack of power, we were unable to consider testing psychotic disorder in mediation analyses. Future longitudinal research conducted on larger samples of participants may seek to perform a mediation analysis of CRP between PRS-T2DM and more clinically relevant psychotic outcomes.

4.2. Strengths and limitations

In this study, we have examined the influence of genetic predispositions for T2DM and schizophrenia on, respectively, psychosis-risk and T2DM-risk using a prospective birth cohort. We provide some evidence that a genetic basis may explain at least part of the variance of the commonly observed comorbidity between the two phenotypes. In addition, we have used childhood inflammatory marker data to test potential mediating effects of inflammation for these associations. Since our exposures were genetic risk, the potential for confounding by environmental and lifestyle factors is limited. However, it is well known that certain antipsychotic medications can have adverse effects on glycaemic indices (Leucht et al., 2013). At present, ALSPAC does not have treatment record linkage and we were thus unable to adjust for antipsychotic treatment. This may have impacted our results for the analyses examining PRS-schizophrenia and IR. We were able to control for potential confounding effects of sex, BMI, social class and for inflammatory disease. Regarding ethnicity, participants of non-European genetic ancestry were removed at the stage of genotyping analysis. We also adjusted our regression analyses for ethnicity, since ethnicity is significantly associated with T2DM-risk (Oldroyd et al., 2005). We further adjusted for PCs (Price et al., 2006) in our PRS analyses, to further reduce the risk of population stratification bias. By including PRS for schizophrenia in our analyses, we help to address a common limitation of research conducted on PEs, that they may not adequately capture schizophrenia liability (Jones et al., 2016); the results of both sets of analyses were consistent. A key limitation is missing data. Over half of the risk set with data on PRS had outcome data missing at follow-up. The missing sample had a higher mean score for PRS-schizophrenia but a lower mean score for PRS-T2DM. Thus, our analyses may underestimate the true association between genetic predisposition for schizophrenia and IR, whilst the opposite might be the case for the association between PRS-T2DM and psychotic outcomes. Furthermore, whilst PEs and psychotic disorder have been shown to reflect an increased risk for psychotic disorders (Sullivan et al., 2020; Zammit et al., 2013), and PEs lie on a continuum with clinical psychosis in the general population (van Os et al., 2009), our data do not allow us to determine whether people meet criteria for specific psychotic disorders as classified in DSM or ICD. The transition from PEs to clinical psychosis is low (Kaymaz et al., 2012), PEs are also associated with other psychiatric phenotypes such as depressive and anxiety disorders (Kelleher et al., 2012), and previous research has found no evidence of an association between PRS-schizophrenia and PEs (Jones et al., 2016). Additionally, since our psychotic outcomes were measured prior to the peak age of onset of clinical psychosis (Eranti et al., 2013), some participants may not have yet developed psychotic symptoms or disorder. This point also applies to our sample of participants meeting the criteria for IR at age 18, since age 18 may be relatively early for the phenotype to become detectable. This may be a further explanation for the weaker evidence for an association between PRS-schizophrenia and IR at age 18. Whilst we attempted to address these limitations by reversing the genotype and phenotype to more accurately capture schizophrenia/T2DM liability, replication of our methods in an adequately powered clinical (and likely older) sample of people with clinically diagnosed psychotic disorders such as schizophrenia, is necessary. Finally, one-off measurements of inflammatory markers in childhood may not reflect lifelong levels of inflammation. However, measurement error, if non-differential, introduces a bias towards the null, so our results may underestimate the true association between PRS-T2DM and IL-6 and CRP.

Future research may complement our work by employing genomic advances which test a greater proportion of genomic information than solely PRS scores, the latter of which are highly dependent on the power of GWAS studies. Such methods might include colocalization analysis (Giambartolomei et al., 2014) or locus-level genetic correlation analysis (Shi et al., 2017). Such research may build on our own since research conducted using PRS scores may be susceptible to type II error due to the phenomenon of 'missing heritability', which is the difference between the known heritability of a trait compared with the currently identified risk-increasing variants (Manolio et al., 2009). It is likely that at least some of the heritability of schizophrenia (Woo et al.,

2017) as well as cardiometabolic disorders (Xia et al., 2016) lies in a number of low-frequency, low-effect-size variants which are therefore difficult to detect with current GWAS methods.

4.3. Implications

Our work provides some evidence that, limitations notwithstanding, a summation of minor genetic variation representing lifetime risk for T2DM or schizophrenia at conception, may contribute a portion of the variance of the comorbidity of these disorders in adulthood. Furthermore, we report that genetic predisposition for T2DM may increase risk of PEs by influencing physiologic changes, such as low-grade inflammation, detectable as early as childhood. It is well known that some commonly prescribed antipsychotics can cause or worsen cardiometabolic indices (Leucht et al., 2013), even after a relatively short length of exposure (Neilsen et al., 2010). Therefore, clinicians who look after people with schizophrenia should ascribe detailed attention to the malleable risk factors for cardiometabolic disorders, such as with the promotion of a healthy lifestyle (Teasdale et al., 2019; Ward et al., 2017), and with careful selection and monitoring of antipsychotic medications. This may help to reduce the excess cardiometabolic illness related morbidity and mortality in people with schizophrenia. In future, similar research may seek to examine the associations between PRS for T2DM and other mental disorders including T2DM and bipolar disorder, both of which are known to have higher rates of cardiometabolic disorders than the general population (Martin et al., 2016). Such research may also help to test the specificity of the findings in this study.

Declaration of competing interest

The authors report no competing interests.

Acknowledgements

The authors are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses.

Funding body agreements and policies

This report is independent research supported by the National Institute for Health Research (NIHR Doctoral Research Fellowship, BIP, DRF-2018-11-ST2-018). The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health and Social Care. GMK acknowledges funding support from the MQ: Transforming Mental Health (Data Science Award; grant code: MQDS17/40), the Wellcome Trust (Intermediate Clinical Fellowship; grant code: 201486/Z/16/Z), and the Medical Research Council (MICA: Mental Health Data Pathfinder; grant code: MC_PC_17213). PBJ acknowledges funding from MRC (DPFS), programmatic funding from National Institute for Health Research (RP-PG- 0616-20003) and support from Collaboration for Leadership in Applied Health Research and Care (CLAHRC) East of England. SZ is supported by the NIHR Biomedical Research Centre at University Hospitals Bristol NHS Foundation Trust and the University of Bristol. TGR is a UKRI Innovation Research Fellow (MR/S003886/1). The UK Medical Research Council and Wellcome Trust (Grant no: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. A comprehensive list of grants funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf/); this research was specifically funded by The Wellcome Trust (Grant no: 08426812/Z/07/Z). GWAS data was generated by Sample Logistics and Genotyping Facilities at Wellcome Sanger Institute and LabCorp (Laboratory Corporation of America) using support from 23andMe.

CRediT authorship contribution statement

BIP and GK conceived the study. Analysis was done by BIP, TR and HJJ. BIP wrote the manuscript, with edits suggested by HJJ, TR, SZ, GL, PBJ, NJW, GMK.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.schres.2020.08.006.

References

Andreassen, O.A., Djurovic, S., Thompson, W.K., Schork, A.J., Kendler, K.S., O'Donovan, M.C., Rujescu, D., Werge, T., van de Bunt, M., Morris, A.P., McCarthy, M.I., International Consortium for Blood Pressure, G, Diabetes Genetics, R, Meta-analysis, C, Psychiatric Genomics Consortium Schizophrenia Working, G, Roddey, J.C., McEvoy, L.K., Desikan, R.S., Dale, A.M., 2013. Improved detection of common variants associated with schizophrenia by leveraging pleiotropy with cardiovascular-disease risk factors. Am. J. Hum. Genet. 92 (2), 197–209.

Bahrami, S., Steen, N.E., Shadrin, A., O'Connell, K., Frei, O., Bettella, F., Wirgenes, K.V., Krull, F., Fan, C.C., Dale, A.M., Smeland, O.B., Djurovic, S., Andreassen, O.A., 2020. Shared genetic loci between body mass index and major psychiatric disorders: a genome-wide association study. JAMA Psychiatry 77 (5), 503–512.

Boyd, A., Golding, J., Macleod, J., Lawlor, D.A., Fraser, A., Henderson, J., Molloy, L., Ness, A., Ring, S., Davey Smith, G., 2013. Cohort profile: the 'Children of the 90s'–the index offspring of the Avon longitudinal study of parents and children. Int. J. Epidemiol. 42 (1), 111–127.

Bulik-Sullivan, B., Finucane, H.K., Anttila, V., Gusev, A., Day, F.R., Loh, P.R., ReproGen, C, Psychiatric Genomics, C, Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control, C, Duncan, L., Perry, J.R., Patterson, N., Robinson, E.B., Daly, M.J., Price, A.L., Neale, B.M., 2015. An atlas of genetic correlations across human diseases and traits. Nat. Genet. 47 (11), 1236–1241.

Calabro, P., Willerson, J.T., Yeh, E.T., 2003. Inflammatory cytokines stimulated C-reactive protein production by human coronary artery smooth muscle cells. Circulation 108 (16), 1930–1932.

Cediel, G., Corvalan, C., Aguirre, C., de Romana, D.L., Uauy, R., 2016. Serum 25-hydroxyvitamin D associated with indicators of body fat and insulin resistance in prepubertal chilean children. Int. J. Obes. 40 (1), 147–152.

Chang, C.C., Chow, C.C., Tellier, L.C.A.M., Vattikuti, S., Purcell, S.M., Lee, J.J., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 4.

Chouinard, V.A., Henderson, D.C., Dalla Man, C., Valeri, L., Gray, B.E., Ryan, K.P., Cypess, A.M., Cobelli, C., Cohen, B.M., Ongur, D., 2019. Impaired insulin signaling in unaffected siblings and patients with first-episode psychosis. Mol. Psychiatry 24 (10), 1513–1522.

Eranti, S.V., MacCabe, J.H., Bundy, H., Murray, R.M., 2013. Gender difference in age at onset of schizophrenia: a meta-analysis. Psychol. Med. 43 (1), 155–167.

Ferentinos, P., Dikeos, D., 2012. Genetic correlates of medical comorbidity associated with schizophrenia and treatment with antipsychotics. Curr Opin Psychiatry 25 (5), 381–390.

Festa, A., D'Agostino Jr., R., Howard, G., Mykkanen, L., Tracy, R.P., Haffner, S.M., 2000. Chronic subclinical inflammation as part of the insulin resistance syndrome: the Insulin Resistance Atherosclerosis Study (IRAS). Circulation 102 (1), 42–47.

Fraser, A., Macdonald-Wallis, C., Tilling, K., Boyd, A., Golding, J., Davey Smith, G., Henderson, J., Macleod, J., Molloy, L., Ness, A., Ring, S., Nelson, S.M., Lawlor, D.A., 2013. Cohort profile: the Avon longitudinal study of parents and children: ALSPAC mothers cohort. Int. J. Epidemiol. 42 (1), 97–110.

Frayling, T.M., Timpson, N.J., Weedon, M.N., Zeggini, E., Freathy, R.M., Lindgren, C.M., Perry, J.R., Elliott, K.S., Lango, H., Rayner, N.W., Shields, B., Harries, L.W., Barrett, J.C., Ellard, S., Groves, C.J., Knight, B., Patch, A.M., Ness, A.R., Ebrahim, S., Lawlor, D.A., Ring, S.M., Ben-Shlomo, Y., Jarvelin, M.R., Sovio, U., Bennett, A.J., Melzer, D., Ferrucci, L., Loos, R.J., Barroso, I., Wareham, N.J., Karpe, F., Owen, K.R., Cardon, L.R., Walker, M., Hitman, G.A., Palmer, C.N., Doney, A.S., Morris, A.D., Smith, G.D., Hattersley, A.T., McCarthy, M.I., 2007. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science 316 (5826), 889–894.

Geloneze, B., Repetto, E.M., Geloneze, S.R., Tambascia, M.A., Ermetice, M.N., 2006. The threshold value for insulin resistance (HOMA-IR) in an admixtured population IR in the Brazilian Metabolic Syndrome Study. Diabetes Res. Clin. Pract. 72 (2), 219–220.

Giambartolomei, C., Vukcevic, D., Schadt, E.E., Franke, L., Hingorani, A.D., Wallace, C., Plagnol, V., 2014. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet. 10 (5), e1004383.

Hackinger, S., Prins, B., Mamakou, V., Zengini, E., Marouli, E., Brčić, L., Serafetinidis, I., Lamnissou, K., Kontaxakis, V., Dedoussis, G., Gonidakis, F., Thanopoulou, A., Tentolouris, N., Tsezou, A., Zeggini, E., 2018. Evidence for genetic contribution to the increased risk of type 2 diabetes in schizophrenia. Transl. Psychiatry 8 (1), 252.

Hartwig, F.P., Borges, M.C., Horta, B.L., Bowden, J., Davey Smith, G., 2017. Inflammatory biomarkers and risk of schizophrenia: a 2-sample Mendelian randomization study. JAMA Psychiatry 74 (12), 1226–1233.

Hedblad, B., Nilsson, P., Janzon, L., Berglund, G., 2000. Relation between insulin resistance and carotid intima-media thickness and stenosis in non-diabetic subjects. Results from a cross-sectional study in Malmo, Sweden. Diabet. Med. 17 (4), 299–307.

Holland, B.S., Copenhaver, M., 1987. An improved sequentially rejective Bonferroni test procedure. Biometrics 43 (2), 417–423.

Holt, R.I., Peveler, R.C., Byrne, C.D., 2004. Schizophrenia, the metabolic syndrome and diabetes. Diabet. Med. 21 (6), 515–523.

Holzmann, J., Hennchen, M., Rohrer, H., 2015. Prox1 identifies proliferating neuroblasts and nascent neurons during neurogenesis in sympathetic ganglia. Dev Neurobiol 75 (12), 1352–1367.

Jones, H.J., Stergiakouli, E., Tansey, K.E., Hubbard, L., Heron, J., Cannon, M., Holmans, P., Lewis, G., Linden, D.E., Jones, P.B., Davey Smith, G., O'Donovan, M.C., Owen, M.J., Walters, J.T., Zammit, S., 2016. Phenotypic manifestation of genetic risk for schizophrenia during adolescence in the general population. JAMA Psychiatry 73 (3), 221–228.

Kaymaz, N., Drukker, M., Lieb, R., Wittchen, H.U., Werbeloff, N., Weiser, M., Lataster, T., van Os, J., 2012. Do subthreshold psychotic experiences predict clinical outcomes in

unselected non-help-seeking population-based samples? A systematic review and meta-analysis, enriched with new results. Psychol. Med. 42 (11), 2239–2253.
Kelleher, I., Keeley, H., Corcoran, P., Lynch, F., Fitzpatrick, C., Devlin, N., Molloy, C., Roddy, S., Clarke, M.C., Harley, M., Arseneault, L., Wasserman, C., Carli, V., Sarchiapone, M., Hoven, C., Wasserman, D., Cannon, M., 2012. Clinicopathological significance of psychotic experiences in non-psychotic young people: evidence from four population-based studies. Br. J. Psychiatry 201 (1), 26–32.
Khandaker, G.M., Pearson, R.M., Zammit, S., Lewis, G., Jones, P.B., 2014. Association of serum interleukin 6 and C-reactive protein in childhood with depression and psychosis in young adult life: a population-based longitudinal study. JAMA Psychiatry 71 (10), 1121–1128.
Kim, J.H., Bachmann, R.A., Chen, J., 2009. Interleukin-6 and insulin resistance. Vitam. Horm. 80, 613–633.
Lappin, J.M., Wijaya, M., Watkins, A., Morell, R., Teasdale, S., Lederman, O., Rosenbaum, S., Dick, S., Ward, P., Curtis, J., 2018. Cardio-metabolic risk and its management in a cohort of clozapine-treated outpatients. Schizophr. Res. 199, 367–373.
Laws, A., Reaven, G.M., 1992. Evidence for an independent relationship between insulin resistance and fasting plasma HDL-cholesterol, triglyceride and insulin concentrations. J. Intern. Med. 231 (1), 25–30.
Leucht, S., Cipriani, A., Spineli, L., Mavridis, D., Orey, D., Richter, F., Samara, M., Barbui, C., Engel, R.R., Geddes, J.R., Kissling, W., Stapf, M.P., Lassig, B., Salanti, G., Davis, J.M., 2013. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. Lancet 382 (9896), 951–962.
Lin, P.I., Shuldiner, A.R., 2010. Rethinking the genetic basis for comorbidity of schizophrenia and type 2 diabetes. Schizophr. Res. 123 (2–3), 234–243.
Mahajan, A., Go, M.J., Zhang, W., Below, J.E., Gaulton, K.J., Ferreira, T., Horikoshi, M., Johnson, A.D., Ng, M.C., Prokopenko, I., Saleheen, D., Wang, X., Zeggini, E., Abecasis, G.R., Adair, L.S., Almgren, P., Atalay, M., Aung, T., Baldassarre, D., Balkau, B., Bao, Y., Barnett, A.H., Barroso, I., Basit, A., Been, L.F., Beilby, J., Bell, G.I., Benediktsson, R., Bergman, R.N., Boehm, B.O., Boerwinkle, E., Bonnycastle, L.L., Burtt, N., Cai, Q., Campbell, H., Carey, J., Cauchi, S., Caulfield, M., Chan, J.C., Chang, L.C., Chang, T.J., Chang, Y.C., Charpentier, G., Chen, C.H., Chen, H., Chen, Y.T., Chia, K.S., Chidambaram, M., Chines, P.S., Cho, N.H., Cho, Y.M., Chuang, L.M., Collins, F.S., Cornelis, M.C., Couper, D.J., Crenshaw, A.T., van Dam, R.M., Danesh, J., Das, D., de Faire, U., Dedoussis, G., Deloukas, P., Dimas, A.S., Dina, C., Doney, A.S., Donnelly, P.J., Dorkhan, M., van Duijn, C., Dupuis, J., Edkins, S., Elliott, P., Emilsson, V., Erbel, R., Eriksson, J.G., Escobedo, J., Esko, T., Eury, E., Florez, J.C., Fontanillas, P., Forouhi, N.G., Forsen, T., Fox, C., Fraser, R.M., Frayling, T.M., Froguel, P., Frossard, P., Gao, Y., Gertow, K., Gieger, C., Gigante, B., Grallert, H., Grant, G.B., Grrop, L.C., Groves, C.J., Grundberg, E., Guiducci, C., Hamsten, A., Han, B.G., Hara, K., Hassanali, N., Hattersley, A.T., Hayward, C., Hedman, A.K., Herder, C., Hofman, A., Holmen, O.L., Hovingh, K., Hreidarsson, A.B., Hu, C., Hu, F.B., Hui, J., Humphries, S.E., Hunt, S.E., Hunter, D.J., Hveem, K., Hydrie, Z.I., Ikegami, H., Illig, T., Ingelsson, E., Islam, M., Isomaa, B., Jackson, A.U., Jafar, T., James, A., Jia, W., Jöckel, K.H., Jonsson, A., Jowett, J.B., Kadowaki, T., Kang, H.M., Kanoni, S., Kao, W.H., Kathiresan, S., Kato, N., Katulanda, P., Keinanen-Kiukaanniemi, K.M., Kelly, A.M., Khan, H., Khaw, K.T., Khor, C.C., Kim, H.L., Kim, S., Kim, Y.J., Kinnunen, L., Klopp, N., Kong, A., Korpi-Hyövälti, E., Kowlessur, S., Kraft, P., Kravic, J., Kristensen, M.M., Krithika, S., Kumar, A., Kumate, J., Kuusisto, J., Kwak, S.H., Laakso, M., Lagou, V., Lakka, T.A., Langenberg, C., Langford, C., Lawrence, R., Leander, K., Lee, J.M., Lee, N.R., Li, M., Li, X., Li, Y., Liang, J., Liju, S., Lim, W.Y., Lind, L., Lindgren, C.M., Lindholm, E., Liu, C.T., Liu, J.J., Lobbens, S., Long, J., Loos, R.J., Lu, W., Luan, J., Lyssenko, V., Ma, R.C., Maeda, S., Mägi, R., Männisto, S., Matthews, D.R., Meigs, J.B., Melander, O., Metspalu, A., Meyer, J., Mirza, G., Mihailov, E., Moebus, S., Mohan, V., Mohlke, K.L., Morris, A.D., Mühleisen, T.W., Müller-Nurasyid, M., Musk, B., Nakamura, J., Nakashima, E., Navarro, P., Ng, P.K., Nica, A.C., Nilsson, P.M., Njølstad, I., Nöthen, M.M., Ohnaka, K., Ong, T.H., Owen, K.R., Palmer, C.N., Pankow, J.S., Park, K.S., Parkin, M., Pechlivanis, S., Pedersen, N.L., Peltonen, L., Perry, J.R., Peters, A., Pinidiyapathirage, J.M., Platou, C.G., Potter, S., Price, J.F., Qi, L., Radha, V., Rallidis, L., Rasheed, A., Rathman, W., Rauramaa, R., Raychaudhuri, S., Rayner, N.W., Rees, S.D., Rehnberg, E., Ripatti, S., Robertson, N., Roden, M., Rossin, E.J., Rudan, I., Rybin, D., Saaristo, T.E., Salomaa, V., Saltevo, J., Samuel, M., Sanghera, D.K., Saramies, J., Scott, J., Scott, L.J., Scott, R.A., Segrè, A.V., Sehmi, J., Sennblad, B., Shah, N., Shah, S., Shera, A.S., Shu, X.O., Shuldiner, A.R., Sigurđsson, G., Sijbrands, E., Silveira, A., Sim, X., Sivapalaratnam, S., Small, K.S., So, W.Y., Stančáková, A., Stefansson, K., Steinbach, G., Steinthorsdottir, V., Stirrups, K., Strawbridge, R.J., Stringham, H.M., Sun, Q., Suo, C., Syvänen, A.C., Takayanagi, R., Takeuchi, F., Tay, W.T., Teslovich, T.M., Thorand, B., Thorleifsson, G., Thorsteinsdottir, U., Tikkanen, E., Trakalo, J., Tremoli, E., Trip, M.D., Tsai, F.J., Tuomi, T., Tuomilehto, J., Uitterlinden, A.G., Valladares-Salgado, A., Vedantam, S., Veglia, F., Voight, B.F., Wang, C., Wareham, N.J., Wennauer, R., Wickremasinghe, A.R., Wilsgaard, T., Wilson, J.F., Wiltshire, S., Winckler, W., Wong, T.Y., Wood, A.R., Wu, J.Y., Wu, Y., Yamamoto, K., Yamauchi, T., Yang, M., Yengo, L., Yokota, M., Young, R., Zabaneh, D., Zhang, F., Zhang, R., Zheng, W., Zimmet, P.Z., Altshuler, D., Bowden, D.W., Cho, Y.S., Cox, N.J., Cruz, M., Hanis, C.L., Kooner, J., Lee, J.Y., Seielstad, M., Teo, Y.Y., Boehnke, M., Parra, E.J., Chambers, J.C., Tai, E.S., McCarthy, M.I., Morris, A.P., Consortium, D.G.R.A.M.-a.D, Consortium, A.G.E.N.T.D.A.-T.D, Consortium, S.A.T.D.S.D, Consortium, M.A.T.D.M.D, Consortium, T.D.G.E.b.N.-g.s.i.m.-E.S.T.D.-G, 2014. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. Nat. Genet. 46 (3), 234–244.
Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., Cho, J.H., Guttmacher, A.E., Kong, A., Kruglyak, L., Mardis, E., Rotimi, C.N., Slatkin, M., Valle, D., Whittemore, A.S., Boehnke, M., Clark, A.G., Eichler, E.E., Gibson, G., Haines, J.L., Mackay, T.F., McCarroll, S.A., Visscher, P.M., 2009. Finding the missing heritability of complex diseases. Nature 461 (7265), 747–753.
Marques-Vidal, P., Mazoyer, E., Bongard, V., Gourdy, P., Ruidavets, J.B., Drouet, L., Ferrieres, J., 2002. Prevalence of insulin resistance syndrome in southwestern France and its relationship with inflammatory and hemostatic markers. Diabetes Care 25 (8), 1371–1377.
Martin, B.C., Warram, J.H., Krolewski, A.S., Bergman, R.N., Soeldner, J.S., Kahn, C.R., 1992. Role of glucose and insulin resistance in development of type 2 diabetes mellitus: results of a 25-year follow-up study. Lancet 340 (8825), 925–929.
Martin, D.J., Ul-Haq, Z., Nicholl, B.I., Cullen, B., Evans, J., Gill, J.M., Roberts, B., Gallacher, J., Mackay, D., McIntosh, A., Hotopf, M., Craddock, N., Deary, I.J., Pell, J.P., Smith, D.J., 2016. Cardiometabolic disease and features of depression and bipolar disorder: population-based, cross-sectional study. Br. J. Psychiatry 208 (4), 343–351.
Matthews, D.R., Hosker, J.P., Rudenski, A.S., Naylor, B.A., Treacher, D.F., Turner, R.C., 1985. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 28 (7), 412–419.
Neilsen, J., Skadhede, S., Correll, C.U., 2010. Antipsychotics associated with the development of type 2 diabetes in antipsychotic-naive schizophrenia patients. Neuropsychopharmacology 35 (9), 1997–2004.
Oldroyd, J., Banerjee, M., Heald, A., Cruickshank, K., 2005. Diabetes and ethnic minorities. Postgrad. Med. J. 81 (958), 486–490.
Padmanabhan, J.L., Nanda, P., Tandon, N., Mothi, S.S., Bolo, N., McCarroll, S., Clementz, B.A., Gershon, E.S., Pearlson, G.D., Sweeney, J.A., Tamminga, C.A., Keshavan, M.S., 2016. Polygenic risk for type 2 diabetes mellitus among individuals with psychosis and their relatives. J. Psychiatr. Res. 77, 52–58.
Perry, B.I., McIntosh, G., Weich, S., Singh, S., Rees, K., 2016. The association between first-episode psychosis and abnormal glycaemic control: systematic review and meta-analysis. Lancet Psychiatry 3 (11), 1049–1058.
Perry, B.I., Upthegrove, R., Thompson, A., Marwaha, S., Zammit, S., Singh, S.P., Khandaker, G., 2018. Dysglycaemia. Findings From the UK ALSPAC Birth Cohort. Schizophr Bull, Inflammation and Psychosis.
Pillinger, T., Beck, K., Gobjila, C., Donocik, J.G., Jauhar, S., Howes, O.D., 2017. Impaired glucose homeostasis in first-episode schizophrenia: a systematic review and meta-analysis. JAMA Psychiatry 74 (3), 261–269.
Poulton, R., Caspi, A., Moffitt, T.E., Cannon, M., Murray, R., Harrington, H., 2000. Children's self-reported psychotic symptoms and adult schizophreniform disorder: a 15-year longitudinal study. Arch. Gen. Psychiatry 57 (11), 1053–1058.
Pradhan, A.D., Manson, J.E., Rifai, N., Buring, J.E., Ridker, P.M., 2001. C-reactive protein, interleukin 6, and risk of developing type 2 diabetes mellitus. JAMA 286 (3), 327–334.
Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A., Reich, D., 2006. Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. 38 (8), 904–909.
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575.
Purcell, S.M., Wray, N.R., Stone, J.L., Visscher, P.M., O'Donovan, M.C., Sullivan, P.F., Sklar, P., International Schizophrenia Consortium, 2009. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature 460, 748–752.
R Core Team, 2017. R: A Language and Environment for Statistical Computing, Vienna, Austria.
Rajkumar, A.P., Horsdal, H.T., Wimberley, T., Cohen, D., Mors, O., Borglum, A.D., Gasse, C., 2017. Endogenous and antipsychotic-related risks for diabetes mellitus in Young people with schizophrenia: a Danish population-based cohort study. Am. J. Psychiatry 174 (7), 686–694.
Reinikainen, J., Laatikainen, T., Karvanen, J., Tolonen, H., 2015. Lifetime cumulative risk factors predict cardiovascular disease mortality in a 50-year follow-up study in Finland. Int. J. Epidemiol. 44 (1), 108–116.
Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014. Biological insights from 108 schizophrenia-associated genetic loci. Nature 511, 421–427.
Shi, H., Mancuso, N., Spendlove, S., Pasaniuc, B., 2017. Local genetic correlation gives insights into the shared genetic architecture of complex traits. Am. J. Hum. Genet. 101 (5), 737–751.
Slopen, N., Kubzansky, L.D., McLaughlin, K.A., Koenen, K.C., 2013. Childhood adversity and inflammatory processes in youth: a prospective study. Psychoneuroendocrinology 38 (2), 188–200.
Stringer, S., Kahn, R.S., de Witte, L.D., Ophoff, R.A., Derks, E.M., 2014. Genetic liability for schizophrenia predicts risk of immune disorders. Schizophr. Res. 159 (2–3), 347–352.
Sullivan, S.A., Kounali, D., Cannon, M., David, A.S., Fletcher, P.C., Holmans, P., Jones, H., Jones, P.B., Linden, D.E.J., Lewis, G., Owen, M.J., O'Donovan, M., Rammos, A., Thompson, A., Wolke, D., Heron, J., Zammit, S., 2020. A population-based cohort study examining the incidence and impact of psychotic experiences from childhood to adulthood, and prediction of psychotic disorder. Am. J. Psychiatry 177 (4), 308–317 appiajp201919060654.
Teasdale, S.B., Curtis, J., Ward, P.B., Watkins, A., Lederman, O., Rosenbaum, S., Kalucy, M., Lappin, J., Samaras, K., 2019. The effectiveness of the keeping the body in mind Xtend pilot lifestyle program on dietary intake in first-episode psychosis: two-year outcomes. Obes. Res. Clin. Pract. 13 (2), 214–216.
Tomasik, J., Lago, S.G., Vazquez-Bourgon, J., Papiol, S., Suarez-Pinilla, P., Crespo-Facorro, B., Bahn, S., 2019. Association of insulin resistance with schizophrenia polygenic risk score and response to antipsychotic treatment. JAMA Psychiatry 76 (8), 864–867.
Tylee, D.S., Sun, J., Hess, J.L., Tahir, M.A., Sharma, E., Malik, R., Worrall, B.B., Levine, A.J., Martinson, J.J., Nejentsev, S., Speed, D., Fischer, A., Mick, E., Walker, B.R., Crawford, A., Grant, S.F.A., Polychronakos, C., Bradfield, J.P., Sleiman, P.M.A., Hakonarson, H., Ellinghaus, E., Elder, J.T., Tsoi, L.C., Trembath, R.C., Barker, J.N., Franke, A., Dehghan, A., Me Research, T, Inflammation Working Group of the, C.C, Consortium, M.C.o.t.I.S. G, Netherlands Twin, R, neuro, C.W.G, Obsessive, C., Tourette Syndrome Working Group of the Psychiatric Genomics, C, Faraone, S.V., Glatt, S.J., 2018. Genetic

correlations among psychiatric and immune-related phenotypes based on genome-wide association data. Am. J. Med. Genet. B Neuropsychiatr. Genet. 177 (7), 641–657.

Upthegrove, R., Manzanares-Teson, N., Barnes, N.M., 2014. Cytokine function in medication-naive first episode psychosis: a systematic review and meta-analysis. Schizophr. Res. 155 (1–3), 101–108.

van Os, J., Linscott, R.J., Myin-Germeys, I., Delespaul, P., Krabbendam, L., 2009. A systematic review and meta-analysis of the psychosis continuum: evidence for a psychosis proneness-persistence-impairment model of psychotic disorder. Psychol. Med. 39 (2), 179–195.

Wallace, T.M., Levy, J.C., Matthews, D.R., 2004. Use and abuse of HOMA modeling. Diabetes Care 27 (6), 1487–1495.

Ward, P.B., Firth, J., Rosenbaum, S., Samaras, K., Stubbs, B., Curtis, J., 2017. Lifestyle interventions to reduce premature mortality in schizophrenia. Lancet Psychiatry 4 (7), e14.

Woo, H.J., Yu, C., Kumar, K., Reifman, J., 2017. Large-scale interaction effects reveal missing heritability in schizophrenia, bipolar disorder and posttraumatic stress disorder. Transl. Psychiatry 7 (4), e1089.

Xi, L., Xiao, C., Bandsma, R.H., Naples, M., Adeli, K., Lewis, G.F., 2011. C-reactive protein impairs hepatic insulin sensitivity and insulin signaling in rats: role of mitogen-activated protein kinases. Hepatology 53 (1), 127–135.

Xia, C., Amador, C., Huffman, J., Trochet, H., Campbell, A., Porteous, D., Generation, S., Hastie, N.D., Hayward, C., Vitart, V., Navarro, P., Haley, C.S., 2016. Pedigree- and SNP-associated genetics and recent environment are the major contributors to anthropometric and cardiometabolic trait variation. PLoS Genet. 12 (2), e1005804.

Zammit, S., Kounali, D., Cannon, M., David, A.S., Gunnell, D., Heron, J., Jones, P.B., Lewis, S., Sullivan, S., Wolke, D., Lewis, G., 2013. Psychotic experiences and psychotic disorders at age 18 in relation to psychotic experiences at age 12 in a longitudinal population-based cohort study. Am. J. Psychiatry 170 (7), 742–750.

PLOS MEDICINE

RESEARCH ARTICLE

The potential shared role of inflammation in insulin resistance and schizophrenia: A bidirectional two-sample mendelian randomization study

Benjamin I. Perry[1,2]*, Stephen Burgess[3], Hannah J. Jones[4,5], Stan Zammit[4,5,6], Rachel Upthegrove[7], Amy M. Mason[8], Felix R. Day[9], Claudia Langenberg[9], Nicholas J. Wareham[9], Peter B. Jones[1,2], Golam M. Khandaker[1,2]

1 Department of Psychiatry, University of Cambridge School of Clinical Medicine, Cambridge, England, **2** Cambridgeshire and Peterborough NHS Foundation Trust, Cambridge, England, **3** MRC Biostatistics Unit, University of Cambridge, Cambridge, England, **4** NIHR Biomedical Research Centre, University Hospitals Bristol NHS Foundation Trust and University of Bristol, Bristol, United Kingdom, **5** Centre for Academic Mental Health, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, England, **6** MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, Wales, **7** Institute for Mental Health, University of Birmingham, Birmingham, England, **8** Department of Public Health and Primary Care, University of Cambridge, Cambridge, England, **9** MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge, England

* bip20@medschl.cam.ac.uk

Citation: Perry BI, Burgess S, Jones HJ, Zammit S, Upthegrove R, Mason AM, et al. (2021) The potential shared role of inflammation in insulin resistance and schizophrenia: A bidirectional two-sample mendelian randomization study. PLoS Med 18(3): e1003455. https://doi.org/10.1371/journal.pmed.1003455

Academic Editor: Cosetta Minelli, Imperial College London, UNITED KINGDOM

Received: November 7, 2019

Accepted: October 26, 2020

Published: March 12, 2021

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: https://doi.org/10.1371/journal.pmed.1003455

Data Availability Statement: The summary data for the 53 insulin resistance SNPs are available in the supporting information of Lotta et al [11] (10.

Abstract

Background

Insulin resistance predisposes to cardiometabolic disorders, which are commonly comorbid with schizophrenia and are key contributors to the significant excess mortality in schizophrenia. Mechanisms for the comorbidity remain unclear, but observational studies have implicated inflammation in both schizophrenia and cardiometabolic disorders separately. We aimed to examine whether there is genetic evidence that insulin resistance and 7 related cardiometabolic traits may be causally associated with schizophrenia, and whether evidence supports inflammation as a common mechanism for cardiometabolic disorders and schizophrenia.

Methods and findings

We used summary data from genome-wide association studies of mostly European adults from large consortia (Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC) featuring up to 108,557 participants; Diabetes Genetics Replication And Meta-analysis (DIAGRAM) featuring up to 435,387 participants; Global Lipids Genetics Consortium (GLGC) featuring up to 173,082 participants; Genetic Investigation of Anthropometric Traits (GIANT) featuring up to 339,224 participants; Psychiatric Genomics Consortium (PGC) featuring up to 105,318 participants; and Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium featuring up to 204,402 participants). We conducted two-sample uni- and multivariable mendelian randomization (MR) analysis to

1038/ng.3714). Full summary statistics for all traits used in the primary analysis are freely and publicly available for download at consortia/group websites. Specifically; for fasting insulin, FPG, HbA1C and glucose tolerance summary data, see https://www.magicinvestigators.org/downloads/; For HDL, LDL and triglycerides summary data, see http://csg.sph.umich.edu/willer/public/lipids2013/; For BMI summary data, see https://portals.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_file s; For T2DM, see https://diagram-consortium.org/downloads.html; For leptin summary data, see ftp://ftp.ebi.ac.uk/pub/databases/gwas/summary_statistics/KilpelainenTO_26833098_GCST0 03368; For schizophrenia summary data, see https://www.med.unc.edu/pgc/download-results/. Summary GWAS data for CRP, which formed part of our post-hoc sensitivity analysis, are also publicly available from the primary GWAS study [35], and inquiries regarding use of CRP summary data can be sent to s.ligthart@erasmusmc.nl.

Funding: This report is independent research supported by the National Institute for Health Research (NIHR Doctoral Research Fellowship, BIP, DRF-2018-11-ST2-018). The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health and Social Care. GMK acknowledges funding support from the Wellcome Trust (Intermediate Clinical Fellowship; grant code: 201486/Z/16/Z), MQ: Transforming Mental Health (Data Science Award; grant code: MQDS17/40), and the Medical Research Council (MICA: Mental Health Data Pathfinder; grant code: MC_PC_17213). SZ & HJ are supported by the NIHR Biomedical Research Centre, University Hospitals Bristol NHS Foundation Trust and University of Bristol. SB is supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (Grant Number 204623/Z/16/Z). PBJ acknowledges funding from the MRC and MQ (as above), programmatic funding from NIHR (RP-PG- 0616-20003) and support from the Applied Research Collaboration East of England. RU acknowledges funding support from the NIHR (HTA grant code): 127700 and MRC (Therapeutic Target Validation in Mental Health grant code: MR/S037675/1). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: We have read the journal's policy and the authors of this manuscript have the following competing interests. SB is a paid

test whether (i) 10 cardiometabolic traits (fasting insulin, high-density lipoprotein and triglycerides representing an insulin resistance phenotype, and 7 related cardiometabolic traits: low-density lipoprotein, fasting plasma glucose, glycated haemoglobin, leptin, body mass index, glucose tolerance, and type 2 diabetes) could be causally associated with schizophrenia; and (ii) inflammation could be a shared mechanism for these phenotypes. We conducted a detailed set of sensitivity analyses to test the assumptions for a valid MR analysis. We did not find statistically significant evidence in support of a causal relationship between cardiometabolic traits and schizophrenia, or vice versa. However, we report that a genetically predicted inflammation-related insulin resistance phenotype (raised fasting insulin (raised fasting insulin (Wald ratio OR = 2.95, 95% C.I, 1.38–6.34, Holm-Bonferroni corrected *p*-value (*p*) = 0.035) and lower high-density lipoprotein (Wald ratio OR = 0.55, 95% C.I., 0.36–0.84; $p = 0.035$)) was associated with schizophrenia. Evidence for these associations attenuated to the null in multivariable MR analyses after adjusting for C-reactive protein, an archetypal inflammatory marker: (fasting insulin Wald ratio OR = 1.02, 95% C.I, 0.37–2.78, $p = 0.975$), high-density lipoprotein (Wald ratio OR = 1.00, 95% C.I., 0.85–1.16; $p = 0.849$), suggesting that the associations could be fully explained by inflammation. One potential limitation of the study is that the full range of gene products from the genetic variants we used as proxies for the exposures is unknown, and so we are unable to comment on potential biological mechanisms of association other than inflammation, which may also be relevant.

Conclusions

Our findings support a role for inflammation as a common cause for insulin resistance and schizophrenia, which may at least partly explain why the traits commonly co-occur in clinical practice. Inflammation and immune pathways may represent novel therapeutic targets for the prevention or treatment of schizophrenia and comorbid insulin resistance. Future work is needed to understand how inflammation may contribute to the risk of schizophrenia and insulin resistance.

Author summary

Why was this study done?

- Cardiometabolic disorders such as diabetes are up to 30% more common in people with schizophrenia than in the general population, and are among the predominant causes of a 10- to 15-year shortened life expectancy in people with schizophrenia.
- Insulin resistance, a precursor to diabetes, is sometimes detectable in young adults suffering their first episode of psychosis, which suggests that chronic lifestyle and clinical factors, such as smoking, physical inactivity, and medication side effects may not fully explain the comorbidity.
- Inflammation has been consistently associated with schizophrenia and cardiometabolic disorders, and so could be a common mechanism for schizophrenia and cardiometabolic disorders. This could help to at least in part explain why people who have

statistical consultant on PLOS Medicine's statistical board. CL is an Academic Editor on PLOS Medicine's editorial board. PBJ has received honoraria for providing scientific advice to Jansen, Ricordati and Lundbeck.

Abbreviations: BMI, body mass index; CHARGE, Cohorts for Heart and Aging Research in Genomic Epidemiology; CI, confidence interval; CRP, C-reactive protein; DAGs, directed acyclic graphs; CVD, cardiovascular disease; DIAGRAM, Diabetes Genetics Replication And Meta-analysis; FEP, first-episode psychosis; FPG, fasting plasma glucose; GIANT, Genetic Investigation of Anthropometric Traits; GLGC, Global Lipids Genetics Consortium; GWAS, genome-wide association study; HOMA, homeostasis model assessment; HbA1C, glycated haemoglobin; HDL, high-density lipoprotein; IVW, inverse variance weighted; LD, linkage disequilibrium; LDL, low-density lipoprotein; MAGIC, Meta-Analyses of Glucose and Insulin-related traits Consortium; MR, mendelian randomization; MR-PRESSO, MR pleiotropy residual sum and outlier; MVMR, multivariable MR; NLR, neutrophil to lymphocyte ratio; OR, odds ratio; PGC, Psychiatric Genomics Consortium; SD, standard deviation; SEs, standard errors; SNPs, single nucleotide polymorphisms; STROBE, Strengthening the Reporting of Observational Studies in Epidemiology; STROBE-MR, Strengthening the Reporting of MR studies; T2DM, type 2 diabetes mellitus.

schizophrenia also have higher rates of cardiometabolic disorders, over and above the commonly attributed lifestyle/clinical factors.

What did the researchers do and find?

- To examine whether insulin resistance and 7 related cardiometabolic traits causally influence schizophrenia risk or vice versa, we conducted bidirectional, two-sample, uni- and multivariable mendelian randomizsation (MR) analyses. The MR approach uses genetic variants as proxies for modifiable exposures to untangle the problems of reverse causation and unmeasured confounding.
- To test a hypothesis that inflammation may be a common mechanism for schizophrenia and cardiometabolic disorders, we also examined a subset of genetic variants which were associated with inflammation as well as the cardiometabolic trait. We also used multivariable MR (MVMR) as a sensitivity analysis to adjust for C-reactive protein (CRP), an archetypal inflammatory marker, as a general downstream marker of systemic inflammation.
- After correction for multiple testing, overall, there was no significant evidence in support of a causal relationship between cardiometabolic traits and schizophrenia, or vice versa. However, we found evidence that supports a causal relationship of an inflammation-related insulin resistance phenotype with schizophrenia.
- Evidence for the association of an inflammation-related insulin resistance phenotype with schizophrenia attenuated fully in MVMR analysis after adjusting for CRP, suggesting that these associations may be underpinned by inflammation.

What do these findings mean?

- These results suggest that cardiometabolic traits are unlikely to have a causal role in the pathogenesis of schizophrenia or vice versa. However, our results suggest that inflammation is related to the risk of both schizophrenia and insulin resistance, which may at least partly explain why they commonly occur in clinical practice.
- Treating or preventing inflammation may be a putative therapeutic option for prevention and/or treatment of both schizophrenia and comorbid insulin resistance.
- In the future, more research is needed to understand the biological mechanisms underpinning how inflammation may increase the risk of schizophrenia and insulin resistance.

Introduction

Schizophrenia is a complex behavioural and cognitive syndrome characterised primarily by disruptions to perception and cognition [1]. It has a lifetime prevalence of around 0.4% [2] but carries a significant global disease burden [3]. Cardiometabolic disorders are up to 30% more common in schizophrenia than the general population [4] and are the leading contributors to premature death in these patients [5]. Their increased prevalence in schizophrenia is

commonly attributed to the adverse effects of antipsychotic medications [6] or lifestyle factors such as physical inactivity and a poor diet [7], but this is unlikely to be the whole story. While the aforementioned factors contribute cumulative risk over time [8], recent meta-analyses of case–control studies suggest that a phenotype of raised fasting insulin, raised triglycerides, and low high-density lipoprotein (HDL) cholesterol, indicative of insulin resistance [9–11], is associated with relatively young antipsychotic-naïve patients with first-episode psychosis (FEP) [12,13], and, cross-sectionally, with psychotic symptoms in young adults [14]. Therefore, insulin resistance, which is a significant risk factor for type 2 diabetes mellitus (T2DM) and obesity, might be causally related to, or share pathophysiologic mechanisms with schizophrenia.

The majority of existing research in the field is cross-sectional, and therefore cannot confirm whether cardiometabolic disorders are a cause or consequence of illness (i.e., reverse causality). For example, 1 longitudinal study found no evidence for an association between insulin resistance in childhood and risk of psychosis in late adolescence [14]. Additionally, while previous studies have adjusted for a number of potential confounders, residual confounding, which is a limitation of both cross-sectional and longitudinal research, could still be relevant. Mendelian randomization (MR) analysis can address these limitations by using genetic variants inherited randomly at conception as unconfounded proxies of a modifiable exposure, to examine whether the exposure may have a causal effect on a disease outcome [15]. MR studies of cardiometabolic traits and schizophrenia are limited, have focused on a very limited set of cardiometabolic exposures, and have reported mixed findings [16,17]. To our knowledge, MR studies examining associations between a wide range of cardiometabolic traits and schizophrenia are lacking. Such studies may help to identify common potentially causal risk factors and pathophysiologic mechanisms for these physical and psychiatric illnesses.

Inflammation could be pathophysiologically related to cardiometabolic disorders and schizophrenia. Higher levels of circulating inflammatory markers have been associated with both psychosis and cardiometabolic disorders, both cross-sectionally and longitudinally [18–20]. MR studies have reported potential causal associations between inflammation, particularly C-reactive protein (CRP) and interleukin-6 (IL-6), and schizophrenia [21,22]. CRP and IL-6 are also implicated in pathogenesis of insulin resistance [23] and may exaggerate the effects of insulin resistance on psychosis risk in young adults [14]. However, to our knowledge, no MR studies have examined whether inflammation could be pathophysiologically related to insulin resistance and schizophrenia, for example, via mediating or common causal mechanisms.

Therefore, we have conducted a study to examine evidence in support of 4 scenarios regarding the potential relationships between inflammation, insulin resistance, and schizophrenia: (1) Inflammation is a common cause (confounder) between insulin resistance and schizophrenia; (2) insulin resistance mediates an association between inflammation and schizophrenia; or vice versa; (3) inflammation is a common cause (confounder) between schizophrenia and insulin resistance; and (4) schizophrenia mediates an association between inflammation and insulin resistance. See S1 Methods for directed acyclic graphs (DAGs) illustrating the proposed mechanisms.

First, we carried out MR analyses to test whether 10 cardiometabolic traits related to insulin resistance (fasting insulin, triglycerides, HDL, low-density lipoprotein (LDL), fasting plasma glucose (FPG), body mass index (BMI), glucose tolerance, leptin, glycated haemoglobin (HbA1C), and T2DM) could be causally associated with schizophrenia. To test the direction of association, we used genetically predicted levels of cardiometabolic traits as exposures and schizophrenia as the outcome and vice versa. Next, we examined whether inflammation could be a shared mechanism linking insulin resistance and schizophrenia using MR analyses including genetic variants for each cardiometabolic trait that were also associated with a marker of inflammation. Finally, we used multivariable MR (MVMR) analysis to control for

genetic associations of cardiometabolic traits with CRP, an archetypal general inflammatory marker, which we used as a general measure for systemic inflammation.

Methods

Selection of genetic variants related to cardiometabolic traits and schizophrenia

For fasting insulin, triglycerides, and HDL, we used a set of 53 single nucleotide polymorphisms (SNPs) reported to be associated with all 3 traits, representative of an insulin resistance phenotype, from a recent meta genome-wide association study (GWAS) of 188,577 European adults, which adjusted for BMI [11]. In our study, we included SNPs reaching genome-wide significance for the corresponding trait. Summary statistics for genome-wide significant SNPs were also obtained for 6 related continuous (FPG, HbA1C, LDL, BMI, leptin, and glucose tolerance) and 1 binary (T2DM) cardiometabolic traits from recent large GWAS (S2–S10 Methods). We obtained summary statistics for schizophrenia from a recent GWAS from the Psychiatric Genomics Consortium (PGC) [24] based on 40,675 cases and 64,643 European controls. The degree of sample overlap between exposure and outcome samples was likely to be low since exposure and outcome data were obtained from different consortia [25].

Ethics statement

Our study was a secondary analysis of the above publicly available data. Informed consent was sought for all participants per the original GWAS protocols, and all ethical approvals for the GWAS were obtained by original GWAS authors.

Statistical analysis

The analysis plan was prospectively conceived by the authors in 2019 but was not formally deposited in a repository or database. All described analyses were planned *a priori* except for the following: a) the analysis of inflammation-related SNPs at a less-stringent significance threshold (see the 'Analysis using inflammation-related SNPs' section below); b) the MVMR analysis including CRP (see the 'Adjustment for Inflammation' section below). These analyses were conceived and conducted in light of findings from the primary analysis, to further probe whether inflammation could explain the results. We obtained summary-level data (SNP rs number, β-coefficient or log odds ratio (OR), standard errors or 95% confidence intervals (CIs), effect allele, other allele, *p*-value, effect allele frequency, sample size, and number of cases/controls) from each GWAS. Where a specific instrument SNP was not available in the outcome dataset, we located proxy SNPs using linkage disequilibrium (LD) tagging ($r^2 > 0.8$) via LDlink [26]. Alleles were harmonised based on matching alleles, and the resulting instruments were clumped for LD to ensure independence (10,000 kb pairs apart, $r^2 < 0.001$). In the event of palindromic SNPs, the forward strand was inferred where possible using allele frequency information. We performed bidirectional analysis (i.e., with schizophrenia as exposure and cardiometabolic traits as outcomes) to examine direction of association. Statistical analysis was conducted using the TwoSampleMR package (v0.5.4) [27] for R (The R Foundation for Statistical Computing, Vienna, Austria) [28]. Our primary MR analysis method was inverse variance weighted (IVW) regression where at least two exposure SNPs were available for analysis. Where one exposure SNP was available for analysis, we used the Wald ratio method. We also conducted weighted median and MR–Egger regression analysis (S11 Methods). For the binary outcome of schizophrenia, the estimates for continuous exposures (fasting insulin, HDL, triglycerides, LDL, FPG, BMI, HbA1C, glucose tolerance, and leptin) represent log-odds

ratios converted into ORs, representing the increase in risk of schizophrenia per standard deviation (SD) of exposure, and 95% CIs. For binary exposures (T2DM), the estimates represent the OR for schizophrenia per unit increase in the log-odds of T2DM. For continuous cardiometabolic outcomes, β-coefficients represent the SD increase in exposure per unit increase in the log-odds of schizophrenia, with standard errors (SEs).

We performed several sensitivity analyses to check the validity of our results. Heterogeneity among SNPs included in each analysis was examined using the Cochran Q test. We checked for horizontal pleiotropy using the MR–Egger regression intercept alongside a more recent and robust method to detect horizontal pleiotropy and outliers, "MR pleiotropy residual sum and outlier" (MR-PRESSO) [29]. Using MR-PRESSO, we used the global test to examine for horizontal pleiotropy, and where evident, used the method to correct the IVW-estimate via outlier removal (S11 Methods). We examined for measurement error in SNP-exposure associations using the I^2_{GX} statistic [30]. This study is reported as per the Strengthening the Reporting of MR studies (STROBE-MR) guideline [31] (S1 Checklist) and the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement [32] (S2 Checklist).

Analysis using inflammation-related SNPs

Next, we repeated MR analysis using only inflammation-related SNPs for each cardiometabolic risk factor as an instrumental variable for the outcome of schizophrenia. We did this to test the hypothesis that these SNPs may represent a mechanism involving inflammation. This could be via, for example, a common causal basis (panel A in S1 Methods) or via vertical (mediating) pleiotropy [27] (panel B in S1 Methods). We used Phenoscanner v2 (University of Cambridge, United Kingdom) [33] to examine each SNP associated with each cardiometabolic risk factor, to identify SNPs that were also associated with a measure of inflammation, defined as blood concentration/count of cytokines (such as chemokines, interferons, interleukins, lymphokines, or tumour necrosis factors), acute phase proteins (e.g., CRP), or immune cells (e.g., neutrophils and lymphocytes). Primarily, we considered inflammation-related SNPs at genome-wide significance ($p<5\times10^{-8}$) to maximise specificity. We also performed a sensitivity analysis by including inflammation-related SNPs at a less-stringent nominal significance threshold ($p<1\text{x}10^{-4}$) to increase sensitivity to inflammation-related SNPs [34] (S12–S17 Methods).

Using the same method, we identified inflammation-related schizophrenia SNPs (S18 Methods) and used them as instrumental variables in MR analysis examining cardiometabolic traits as outcomes.

Adjustment for inflammation

As a post hoc sensitivity analysis to estimate whether any associations evident above may be explained by inflammation, we carried out MVMR analysis [34,35] using the 53 SNPs for fasting insulin, triglycerides, and HDL, representative of an insulin resistance phenotype, as exposures with schizophrenia as the outcome, after conditioning on the associations of these 53 SNPs with CRP. We chose CRP because it is a widely used downstream measure of systemic inflammation, and publicly available data from large-scale GWAS for CRP are available. Summary statistics for CRP were obtained from a recent large GWAS based on 204,402 participants [36]. For CRP as an exposure in MVMR, we used all independent (10,000 kb pairs apart, $r^2 < 0.001$) SNPs reported to be conditionally associated with CRP and located within the *CRP* coding region (S19 Methods).

Correction for multiple testing

Statistical significance was estimated using the Holm–Bonferroni correction method [37], correcting for the number of exposures tested at each stage of analysis.

Results

MR analyses using all genetic variants associated with insulin resistance and other cardiometabolic traits

We did not find significant evidence for associations between genetically-predicted levels of cardiometabolic traits and schizophrenia, using the primary IVW analysis method. Evidence using the weighted median method for associations between genetically-predicted levels of triglycerides (weighted median OR = 1.26; 95% C.I., 1.06–1.50; corrected $p = 0.090$) and HDL (weighted median OR = 0.79; 95% C.I., 0.65–0.95; corrected $p = 0.126$) with schizophrenia did not survive correction for multiple testing (Table 1).

Table 1. MR analyses of cardiometabolic traits and schizophrenia using all SNPs.

Risk Factor	SNPs, No.[a]	Method	Odds Ratio (95% C.I.)	*p*-value	Corrected *p*-value[b]
Fasting Insulin	9	IVW	1.13 (0.76–1.70)	0.548	1.000
		Weighted Median	0.98 (0.68–1.41)	0.920	1.000
		MR Egger	9.24 (1.82–46.97)	0.028	0.280
Triglycerides	9	IVW	1.16 (0.86–1.56)	0.334	1.000
		Weighted Median	1.26 (1.06–1.50)	0.009	0.090
		MR Egger	1.31 (0.84–2.03)	0.308	1.000
HDL	14	IVW	0.94 (0.71–1.23)	0.649	1.000
		Weighted Median	0.79 (0.65–0.95)	0.010	0.126
		MR Egger	0.67 (0.45–0.99)	0.067	0.670
Fasting Plasma Glucose	18	IVW	1.07 (0.87–1.31)	0.522	1.000
		Weighted Median	1.01 (0.84–1.23)	0.887	1.000
		MR Egger	1.13 (0.74–1.74)	0.584	1.000
Type 2 Diabetes Mellitus	27	IVW	0.93 (0.78–1.12)	0.470	1.000
		Weighted Median	0.93 (0.80–1.09)	0.375	1.000
		MR Egger	1.03 (0.66–1.62)	0.895	1.000
Body Mass Index	81	IVW	1.05 (0.89–1.24)	0.554	1.000
		Weighted Median	1.07 (0.92–1.24)	0.383	1.000
		MR Egger	1.43 (0.97–2.10)	0.103	1.000
HbA1C	36	IVW	1.01 (0.76–1.32)	0.956	1.000
		Weighted Median	1.12 (0.82–1.51)	0.483	1.000
		MR Egger	1.33 (0.79–2.23)	0.295	1.000
Glucose Tolerance	7	IVW	0.98 (0.85–1.14)	0.800	1.000
		Weighted Median	1.10 (0.87–1.15)	0.993	1.000
		MR Egger	1.85 (0.95–3.32)	0.094	0.940
LDL	74	IVW	0.99 (0.93–1.05)	0.679	1.000
		Weighted Median	0.97 (0.90–1.03)	0.322	1.000
		MR Egger	0.98 (0.90–1.07)	0.692	1.000
Leptin	4	IVW	1.97 (0.90–4.31)	0.091	0.910
		Weighted Median	1.18 (0.66–2.11)	0.579	1.000
		MR Egger	3.29 (0.56–17.22)	0.358	1.000

HDL = high-density lipoprotein; HbA1C = glycated haemoglobin; LDL = low-density lipoprotein; IVW = inverse variance weighted regression; SNPs = single nucleotide polymorphisms

[a]Number of SNPs remaining after clumping for independence

[b] Each analysis method (IVW, Weighted Median and MR Egger) corrected using the Holm-Bonferroni method for 10 cardiometabolic markers

Estimates represent ORs for schizophrenia per SD increase in exposure (per unit-increase in log-odds of exposure for T2DM).

https://doi.org/10.1371/journal.pmed.1003455.t001

MR analyses using inflammation-related genetic variants for insulin resistance and other cardiometabolic traits

After testing only genome-wide significant inflammation-related SNPs for cardiometabolic traits, we found evidence for associations of inflammation-related genetically-predicted fasting insulin (Wald Ratio OR = 2.95; 95% C.I., 1.38–6.34; corrected $p = 0.035$) and HDL (Wald Ratio OR = 0.55; 95% CI, 0.36–0.84; corrected $p = 0.035$) with schizophrenia. We could not include any genome-wide significant inflammation-related variants for triglycerides, leptin or glucose tolerance. In our sensitivity analysis featuring inflammation-related cardiometabolic variants at a less stringent significance threshold, evidence persisted for associations of inflammation-related genetically-predicted fasting insulin (IVW OR = 1.74; 95% C.I., 1.08–2.98; corrected $p = 0.030$) and HDL (IVW OR = 0.78; 95% C.I., 0.62–0.92; corrected $p = 0.036$) with schizophrenia. In addition, we found evidence for an association of genetically-predicted inflammation-related triglycerides (IVW OR = 1.24; 95% C.I., 1.07–1.55; corrected $p = 0.036$) with schizophrenia (Table 2; Fig 1 & Fig 2).

Adjustment for inflammation

MVMR analysis for inflammation-related SNPs of fasting insulin, triglycerides, and HDL with schizophrenia showed that the univariable associations fully attenuated after controlling for the genetic associations of these variants with CRP, in analyses involving both inflammation-related SNPs at genome-wide and nominal significance levels. Controlling for CRP had negligible effect on MR estimates based on all genetic variants (Fig 3, S1 and S2 Results).

Test for bidirectionality using schizophrenia as exposure

We did not find statistically significant MR associations between schizophrenia and any cardiometabolic trait after correction for multiple testing (S3 Results, S1 Fig). Similarly, we did not find statistically significant MR associations of inflammation-related schizophrenia variants with cardiometabolic traits after correction for multiple testing (S4 Results, S1 Fig).

Test for horizontal pleiotropy

Using the MR-Egger regression intercept test, we found evidence of potential horizontal pleiotropy for BMI and HDL in the all-SNP analysis, but no evidence for horizontal pleiotropy for any cardiometabolic exposure in the inflammation-related SNP analysis. Using MR-PRESSO however, we found evidence that horizontal pleiotropy was likely to have affected estimates for all cardiometabolic exposures in the all-SNP analysis (p value for global test all ≤ 0.020), and both LDL and T2DM in the inflammation-related SNP analysis. Following MR-PRESSO outlier correction, evidence strengthened for the association of triglycerides with schizophrenia in the all-SNP analysis (MR-PRESSO IVW $\beta = 0.23$, S.E. 0.06, $p = 0.008$), but outlier-corrected IVW estimates for other exposures were not significantly altered.

In the bidirectional analyses, both MR-PRESSO and the MR-Egger regression intercept suggested horizontal pleiotropy affecting the outcomes of HDL, BMI and LDL (all $p<0.05$). Following outlier correction, there was evidence for a weak protective effect of schizophrenia on BMI ($\beta = -0.04$, S.E. 0.02, $p = 0.014$). MR-PRESSO additionally revealed possible horizontal pleiotropy affecting the outcomes of fasting insulin, triglycerides and T2DM (p for MR-PRESSO global test all <0.05) (S5–S12 Results), but outlier-corrected IVW estimates were not significantly altered.

Test for heterogeneity of instruments

In the analyses based on all SNPs, the majority of cardiometabolic traits demonstrated evidence of heterogeneity, which was reduced in the inflammation-related SNP analysis (S5–S8

Table 2. MR analyses of inflammatory-related cardiometabolic SNPs and schizophrenia.

Risk Factor	Method	Genome-Wide Significant Inflammatory-Related SNPs				Nominally Significant Inflammatory-Related SNPs			
		SNPs, No.	Odds Ratio (95% C.I.)	*p*-value	Corrected *p*-value[a]	SNPs, No.	Odds Ratio (95% C.I.)	*p*-value	Corrected *p*-value[a]
Fasting Insulin	IVW / Wald Ratio	1	2.95 (1.38–6.34)	0.005	0.035	5	1.74 (1.08–2.98)	0.003	0.030
	Weighted Median						1.40 (0.83–2.34)	0.203	1.000
	MR Egger						7.20 (1.03–50.54)	0.141	0.987
Triglycerides	IVW / Wald Ratio	0	*	*	*	4	1.24 (1.07–1.55)	0.004	0.036
	Weighted Median						1.26 (1.06–1.50)	0.009	0.063
	MR Egger						1.29 (1.02–1.63)	0.167	0.987
HDL	IVW / Wald Ratio	1	0.55 (0.36–0.84)	0.005	0.035	7	0.78 (0.62–0.92)	0.004	0.036
	Weighted Median						0.77 (0.64–0.94)	0.008	0.056
	MR Egger						0.68 (0.51–0.91)	0.047	0.288
Fasting Plasma Glucose	IVW	2	1.53 (0.39–5.97)	0.537	1.000	4	1.04 (0.36–2.98)	0.945	1.000
	Weighted Median						1.08 (0.63–1.86)	0.776	1.000
	MR Egger						8.44 (0.65–120.54)	0.409	1.000
Type 2 Diabetes Mellitus	IVW	7	0.94 (0.59–1.48)	0.776	1.000	10	0.97 (0.71–1.33)	0.850	1.000
	Weighted Median		1.05 (0.26–4.32)	0.941	1.000		1.05 (0.74–1.48)	0.781	1.000
	MR Egger		1.40 (0.32–6.08)	0.668	1.000		1.42 (0.59–3.38)	0.458	1.000
HbA1C	IVW	7	1.20 (0.67–2.13)	0.546	1.000	10	1.02 (0.64–1.61)	0.942	1.000
	Weighted Median		0.93 (0.46–1.85)	0.832	1.000		0.95 (0.54–1.69)	0.865	1.000
	MR Egger		1.68 (0.39–7.21)	0.508	1.000		1.18 (0.41–3.37)	0.767	1.000
Body Mass Index	IVW	4	1.23 (0.88–1.71)	0.229	1.000	12	1.48 (0.76–2.87)	0.249	1.000
	Weighted Median		1.15 (0.80–1.65)	0.451	1.000		1.16 (0.85–1.58)	0.350	1.000
	MR Egger		0.77 (0.33–1.79)	0.650	1.000		3.36 (0.61–18.45)	0.399	1.000
LDL	IVW	13	0.96 (0.79–1.17)	0.687	1.000	23	0.93 (0.79–1.10)	0.420	1.000
	Weighted Median		0.91 (0.80–1.04)	0.181	1.000		0.91 (0.80–1.04)	0.129	0.987
	MR Egger		0.81 (0.58–1.14)	0.254	1.000		0.82 (0.62–1.11)	0.220	0.987
Leptin	IVW	0	*	*	*	2	1.56 (0.77–3.17)	0.221	0.987
Glucose Tolerance	IVW	0	*	*	*	2	1.06 (0.82–1.56)	0.882	1.000

HDL = high-density lipoprotein; HbA1C = glycated haemoglobin; LDL = low-density lipoprotein; IVW = inverse variance weighted regression; SNPs = single nucleotide polymorphisms

[a]Each analysis method (IVW, Weighted Median and MR Egger) corrected using the Holm-Bonferroni method

*no inflammatory-related SNPs includedEstimates represent ORs for schizophrenia per SD increase in exposure (or per unit-increase in log-odds of binary exposures e.g. T2DM).

https://doi.org/10.1371/journal.pmed.1003455.t002

Results). There was limited evidence of heterogeneity in the sensitivity analyses based on inflammation-related SNPs for T2DM, BMI, and HbA1C only.

Test for measurement error

Results for the I^2_{GX} tests for SNP-exposure associations revealed some evidence for potential measurement error, which may have biased MR–Egger analyses in the analyses with leptin, glucose tolerance, T2DM, and schizophrenia as exposures (S13 Results).

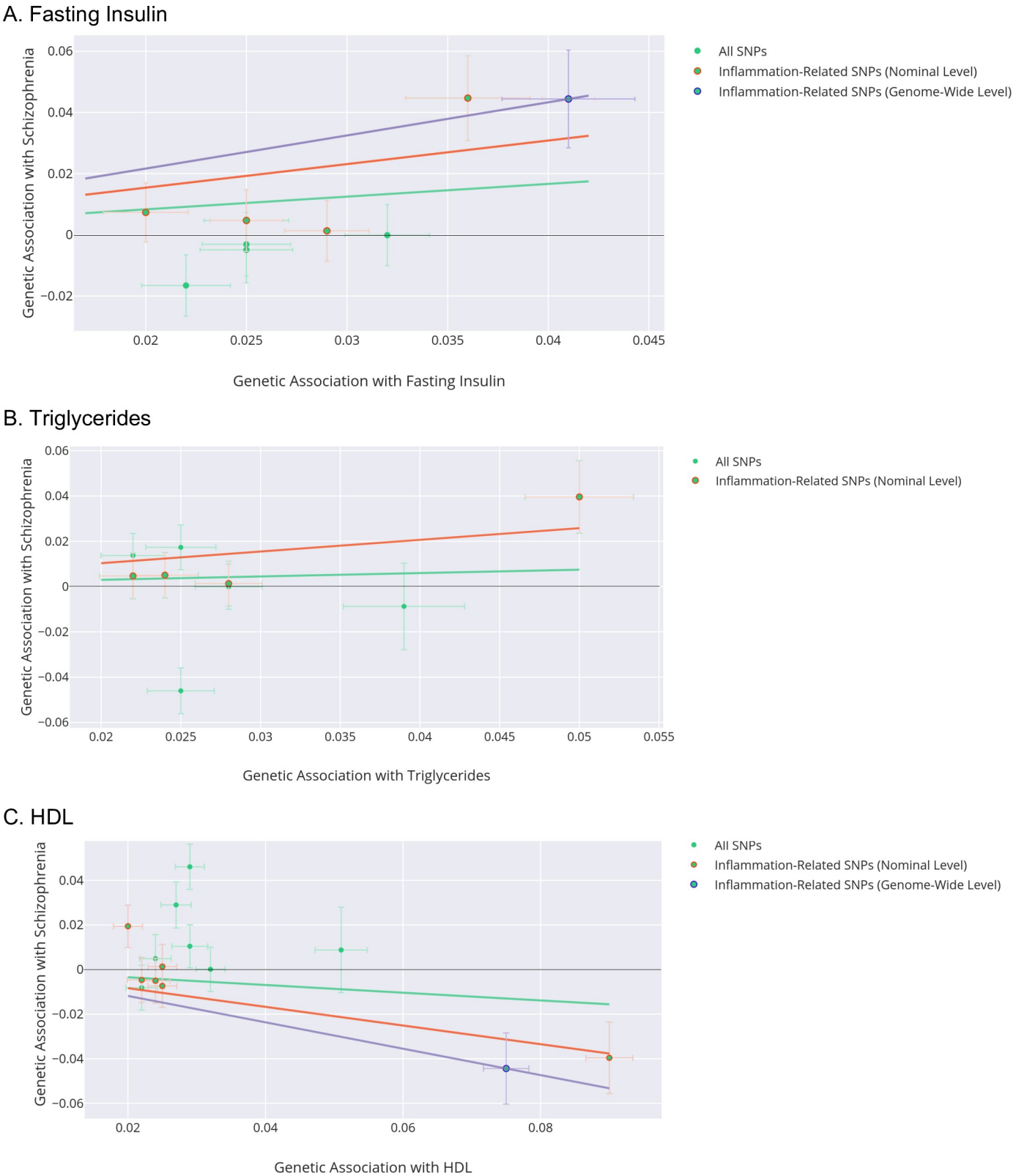


Fig 1. MR Analyses Testing Associations between Insulin Resistance Phenotypes (Fasting Insulin (A), Triglycerides (B) and HDL (C)) and Schizophrenia, Highlighting Inflammation-Related SNPs. Points in plots represent the association of the genome-wide significant insulin-resistance single nucleotide polymorphisms (SNPs) and their association with schizophrenia (Y axis) and the exposure (X axis). SNPs are denoted by green points in the plot. Inflammation-

related SNPs at genome-wide significance are denoted by a purple border. Inflammation-related SNPs at nominal significance are denoted by a red border. Whiskers represent standard errors. Lines on the plot represent inverse-variance weighted (>1 SNP) or linear regression (1 SNP) of all-SNPs (green line), inflammation-related SNPs at genome-wide significance (purple line) and inflammation-related SNPs at nominal significance (red line).

https://doi.org/10.1371/journal.pmed.1003455.g001

Discussion

Main findings

We conducted bidirectional uni- and multivariable two-sample MR analyses using large publicly available genomic datasets to first examine for associations that support a causal relationship between insulin resistance and related cardiometabolic traits and schizophrenia, and second, to examine whether there is evidence to support that inflammation may be a common causal mechanism for insulin resistance and schizophrenia. Using our primary IVW analysis method, we did not find evidence in support of a causal association between genetically-predicted cardiometabolic traits and schizophrenia. However, we found weak evidence using the weighted median method in support of a causal association of genetically-predicted levels of triglycerides and HDL with schizophrenia, but this association did not survive correction for multiple testing and the estimate may have been affected by horizontal pleiotropy. We found more consistent evidence for an association of the insulin resistance phenotype of fasting insulin, triglycerides, and HDL [11] with schizophrenia when we examined only genetic variants also associated with inflammation. Using two *p*-value cut-offs for inflammation-related SNPs, we found that the strength of association with schizophrenia increased as the specificity toward inflammation-related SNPs increased. In MVMR analyses adjusting for CRP, those estimates attenuated fully to the null. We found no evidence in bidirectional analyses in support of a causal relationship of schizophrenia with insulin resistance (panels C and D in S1 Methods). Together, our results are therefore most consistent with inflammation as a common cause for insulin resistance and schizophrenia (panel A in S1 Methods).

Inflammation as a common cause for schizophrenia and insulin resistance

Three aspects of our results point towards inflammation as a common cause for insulin resistance and schizophrenia (panel A in S1 Methods). First, we did not find convincing overall evidence for a causal relationship between insulin resistance and schizophrenia (likely ruling out panel B in S1 Methods). Second, in our analyses of inflammation-related variants for the cardiometabolic traits, we found strong and consistent evidence in support of a potential causal relationship of fasting insulin, HDL and triglycerides with schizophrenia, and the strength of association with schizophrenia increased as the specificity toward inflammation-related SNPs increased. Third, we used MVMR to evidence that after controlling for CRP, an archetypal general inflammatory marker, the associations between inflammation-related genetic variants for insulin resistance and schizophrenia completely attenuated. This result suggests that the observed associations for the inflammation-related variants are at least in part explained by inflammation. Together, the results are consistent with the idea that inflammation may be a common causal mechanism for insulin resistance and schizophrenia.

Evidence for a common causal mechanism between insulin resistance and schizophrenia may help to explain why schizophrenia is associated with higher rates of insulin resistance even in early stages of illness, when the cumulative effects of medication and lifestyle factors are relatively small [12,38]. Anti-inflammatory agents, of which several have shown promise in treating the symptoms of schizophrenia [39], should therefore be considered as a putative therapeutic target for prevention and treatment of cardiometabolic disorders in schizophrenia.

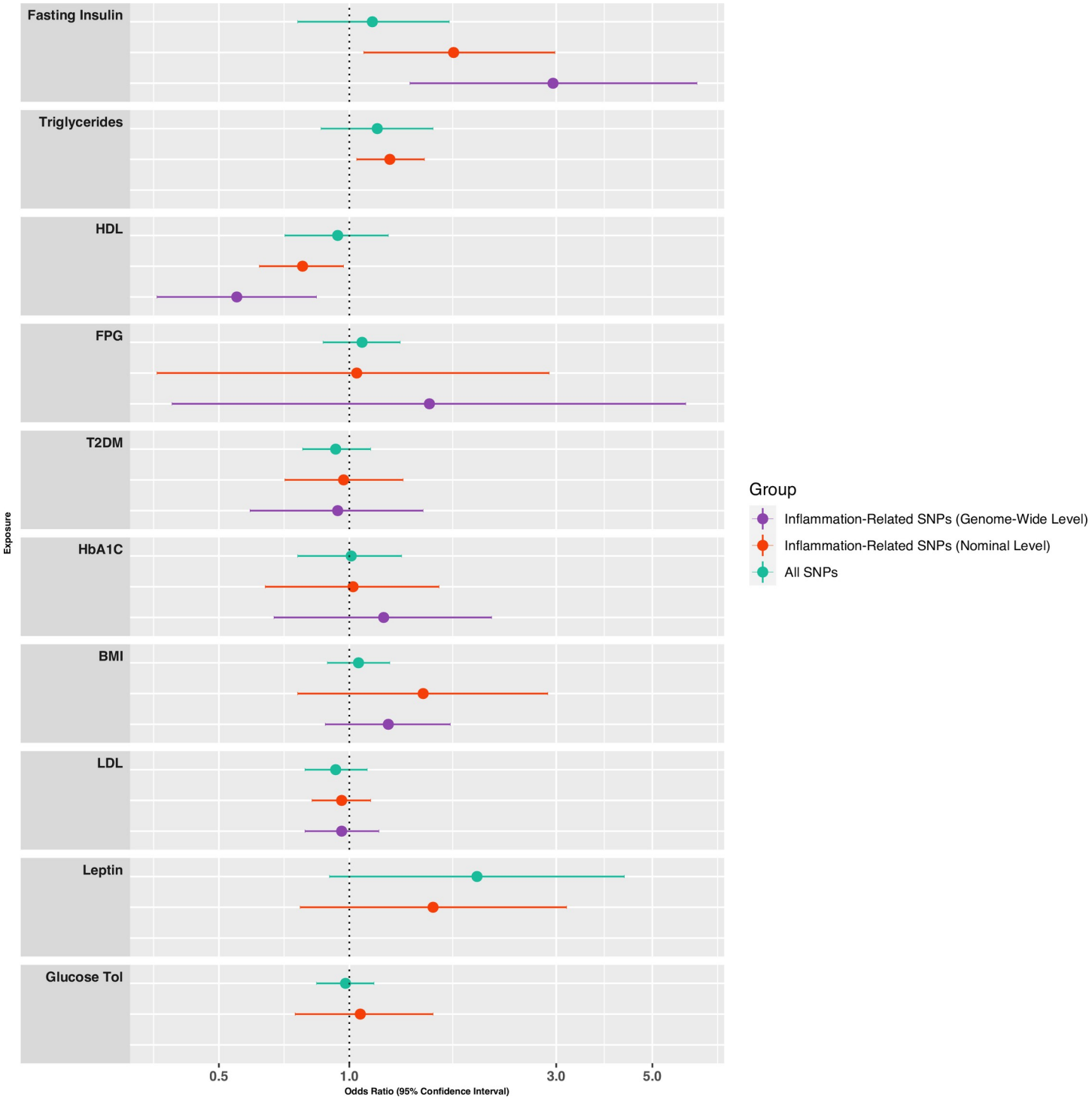


Fig 2. MR analyses testing associations between cardiometabolic traits and schizophrenia. Forest plot presents ORs and 95% CIs for associations between cardiometabolic traits and schizophrenia using IVW / Wald Ratio MR analyses based on all single nucleotide polymorphisms (SNPs) associated with each risk factor (green), inflammation-related SNPs at genome-wide significance (purple), and inflammation-related SNPs at nominal significance (red). See Tables 1 and 2 for the number of SNPs used in each analysis. HDL = High Density Lipoprotein; T2DM = Type 2 Diabetes Mellitus; BMI = Body Mass Index; FPG = Fasting Plasma Glucose; LDL = Low-Density Lipoprotein; HbA1C = Glycated Haemoglobin; Glucose Tol = Glucose Tolerance.

https://doi.org/10.1371/journal.pmed.1003455.g002

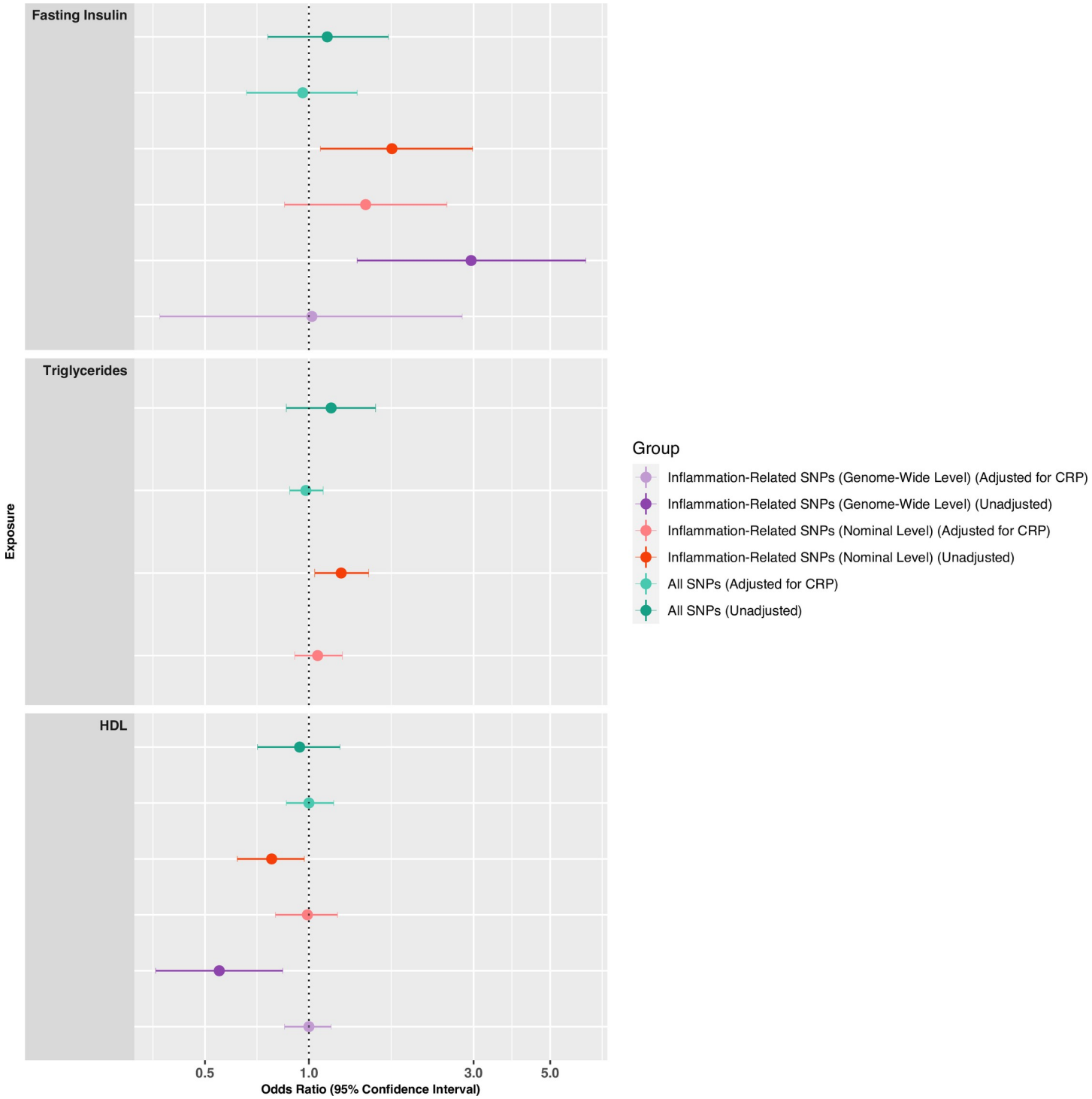


Fig 3. Multivariable MR analysis testing associations between insulin resistance phenotypes and schizophrenia after controlling for genetic associations for CRP. Forest plot presents ORs and 95% CIs for inverse-variance weighted regression (IVW) MR associations between insulin resistance phenotypes and schizophrenia using all single nucleotide polymorphisms (SNPs) (dark green), and after controlling for association of these SNPs with C-reactive protein (CRP) using multivariable MR (MVMR) (light green). The forest plot also presents ORs and 95% CIs for IVW / Wald Ratio MR associations between insulin resistance phenotypes and schizophrenia using inflammation-related SNPs (genome-wide significance = dark purple; nominal significance = red), and after controlling for association of these SNPs with CRP using MVMR (genome-wide significance = light purple; nominal significance = light red). HDL = high-density lipoprotein.

https://doi.org/10.1371/journal.pmed.1003455.g003

We used CRP, an archetypal downstream inflammatory marker, as a means of gauging the effect of systemic inflammation in MVMR analysis, rather than hypothesising a specific role for CRP in the relationship between insulin resistance and schizophrenia. Nevertheless, CRP has observationally shown in both cross-sectional [40] and longitudinal [41] research to be associated with schizophrenia, although such findings are limited by the potential of residual confounding and reverse causality. Interestingly, however, MR findings have reported that genetically predicted CRP may have a protective effect on schizophrenia [21], with authors positing that a genetically attenuated ability to produce CRP may predispose to more insidious and chronic infections. In our MVMR analysis, attenuation of insulin resistance schizophrenia associations after controlling for CRP is consistent with inflammation being associated with both exposure and outcome, albeit "negatively" with the latter. Further research is needed to explore potential mechanisms of association between CRP and schizophrenia.

Many of the SNPs included in the inflammation-related analysis were associated with neutrophils and/or lymphocytes. A raised neutrophil to lymphocyte ratio (NLR) is a marker of systemic inflammation and is known to be associated with schizophrenia [42] and insulin resistance [43]. We were unable to find large GWAS studies conducted in European populations for NLR or for other inflammatory markers, which we might have used in MVMR analyses in place of CRP.

Based on our findings, we are unable to completely rule out the possibility that insulin resistance may mediate an inflammation–schizophrenia association (panel B in S1 Methods), since there was weak evidence that did not survive correction for multiple testing for an association of triglycerides and HDL with schizophrenia using the weighted median method, and in our MR-PRESSO sensitivity analysis, evidence from the outlier-corrected IVW analysis suggested a possible association between triglycerides and schizophrenia. These findings are broadly similar to 1 previous MR study [17], which reported only weak evidence of an association between the homeostasis model assessment (HOMA), a measure of insulin resistance, on schizophrenia. Another MR study [16] reported a genetic association between fasting insulin and schizophrenia, although the evidence attenuated after adjustment for BMI. To account for BMI, we obtained summary statistics for genetic variants related to insulin resistance after controlling for BMI [11]. The previous MR study included an ethnically heterogeneous sample, increasing the potential for population stratification bias. We used genetic data from a more ethnically homogenous GWAS of schizophrenia [24]. Nevertheless, while our results in the all-SNP analysis suggested weak evidence for triglycerides and HDL, which may reflect an insulin resistance phenotype, the evidence did not survive correction for multiple testing and requires replication in future when larger GWAS samples are available.

The implications of our findings with regard to shared causal mechanisms should not distract clinicians from focusing on the assessment and management of malleable lifestyle factors related to cardiometabolic disorders in people with schizophrenia. Factors such as poorer diet, reduced exercise and smoking, which are associated with schizophrenia [7,44,45], may predispose to an inflammatory state [46]. Therefore, it is possible that lifestyle factors exacerbate a feedback loop between inflammation, insulin resistance, and schizophrenia by increasing both inflammation and insulin resistance, eventually leading to T2DM and other cardiometabolic disorders such as obesity and cardiovascular disease (CVD). In addition to the potential therapeutic potential of anti-inflammatory medications, malleable lifestyle factors must continue to remain crucial targets [47,48] for the prevention of cardiometabolic morbidity in people with schizophrenia.

Additional findings

We report that after outlier correction, schizophrenia had a weak protective effect on BMI. This finding complements estimates from previous research [53] using LD score regression,

though we are able to advance previous findings since genetic correlation analyses are unable to test direction of association. This finding suggests that weight gain associated with schizophrenia is unlikely to be a feature of the illness itself but could be attributed to iatrogenic or lifestyle effects. Moreover, the "lean insulin-resistance" phenotype may be associated with higher levels of inflammation [54] and warrants further research in the context of schizophrenia, particularly since in younger patients, the "lean" nature of the phenotype may mean that important cardiometabolic investigations may be overlooked in the clinic.

Strengths and limitations

Strengths of this study include the use of a large set of cardiometabolic traits and large GWAS datasets, through which we were able to test specific biological mechanisms. We chose SNPs reaching genome-wide significance from large GWAS and meta-GWAS for insulin resistance and related cardiometabolic traits. We performed a comprehensive set of sensitivity analyses to check MR assumptions. Furthermore, while weak instrument bias may be a factor in MR analysis, in two-sample MR, this bias tends towards the null [55], thus would not explain the positive associations we describe. We corrected for multiple testing to minimise potential type I error.

Our study has some limitations. We did not select SNPs in known coding regions for the exposures, for example, the *IRS-1* gene for insulin resistance [56]. We took this step on the assumption that many mechanisms at play may not yet be fully understood. For example, while the heritability of cardiometabolic traits such as obesity is as high as 70%, the variance currently explained by known genetic variants is but a small fraction of this [57]. In addition, selecting SNPs from many different GWAS studies featuring large sample sizes may increase the risk of sample overlap between exposure and outcome variables and can bias the results in either direction, depending on the proportion of overlap [27]. Also, for our primary inflammation-related SNP analysis, we chose a stringent *p*-value threshold to define inflammation-related SNPs. In doing so, we may have overlooked some SNPs with true inflammatory associations. As a result, only one genome-wide significant inflammation-related genetic variant was included in the analysis of fasting insulin and HDL, and none were included for triglycerides. Therefore, these results be considered with caution. However, we attempted to address this limitation by relaxing the *p*-value threshold for inflammation-related SNPs, thereby allowing a greater number of SNPs to be included, and the results for fasting insulin, HDL and triglycerides were consistent. Yet, the inclusion of inflammation-related genetic variants at a relaxed significance threshold may have increased the risk of weak instrument bias for those analyses. In the future, larger and better-powered GWAS may identify more SNPs for analysis and at greater resolution, potentially unearthing a larger number of inflammation-related SNPs, which would be helpful to confirm our findings. Additionally, the full range of gene products from the genetic variants we used as proxies for the cardiometabolic traits is unknown, and so we are unable to comment on potential biological mechanisms of association other than inflammation, which may also be relevant. Finally, our analyses were based on data from mostly European participants, so it is unclear whether our results apply to other populations. Large-scale GWAS and replication of our analyses in other populations are required to answer this question.

Conclusions

It is well established that certain antipsychotic drugs and lifestyle factors such as smoking, lack of exercise, and poor diet are important contributors to cardiometabolic comorbidity in people with schizophrenia. In addition to this, our findings suggest that inflammation may be a

common cause for schizophrenia and cardiometabolic disorders, which may at least partly explain why they so commonly co-occur in clinical practice. Lifestyle modification and careful prescription of certain antipsychotic medications remain crucial malleable targets to reduce the significant impact that comorbid cardiometabolic disorders place on the quality and length of life in people with schizophrenia. In addition, our findings suggest that targeting inflammation could be an important therapeutic target for the treatment and prevention of cardiometabolic disorders in people with schizophrenia. Future research should seek to examine the biological mechanisms, which underpin how inflammation can simultaneously increase the risk of both insulin resistance and schizophrenia.

Supporting information

S1 Methods. Directed acyclic graphs outlining potential mechanisms of association between inflammation, insulin resistance, and schizophrenia.
(DOCX)

S2 Methods. GWAS used for SNP selection.
(DOCX)

S3 Methods. SNPs used as instruments for fasting insulin, triglycerides, and high-density lipoprotein.
(DOCX)

S4 Methods. SNPs used as instruments for fasting plasma glucose.
(DOCX)

S5 Methods. SNPs used as instruments for type 2 diabetes mellitus.
(DOCX)

S6 Methods. SNPs used as instruments for body mass index.
(DOCX)

S7 Methods. SNPs used as instruments for glucose tolerance.
(DOCX)

S8 Methods. SNPs used as instruments for low density lipoprotein.
(DOCX)

S9 Methods. SNPs used as instruments for glycated haemoglobin.
(DOCX)

S10 Methods. SNPs used as instruments for leptin.
(DOCX)

S11 Methods. MR analysis methods.
(DOCX)

S12 Methods. Inflammation-related SNPs for fasting insulin, triglycerides, and high-density lipoprotein.
(DOCX)

S13 Methods. Inflammation-related SNPs for low density lipoprotein.
(DOCX)

S14 Methods. Inflammation-related SNPs for fasting plasma glucose.
(DOCX)

S15 Methods. Inflammation-related SNPs for glycated haemoglobin.
(DOCX)

S16 Methods. Inflammation-related SNPs for type 2 diabetes mellitus.
(DOCX)

S17 Methods. Inflammation-related SNPs for body mass index.
(DOCX)

S18 Methods. Inflammation-related SNPs for schizophrenia.
(DOCX)

S19 Methods. SNPs used for CRP in MVMR analysis.
(DOCX)

S1 Results. Multivariable MR (MVMR) results for insulin resistance phenotype exposures (all-SNP analysis) with addition of CRP as exposure.
(DOCX)

S2 Results. Multivariable MR (MVMR) results for insulin resistance phenotype exposures (inflammation-related-SNP analysis) with addition of CRP as exposure.
(DOCX)

S3 Results. MR analyses using all SNPs for schizophrenia and cardiometabolic outcomes.
(DOCX)

S4 Results. The association between inflammation-related schizophrenia SNPs and cardiometabolic outcomes.
(DOCX)

S5 Results. Cochran Q tests for heterogeneity and MR–Egger intercept tests for horizontal pleiotropy for the association between all cardiometabolic SNPs and schizophrenia.
(DOCX)

S6 Results. Cochran Q tests for heterogeneity and MR–Egger intercept tests for horizontal pleiotropy for the association between inflammation-related cardiometabolic SNPs and schizophrenia.
(DOCX)

S7 Results. Cochran Q tests for heterogeneity and MR–Egger intercept tests for horizontal pleiotropy for the association between schizophrenia SNPs and cardiometabolic outcomes.
(DOCX)

S8 Results. Cochran Q tests for heterogeneity and MR–Egger intercept tests for horizontal pleiotropy for the association between inflammation-related schizophrenia SNPs and cardiometabolic outcomes.
(DOCX)

S9 Results. MR-PRESSO tests of cardiometabolic all-SNP analysis to examine for and correct horizontal pleiotropy.
(DOCX)

S10 Results. MR-PRESSO tests of inflammation-related cardiometabolic SNPs to examine for and correct horizontal pleiotropy.
(DOCX)

S11 Results. MR-PRESSO tests of schizophrenia all-SNP analysis to examine for and correct horizontal pleiotropy.
(DOCX)

S12 Results. MR-PRESSO tests of inflammation-related schizophrenia SNP analysis to examine for and correct horizontal pleiotropy.
(DOCX)

S13 Results. I_{2GX} statistics to examine for potential violation of the "No Measurement Error" (NOME) assumption for MR–Egger analyses.
(DOCX)

S1 Checklist. STROBE-MR: Guidelines for strengthening the reporting of mendelian randomization studies.
(DOCX)

S2 Checklist. STROBE: Guidelines for reporting observational studies.
(DOCX)

S1 Fig. Forest plot illustrating MR analyses of schizophrenia as outcome using all SNPs (green) and inflammation-related SNPs (purple).
(DOCX)

Acknowledgments

The authors wish to thank Dr Isobel Stewart (University of Cambridge) for her methodological advice and support.

Author Contributions

Conceptualization: Benjamin I. Perry, Stephen Burgess, Rachel Upthegrove, Claudia Langenberg, Nicholas J. Wareham, Peter B. Jones, Golam M. Khandaker.

Data curation: Benjamin I. Perry.

Formal analysis: Benjamin I. Perry, Hannah J. Jones, Amy M. Mason.

Funding acquisition: Benjamin I. Perry.

Investigation: Benjamin I. Perry, Stan Zammit, Rachel Upthegrove, Amy M. Mason.

Methodology: Benjamin I. Perry, Stephen Burgess, Hannah J. Jones, Stan Zammit, Rachel Upthegrove, Amy M. Mason, Felix R. Day, Claudia Langenberg, Nicholas J. Wareham, Peter B. Jones, Golam M. Khandaker.

Resources: Golam M. Khandaker.

Supervision: Stephen Burgess, Hannah J. Jones, Stan Zammit, Rachel Upthegrove, Felix R. Day, Nicholas J. Wareham, Peter B. Jones, Golam M. Khandaker.

Visualization: Benjamin I. Perry.

Writing – original draft: Benjamin I. Perry.

Writing – review & editing: Benjamin I. Perry, Stephen Burgess, Hannah J. Jones, Stan Zammit, Rachel Upthegrove, Amy M. Mason, Felix R. Day, Claudia Langenberg, Nicholas J. Wareham, Peter B. Jones, Golam M. Khandaker.

References

1. Owen MJ, Sawa A, Mortensen PB. Schizophrenia. Lancet. 2016; 388 (10039):86–97. https://doi.org/10.1016/S0140-6736(15)01121-6 PMID: 26777917
2. Bhugra D. The global prevalence of schizophrenia. PLoS Med. 2005; 2(5):e151; quiz e75. https://doi.org/10.1371/journal.pmed.0020151 PMID: 15916460
3. Charlson FJ, Ferrari AJ, Santomauro DF, Diminic S, Stockings E, Scott JG, et al. Global Epidemiology and Burden of Schizophrenia: Findings From the Global Burden of Disease Study 2016. Schizophr Bull. 2018; 44 (6):1195–1203. https://doi.org/10.1093/schbul/sby058 PMID: 29762765
4. Mitchell AJ, Vancampfort D, Sweers K, van Winkel R, Yu W, De Hert M. Prevalence of metabolic syndrome and metabolic abnormalities in schizophrenia and related disorders—a systematic review and meta-analysis. Schizophr Bull. 2013; 39 (2):306–318. https://doi.org/10.1093/schbul/sbr148 PMID: 22207632
5. Saha S, Chant D, McGrath J. A systematic review of mortality in schizophrenia: is the differential mortality gap worsening over time? Arch Gen Psychiatry. 2007; 64 (10):1123–1131. https://doi.org/10.1001/archpsyc.64.10.1123 PMID: 17909124
6. Leucht S, Cipriani A, Spineli L, Mavridis D, Orey D, Richter F, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. Lancet. 2013; 382 (9896):951–962. https://doi.org/10.1016/S0140-6736(13)60733-3 PMID: 23810019
7. Heald A, Pendlebury J, Anderson S, Narayan V, Guy M, Gibson M, et al. Lifestyle factors and the metabolic syndrome in Schizophrenia: a cross-sectional study. Ann General Psychiatry. 2017; 16:12. https://doi.org/10.1186/s12991-017-0134-6 PMID: 28289436
8. Reinikainen J, Laatikainen T, Karvanen J, Tolonen H. Lifetime cumulative risk factors predict cardiovascular disease mortality in a 50-year follow-up study in Finland. Int J Epidemiol. 2015; 44 (1):108–116. https://doi.org/10.1093/ije/dyu235 PMID: 25501686
9. Iwani NA, Jalaludin MY, Zin RM, Fuziah MZ, Hong JY, Abqariyah Y, et al. Triglyceride to HDL-C Ratio is Associated with Insulin Resistance in Overweight and Obese Children. Sci Rep. 2017; 7:40055. https://doi.org/10.1038/srep40055 PMID: 28059134
10. Glueck CJ, Khan NA, Umar M, Uppal MS, Ahmed W, Morrison JA, et al. Insulin resistance and triglycerides. J Investig Med. 2009; 57 (8):874–881. https://doi.org/10.2310/JIM.0b013e3181bca9d2 PMID: 19809367
11. Lotta LA, Gulati P, Day FR, Payne F, Ongen H, van de Bunt M, et al. Integrative genomic analysis implicates limited peripheral adipose storage capacity in the pathogenesis of human insulin resistance. Nat Genet. 2017; 49 (1):17–26. https://doi.org/10.1038/ng.3714 PMID: 27841877
12. Perry BI, McIntosh G, Weich S, Singh S, Rees K. The association between first-episode psychosis and abnormal glycaemic control: systematic review and meta-analysis. Lancet Psychiatry. 2016; 3 (11):1049–1058. https://doi.org/10.1016/S2215-0366(16)30262-0 PMID: 27720402
13. Pillinger T, Beck K, Stubbs B, Howes OD. Cholesterol and triglyceride levels in first-episode psychosis: systematic review and meta-analysis. Br J Psychiatry. 2017; 211 (6):339–349. https://doi.org/10.1192/bjp.bp.117.200907 PMID: 28982658
14. Perry BI, Upthegrove R, Thompson A, Marwaha S, Zammit S, Singh SP, et al. Dysglycaemia, Inflammation and Psychosis: Findings From the UK ALSPAC Birth Cohort. Schizophr Bull. 2018.
15. Smith GD. Mendelian Randomization for Strengthening Causal Inference in Observational Studies: Application to Gene x Environment Interactions. Perspect Psychol Sci. 2010; 5 (5):527–545. https://doi.org/10.1177/1745691610383505 PMID: 26162196
16. Li Z, Chen P, Chen J, Xu Y, Wang Q, Li X, et al. Glucose and Insulin-Related Traits, Type 2 Diabetes and Risk of Schizophrenia: A Mendelian Randomization Study. EBioMedicine. 2018; 34:182–188. https://doi.org/10.1016/j.ebiom.2018.07.037 PMID: 30100396
17. Polimanti R, Gelernter J, Stein DJ. Genetically determined schizophrenia is not associated with impaired glucose homeostasis. Schizophr Res. 2017.
18. Dandona P, Aljada A, Bandyopadhyay A. Inflammation: the link between insulin resistance, obesity and diabetes. Trends Immunol. 2004; 25 (1):4–7. https://doi.org/10.1016/j.it.2003.10.013 PMID: 14698276
19. Khandaker GM, Pearson RM, Zammit S, Lewis G, Jones PB. Association of serum interleukin 6 and C-reactive protein in childhood with depression and psychosis in young adult life: a population-based longitudinal study. JAMA Psychiat. 2014; 71 (10):1121–1128. https://doi.org/10.1001/jamapsychiatry.2014.1332 PMID: 25133871
20. Upthegrove R, Manzanares-Teson N, Barnes NM. Cytokine function in medication-naive first episode psychosis: a systematic review and meta-analysis. Schizophr Res. 2014; 155 (1–3):101–108. https://doi.org/10.1016/j.schres.2014.03.005 PMID: 24704219

21. Hartwig FP, Borges MC, Horta BL, Bowden J, Davey SG. Inflammatory Biomarkers and Risk of Schizophrenia: A 2-Sample Mendelian Randomization Study. JAMA Psychiat. 2017; 74 (12):1226–1233. https://doi.org/10.1001/jamapsychiatry.2017.3191 PMID: 29094161

22. Khandaker GM, Zammit S, Burgess S, Lewis G, Jones PB. Association between a functional interleukin 6 receptor genetic variant and risk of depression and psychosis in a population-based birth cohort. Brain Behav Immun. 2017.

23. Kim JH, Bachmann RA, Chen J. Interleukin-6 and insulin resistance. Vitam Horm. 2009; 80:613–633. https://doi.org/10.1016/S0083-6729(08)00621-3 PMID: 19251052

24. Pardinas AF, Holmans P, Pocklington AJ, Escott-Price V, Ripke S, Carrera N, et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nat Genet. 2018; 50 (3):381–389. https://doi.org/10.1038/s41588-018-0059-2 PMID: 29483656

25. Shi H, Mancuso N, Spendlove S, Pasaniuc B. Local Genetic Correlation Gives Insights into the Shared Genetic Architecture of Complex Traits. Am J Hum Genet. 2017; 101 (5):737–751. https://doi.org/10.1016/j.ajhg.2017.09.022 PMID: 29100087

26. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics. 2015; 31 (21):3555–3557. https://doi.org/10.1093/bioinformatics/btv402 PMID: 26139635

27. Hemani G, Bowden J, Davey SG. Evaluating the potential role of pleiotropy in Mendelian randomization studies. Hum Mol Genet. 2018; 27 (R2):R195–R208. https://doi.org/10.1093/hmg/ddy163 PMID: 29771313

28. R Core Team. R: A Language and Environment for Statistical Computing.2017. Available from: https://www.R-project.org/.

29. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet. 2018; 50 (5):693–698. https://doi.org/10.1038/s41588-018-0099-7 PMID: 29686387

30. Bowden J, Del Greco MF, Minelli C, Davey Smith G, Sheehan NA, Thompson JR. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I2 statistic. Int J Epidemiol. 2016; 45 (6):1961–1974. https://doi.org/10.1093/ije/dyw220 PMID: 27616674

31. Burgess S, David Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, et al. Guidelines for performing Mendelian randomization investigations [version 2; peer review: 2 approved]. Wellcome Open Res. 2020; 4(186). https://doi.org/10.12688/wellcomeopenres.15555.2 PMID: 32760811

32. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med. 2007; 4 (10):e296. https://doi.org/10.1371/journal.pmed.0040296 PMID: 17941714

33. Staley JR, Blackshaw J, Kamat MA, Ellis S, Surendran P, Sun BB, et al. PhenoScanner: a database of human genotype-phenotype associations. Bioinformatics. 2016; 32 (20):3207–3209. https://doi.org/10.1093/bioinformatics/btw373 PMID: 27318201

34. Burgess S, Thompson SG. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. Am J Epidemiol. 2015; 181 (4):251–260. https://doi.org/10.1093/aje/kwu283 PMID: 25632051

35. Sanderson E, Davey Smith G, Windmeijer F, Bowden J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. Int J Epidemiol. 2018.

36. Ligthart S, Vaez A, Võsa U, Stathopoulou MG, de Vries PS, Prins BP, et al. Genome Analyses of >200,000 Individuals Identify 58 Loci for Chronic Inflammation and Highlight Pathways that Link Inflammation and Complex Disorders. Am J Hum Genet. 2018; 103 (5):691–706. https://doi.org/10.1016/j.ajhg.2018.09.009 PMID: 30388399

37. Holm S. A Simple Sequentially Rejective Multiple Test Procedure. 1979; 6 (2):65–70.

38. Pillinger T, Beck K, Gobjila C, Donocik JG, Jauhar S, Howes OD. Impaired Glucose Homeostasis in First-Episode Schizophrenia: A Systematic Review and Meta-analysis. JAMA Psychiat. 2017; 74 (3):261–269. https://doi.org/10.1001/jamapsychiatry.2016.3803 PMID: 28097367

39. Cakici N, van Beveren NJM, Judge-Hundal G, Koola MM, Sommer IEC. An update on the efficacy of anti-inflammatory agents for patients with schizophrenia: a meta-analysis. Psychol Med. 2019:1–13. https://doi.org/10.1017/S0033291719001995 PMID: 31439071

40. Fernandes BS, Steiner J, Bernstein HG, Dodd S, Pasco JA, Dean OM, et al. C-reactive protein is increased in schizophrenia but is not altered by antipsychotics: meta-analysis and implications. Mol Psychiatry. 2016; 21 (4):554–564. https://doi.org/10.1038/mp.2015.87 PMID: 26169974

41. Metcalf SA, Jones PB, Nordstrom T, Timonen M, Maki P, Miettunen J, et al. Serum C-reactive protein in adolescence and risk of schizophrenia in adulthood: A prospective birth cohort study. Brain Behav Immun. 2017; 59:253–259. https://doi.org/10.1016/j.bbi.2016.09.008 PMID: 27622678

42. Karageorgiou V, Milas GP, Michopoulos I. Neutrophil-to-lymphocyte ratio in schizophrenia: A systematic review and meta-analysis. Schizophr Res. 2018. https://doi.org/10.1016/j.schres.2018.12.017 PMID: 30573407

43. Lou M, Luo P, Tang R, Peng Y, Yu S, Huang W, et al. Relationship between neutrophil-lymphocyte ratio and insulin resistance in newly diagnosed type 2 diabetes mellitus patients. BMC Endocr Disord. 2015; 15:9. https://doi.org/10.1186/s12902-015-0002-9 PMID: 25887236

44. McCreadie RG,. smoking and cardiovascular risk in people with schizophrenia: descriptive study. Br J Psychiatry. 2003; 183:534–539. https://doi.org/10.1192/bjp.183.6.534 PMID: 14645025

45. Bly MJ, Taylor SF, Dalack G, Pop-Busui R, Burghardt KJ, Evans SJ, et al. Metabolic syndrome in bipolar disorder and schizophrenia: dietary and lifestyle factors compared to the general population. Bipolar Disord. 2014; 16 (3):277–288. https://doi.org/10.1111/bdi.12160 PMID: 24330321

46. Jarvandi S, Davidson NO, Jeffe DB, Schootman M. Influence of lifestyle factors on inflammation in men and women with type 2 diabetes: results from the National Health and Nutrition Examination Survey, 1999–2004. Ann Behav Med. 2012; 44 (3):399–407. https://doi.org/10.1007/s12160-012-9397-y PMID: 22865468

47. Teasdale SB, Curtis J, Ward PB, Watkins A, Lederman O, Rosenbaum S, et al. The effectiveness of the Keeping the Body in Mind Xtend pilot lifestyle program on dietary intake in first-episode psychosis: Two-year outcomes. Obes Res Clin Pract. 2019; 13 (2):214–216. https://doi.org/10.1016/j.orcp.2019.02.003 PMID: 30826255

48. Ward PB, Firth J, Rosenbaum S, Samaras K, Stubbs B, Curtis J. Lifestyle interventions to reduce premature mortality in schizophrenia. Lancet Psychiatry. 2017; 4 (7):e14. https://doi.org/10.1016/S2215-0366(17)30235-3 PMID: 28652047

49. Lord GM. Leptin as a proinflammatory cytokine. Contrib Nephrol. 2006; 151:151–164. https://doi.org/10.1159/000095326 PMID: 16929139

50. Bulló M, García-Lorda P, Megias I, Salas-Salvadó J. Systemic inflammation, adipose tissue tumor necrosis factor, and leptin expression. Obes Res. 2003; 11 (4):525–531. https://doi.org/10.1038/oby.2003.74 PMID: 12690081

51. Stubbs B, Wang AK, Vancampfort D, Miller BJ. Are leptin levels increased among people with schizophrenia versus controls? A systematic review and comparative meta-analysis. Psychoneuroendocrinology. 2016; 63:144–154. https://doi.org/10.1016/j.psyneuen.2015.09.026 PMID: 26444588

52. Neelamekam S, Nurjono M, Lee J. Regulation of interleukin-6 and leptin in schizophrenia patients: a preliminary analysis. Clin Psychopharmacol Neurosci. 2014; 12 (3):209–214. https://doi.org/10.9758/cpn.2014.12.3.209 PMID: 25598824

53. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh PR, et al. An atlas of genetic correlations across human diseases and traits. Nat Genet. 2015; 47 (11):1236–1241. https://doi.org/10.1038/ng.3406 PMID: 26414676

54. Ding C, Chan Z, Magkos F. Lean, but not healthy: the 'metabolically obese, normal-weight' phenotype. Curr Opin Clin Nutr Metab Care. 2016; 19 (6):408–417. https://doi.org/10.1097/MCO.0000000000000317 PMID: 27552473

55. Davies NM. von Hinke Kessler Scholder S, Farbmacher H, Burgess S, Windmeijer F, Smith GD. The many weak instruments problem and Mendelian randomization. Stat Med. 2015; 34 (3):454–468. https://doi.org/10.1002/sim.6358 PMID: 25382280

56. Carvalho E, Jansson PA, Axelsen M, Eriksson JW, Huang X, Groop L, et al. Low cellular IRS 1 gene and protein expression predict insulin resistance and NIDDM. FASEB J. 1999; 13 (15):2173–2178. https://doi.org/10.1096/fasebj.13.15.2173 PMID: 10593864

57. Herrera BM, Keildson S, Lindgren CM. Genetics and epigenetics of obesity. Maturitas. 2011; 69 (1):41–49. https://doi.org/10.1016/j.maturitas.2011.02.018 PMID: 21466928

Appendix D

Appendix D Methods

Predictors Included in Exploratory Validation Analysis of QRISK3, QDiabetes and PRIMROSE

Age

The actual age of participants at the time of attendance at the F17 clinic was between 17.21 and 18.67 years. I chose to uniformly consider age as 18 years for all participants at the time of the exposure. I did this under consideration that the relatively small variation of age in the sample may have been disproportionately amplified by the tested algorithms.

Ethnicity

Ethnicity was recorded from questionnaire data (Missing; White; Black Caribbean; Black African; Other Black; Indian; Pakistani; Bangladeshi; Chinese; Other; Don't Know) and recoded to match QDiabetes/QRISK3 defined categories. Ethnic categories in QRISK3 and QDiabetes differed in also including 'Other Asian', which I was unable to recode.

Townsend Scores

Townsend Scores were calculated by ALSPAC and obtained in quintiles, based upon self-report data.

Family History of Cardiometabolic Disorders

A positive family history of cardiometabolic disorders obtained from self-report questionnaire data encompassing hypertension, T2DM, hypercholesterolaemia, or cardiovascular diseases.

Smoking

Smoking status was recoded from self-report questionnaire data at age 18 years to match the categories denoted by QRISK3, QDdiabetes and PRIMROSE.

Body Mass Index (BMI)

BMI was calculated during clinic assessment from measurements of height (m) and weight (kg) at age 18 years.

Systolic Blood Pressure

Systolic blood pressure was obtained during clinic assessments in both arms, twice. I included the second measure of systolic blood pressure recorded in the left arm, unless only the right arm, or only one measurement was available.

Blood-Based Predictors

For blood-based predictors (FPG, HDL and triglycerides), fasting samples were taken at 0900 after a 10-hour fast (water only) at age 18 years. Samples were immediately spun, frozen and stored at –80^{o}C and measurements were assayed within 3 to 9 months of the samples being taken with no previous freeze-thaw cycles. FPG was measured by an ultrasensitive ELISA (Mercodia, Uppsala, Sweden) automated microparticle enzyme immunoassay. Its sensitivity was 0.07 mU/L, and inter- and intra-assay coefficients of variation were <6%. Plasma lipid concentrations were measured by modification of the standard Lipid Research Clinics Protocol by using enzymatic reagents for lipid determination.

Depression

Depression was measured using the CIS-R and coded by ALSPAC to meet ICD-10 criteria. I considered depression to be present for participants meeting ICD-10 criteria for mild, moderate or severe depression.

Psychosis

Since hospital record linkage was not available in the dataset, for a measure of clinical psychosis, we used the presence of psychotic disorder at age 18y.

Medication

Medication data was coded by the authors from self-report free-text data collected from a 'white space' box in a questionnaire. There was wide variation in spellings of medications in this data. Participants were coded as "1" for any particular medication if spellings either matched the spellings in the British National Formulary, or if the medication could be easily deciphered from the free-text. In the event of ambiguity, participants were coded "0". I was unable to include prescription of statins since ALSPAC guidelines state variables can only be coded if cell counts are greater than 5, to reduce the risk of participant identification. Less than 5 participants in the cohort self-reported a prescription for statins.

Unavailable Predictors

Due to the lack of hospital record linkage, I could not include personal history of either cardiovascular disease, gestational diabetes, systemic lupus erythematosus, learning disabilities, migraines, erectile dysfunction, atrial fibrillation, chronic kidney disease, or polycystic ovarian syndrome.

Appendix D Results

Reference List of Studies Included In Systematic Review Presented in Chapter 6

Abd El-Wahab, E. W., Shatat, H. Z., & Charl, F. (2019). Adapting a Prediction Rule for Metabolic Syndrome Risk Assessment Suitable for Developing Countries. J Prim Care Community Health, 10, 2150132719882760. doi:10.1177/2150132719882760

Abdul-Ghani, M. A., Abdul-Ghani, T., Stern, M. P., Karavic, J., Tuomi, T., Bo, I., . . . Groop, L. (2011). Two-step approach for the prediction of future type 2 diabetes risk. Diabetes Care, 34(9), 2108-2112. doi:10.2337/dc10-2201

Addoh, O., Edwards, M. K., & Loprinzi, P. D. (2016). Predictive Validity of a Medical-Related Cardiorespiratory Fitness Algorithm in Predicting Cardiovascular Disease- and All-Cause Mortality: Implications for Integration Into Clinical Practice. Mayo Clin Proc, 91(9), 1320-1321. doi:10.1016/j.mayocp.2016.06.008

Aekplakorn, W., Bunnag, P., Woodward, M., Sritara, P., Cheepudomwit, S., Yamwong, S., . . . Rajatanavin, R. (2006). A risk score for predicting incident diabetes in the Thai population. Diabetes Care, 29(8), 1872-1877. doi:10.2337/dc05-2141

Alaa, A. M., Bolton, T., Di Angelantonio, E., Rudd, J. H. F., & van der Schaar, M. (2019). Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. PLoS One, 14(5), e0213653. doi:10.1371/journal.pone.0213653

Alghwiri, A., Alghadir, A., Awad, H. (2014). The Arab Risk (ARABRISK): Translation and Validation. Biomedical Research, 25(2), 271-275.

Alssema, M., Newson, R. S., Bakker, S. J., Stehouwer, C. D., Heymans, M. W., Nijpels, G., . . . Dekker, J. M. (2012). One risk assessment tool for cardiovascular disease, type 2 diabetes, and chronic kidney disease. Diabetes Care, 35(4), 741-748. doi:10.2337/dc11-1417

Alssema, M., Vistisen, D., Heymans, M. W., Nijpels, G., Glumer, C., Zimmet, P. Z., . . . collaboration, D.-. (2011). The Evaluation of Screening and Early Detection Strategies for Type 2 Diabetes and Impaired Glucose Tolerance (DETECT-2) update of the Finnish diabetes risk score for prediction of incident type 2 diabetes. Diabetologia, 54(5), 1004-1012. doi:10.1007/s00125-010-1990-7

Anderson, J. P., Parikh, J. R., Shenfeld, D. K., Ivanov, V., Marks, C., Church, B. W., . . . Rublee, D. A. (2015). Reverse Engineering and Evaluation of Prediction Models for Progression to Type 2 Diabetes: An Application of Machine Learning Using Electronic Health Records. J Diabetes Sci Technol, 10(1), 6-18. doi:10.1177/1932296815620200

Arima, H., Yonemoto, K., Doi, Y., Ninomiya, T., Hata, J., Tanizaki, Y., . . . Kiyohara, Y. (2009). Development and validation of a cardiovascular risk prediction model for Japanese: the Hisayama study. Hypertens Res, 32(12), 1119-1122. doi:10.1038/hr.2009.161

Artero, E. G., Jackson, A. S., Sui, X., Lee, D. C., O'Connor, D. P., Lavie, C. J., . . . Blair, S. N. (2014). Longitudinal algorithms to estimate cardiorespiratory fitness: associations with nonfatal cardiovascular disease and disease-specific mortality. J Am Coll Cardiol, 63(21), 2289-2296. doi:10.1016/j.jacc.2014.03.008

Artigao-Rodenas, L. M., Carbayo-Herencia, J. A., Divison-Garrote, J. A., Gil-Guillen, V. F., Masso-Orozco, J., Simarro-Rueda, M., . . . Grupo de Enfermedades Vasculares de, A. (2013). Framingham risk score for prediction of cardiovascular diseases: a population-based study from southern Europe. PLoS One, 8(9), e73529. doi:10.1371/journal.pone.0073529

Aslibekyan, S., Campos, H., Loucks, E. B., Linkletter, C. D., Ordovas, J. M., & Baylin, A. (2011). Development of a cardiovascular risk score for use in low- and middle-income countries. J Nutr, 141(7), 1375-1380. doi:10.3945/jn.110.133140

Assmann, G., Cullen, P., & Schulte, H. (2002). Simple scoring scheme for calculating the risk of acute coronary events based on the 10-year follow-up of the prospective cardiovascular Munster (PROCAM) study. Circulation, 105(3), 310-315. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11804985

Ayala Solares, J. R., Canoy, D., Raimondi, F. E. D., Zhu, Y., Hassaine, A., Salimi-Khorshidi, G., . . . Rahimi, K. (2019). Long-Term Exposure to Elevated Systolic Blood Pressure in Predicting Incident Cardiovascular Disease: Evidence From Large-Scale Routine Electronic Health Records. J Am Heart Assoc, 8(12), e012129. doi:10.1161/JAHA.119.012129

Backholer, K., Hirakawa, Y., Tonkin, A., Giles, G., Magliano, D. J., Colagiuri, S., . . . Woodward, M. (2017). Development of an Australian cardiovascular disease mortality risk score using multiple imputation and recalibration from national statistics. BMC Cardiovasc Disord, 17(1), 17. doi:10.1186/s12872-016-0462-5

Balkau, B., Hu, G., Qiao, Q., Tuomilehto, J., Borch-Johnsen, K., Pyörälä, K., . . . Group, E. D. E. (2004). Prediction of the risk of cardiovascular mortality using a score that includes glucose as a risk factor. The DECODE Study. Diabetologia, 47(12), 2118-2128. doi:10.1007/s00125-004-1574-5

Barazzoni, R., Gortan Cappellari, G., Semolic, A., Ius, M., Zanetti, M., Gabrielli, A., . . . Simon, G. (2019). Central adiposity markers, plasma lipid profile and cardiometabolic risk prediction in overweight-obese individuals. Clin Nutr, 38(3), 1171-1179. doi:10.1016/j.clnu.2018.04.014

Bell, K., Hayen, A., McGeechan, K., Neal, B., & Irwig, L. (2012). Effects of additional blood pressure and lipid measurements on the prediction of cardiovascular risk. Eur J Prev Cardiol, 19(6), 1474-1485. doi:10.1177/1741826711424494

Boland, B., De Muylder, R., Goderis, G., Degryse, J., Gueuning, Y., Paulus, D., & Jeanjean, M. (2004). Cardiovascular prevention in general practice: development and validation of an algorithm. Acta Cardiol, 59(6), 598-605. doi:10.2143/AC.59.6.2005241

Boucher, H. (2019). Exploring the Utility of MUAC in Classifying Adult Metabolic Syndrome Using NHANES 2015-16. William Honors College, Honors Research Projects, University of Akron.

Brand, R. J., Rosenman, R.H., Sholtz, R.I., Friedman, M. (1976). Multivariate Prediction of Coronary Heart Disease in the Western Collaborative Group Study Compared to the Findings of the Framingham Study. Circulation, 53(2), 348-355.

Brautbar, A., Ballantyne, C. M., Lawson, K., Nambi, V., Chambless, L., Folsom, A. R., . . . Boerwinkle, E. (2009). Impact of adding a single allele in the 9p21 locus to traditional risk factors on reclassification of coronary heart disease risk and implications for lipid-modifying therapy in the Atherosclerosis Risk in Communities study. Circ Cardiovasc Genet, 2(3), 279-285. doi:10.1161/CIRCGENETICS.108.817338

Chambless, L. E., Folsom, A. R., Sharrett, A. R., Sorlie, P., Couper, D., Szklo, M., & Nieto, F. J. (2003). Coronary heart disease risk prediction in the Atherosclerosis Risk in Communities (ARIC) study. J Clin Epidemiol, 56(9), 880-890. doi:10.1016/s0895-4356(03)00055-6

Chen, L., Magliano, D. J., Balkau, B., Colagiuri, S., Zimmet, P. Z., Tonkin, A. M., . . . Shaw, J. E. (2010). AUSDRISK: an Australian Type 2 Diabetes Risk Assessment Tool based on demographic, lifestyle and simple anthropometric measures. Med J Aust, 192(4), 197-202. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/20170456

Chen, L., Tonkin, A. M., Moon, L., Mitchell, P., Dobson, A., Giles, G., . . . De Bacquer, D. (2009). Recalibration and validation of the SCORE risk chart in the Australian population: the AusSCORE chart. Eur J Cardiovasc Prev Rehabil, 16(5), 562-570. doi:10.1097/HJR.0b013e32832cd9cb

Chen, X., Wu, Z., Chen, Y., Wang, X., Zhu, J., Wang, N., . . . Fu, C. (2017). Risk score model of type 2 diabetes prediction for rural Chinese adults: the Rural Deqing Cohort Study. J Endocrinol Invest, 40(10), 1115-1123. doi:10.1007/s40618-017-0680-4

Chien, K., Cai, T., Hsu, H., Su, T., Chang, W., Chen, M., . . . Hu, F. B. (2009). A prediction model for type 2 diabetes risk among Chinese people. Diabetologia, 52(3), 443-450. doi:10.1007/s00125-008-1232-4

Chien, K. L., Hsu, H. C., Su, T. C., Chang, W. T., Chen, P. C., Sung, F. C., . . . Lee, Y. T. (2012). Constructing a point-based prediction model for the risk of coronary artery disease in a Chinese community: a report from a cohort study in Taiwan. Int J Cardiol, 157(2), 263-268. doi:10.1016/j.ijcard.2012.03.017

Choe, E. K., Rhee, H., Lee, S., Shin, E., Oh, S. W., Lee, J. E., & Choi, S. H. (2018). Metabolic Syndrome Prediction Using Machine Learning Models with Genetic and Clinical Information from a Nonobese Healthy Population. Genomics Inform, 16(4), e31. doi:10.5808/GI.2018.16.4.e31

Cross, D. S., McCarty, C. A., Hytopoulos, E., Beggs, M., Nolan, N., Harrington, D. S., . . . Quertermous, T. (2012). Coronary risk assessment among intermediate risk patients using a clinical and biomarker based algorithm developed and validated in two population cohorts. Curr Med Res Opin, 28(11), 1819-1830. doi:10.1185/03007995.2012.742878

D'Agostino, R. B., Sr., Grundy, S., Sullivan, L. M., Wilson, P., & Group, C. H. D. R. P. (2001). Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. JAMA, 286(2), 180-187. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11448281

D'Agostino, R. B., Sr., Vasan, R. S., Pencina, M. J., Wolf, P. A., Cobain, M., Massaro, J. M., & Kannel, W. B. (2008). General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation, 117(6), 743-753. doi:10.1161/CIRCULATIONAHA.107.699579

Davies, R. W., Dandona, S., Stewart, A. F., Chen, L., Ellis, S. G., Tang, W. H., . . . Wells, G. A. (2010). Improved prediction of cardiovascular disease based on a panel of single nucleotide polymorphisms identified through genome-wide association studies. Circ Cardiovasc Genet, 3(5), 468-474. doi:10.1161/CIRCGENETICS.110.946269

De Bacquer, D., & De Backer, G. (2010). Predictive ability of the SCORE Belgium risk chart for cardiovascular mortality. Int J Cardiol, 143(3), 385-390. doi:10.1016/j.ijcard.2009.03.101

Dimopoulos, A. C., Nikolaidou, M., Caballero, F. F., Engchuan, W., Sanchez-Niubo, A., Arndt, H., . . . Panagiotakos, D. B. (2018). Machine learning methodologies versus cardiovascular risk scores, in predicting disease risk. BMC Med Res Methodol, 18(1), 179. doi:10.1186/s12874-018-0644-1

Dugee, O., Janchiv, O., Jousilahti, P., Sakhiya, A., Palam, E., Nuorti, J. P., & Peltonen, M. (2015). Adapting existing diabetes risk scores for an Asian population: a risk score for detecting undiagnosed diabetes in the Mongolian population. Bmc Public Health, 15, 938. doi:10.1186/s12889-015-2298-9

Dunder, K., Lind, L., Zethelius, B., Berglund, L., & Lithell, H. (2004). Evaluation of a scoring scheme, including proinsulin and the apolipoprotein B/apolipoprotein A1 ratio, for the risk of acute coronary events in middle-aged men: Uppsala Longitudinal Study of Adult Men (ULSAM). Am Heart J, 148(4), 596-601. doi:10.1016/j.ahj.2004.03.021

Ferrario, M., Chiodini, P., Chambless, L. E., Cesana, G., Vanuzzo, D., Panico, S., . . . Group, C. P. R. (2005). Prediction of coronary events in a low incidence population. Assessing accuracy of the CUORE Cohort Study prediction equation. Int J Epidemiol, 34(2), 413-421. doi:10.1093/ije/dyh405

Friedland, D. R., Cederberg, C., & Tarima, S. (2009). Audiometric pattern as a predictor of cardiovascular status: development of a model for assessment of risk. Laryngoscope, 119(3), 473-486. doi:10.1002/lary.20130

Gabriel, R., Brotons, C., Tormo, M. J., Segura, A., Rigo, F., Elosua, R., . . . group, E. s. (2015). The ERICE-score: the new native cardiovascular score for the low-risk and aged Mediterranean population of Spain. Rev Esp Cardiol (Engl Ed), 68(3), 205-215. doi:10.1016/j.rec.2014.03.019

Gao, W. G., Dong, Y. H., Pang, Z. C., Nan, H. R., Wang, S. J., Ren, J., . . . Qiao, Q. (2010). A simple Chinese risk score for undiagnosed diabetes. Diabet Med, 27(3), 274-281. doi:10.1111/j.1464-5491.2010.02943.x

Gao, W. G., Qiao, Q., Pitkaniemi, J., Wild, S., Magliano, D., Shaw, J., . . . Tuomilehto, J. (2009). Risk prediction models for the development of diabetes in Mauritian Indians. Diabet Med, 26(10), 996-1002. doi:10.1111/j.1464-5491.2009.02810.x

Gaziano, T. A., Young, C. R., Fitzmaurice, G., Atwood, S., & Gaziano, J. M. (2008). Laboratory-based versus non-laboratory-based method for assessment of cardiovascular disease risk: the NHANES I Follow-up Study cohort. Lancet, 371(9616), 923-931. doi:10.1016/S0140-6736(08)60418-3

Glumer, C., Carstensen, B., Sandbaek, A., Lauritzen, T., Jorgensen, T., Borch-Johnsen, K., & inter, s. (2004). A Danish diabetes risk score for targeted screening: the Inter99 study. Diabetes Care, 27(3), 727-733. doi:10.2337/diacare.27.3.727

Griffin, S. J., Little, P. S., Hales, C. N., Kinmonth, A. L., & Wareham, N. J. (2000). Diabetes risk score: towards earlier detection of type 2 diabetes in general practice. Diabetes Metab Res Rev, 16(3), 164-171. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/10867715

Group, W. C. R. C. W. (2019). World Health Organization cardiovascular disease risk charts: revised models to estimate risk in 21 global regions. Lancet Glob Health, 7(10), e1332-e1345. doi:10.1016/S2214-109X(19)30318-3

Gupta, P., Prieto-Merino, D., Ajay, V.S., Singh, K., Roy, A., Krishnan, A., Narayan, V.N., Ali, M.K., Tandon, N., Dorairaj, P., Perel, P. (2019). Cardiovascular Risk Prediction in India: Comparison of the Original and Recalibrated Framingham Prognostic Models in Urban

Populations. [Version 1; peer review : 2 approved with reservations]. Wellcome Open Research, 4(71), 1-14.

Ha, K. H., Lee, Y. H., Song, S. O., Lee, J. W., Kim, D. W., Cho, K. H., & Kim, D. J. (2018). Development and Validation of the Korean Diabetes Risk Score: A 10-Year National Cohort Study. Diabetes Metab J, 42(5), 402-414. doi:10.4093/dmj.2018.0014

Hamer, M., Chida, Y., & Stamatakis, E. (2009). Utility of C-reactive protein for cardiovascular risk stratification across three age groups in subjects without existing cardiovascular diseases. Am J Cardiol, 104(4), 538-542. doi:10.1016/j.amjcard.2009.04.020

Heianza, Y., Arase, Y., Saito, K., Hsieh, S. D., Tsuji, H., Kodama, S., . . . Sone, H. (2013). Development of a screening score for undiagnosed diabetes and its application in estimating absolute risk of future type 2 diabetes in Japan: Toranomon Hospital Health Management Center Study 10 (TOPICS 10). J Clin Endocrinol Metab, 98(3), 1051-1060. doi:10.1210/jc.2012-3092

Heikes, K. E., Eddy, D. M., Arondekar, B., & Schlessinger, L. (2008). Diabetes Risk Calculator: a simple tool for detecting undiagnosed diabetes and pre-diabetes. Diabetes Care, 31(5), 1040-1045. doi:10.2337/dc07-1150

Hippisley-Cox, J., & Coupland, C. (2017). Development and validation of QDiabetes-2018 risk prediction algorithm to estimate future risk of type 2 diabetes: cohort study. BMJ, 359, j5019. doi:10.1136/bmj.j5019

Hippisley-Cox, J., Coupland, C., & Brindle, P. (2013). Derivation and validation of QStroke score for predicting risk of ischaemic stroke in primary care and comparison with other risk scores: a prospective open cohort study. BMJ, 346, f2573. doi:10.1136/bmj.f2573

Hippisley-Cox, J., Coupland, C., & Brindle, P. (2017). Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. BMJ, 357, j2099. doi:10.1136/bmj.j2099

Hippisley-Cox, J., Coupland, C., Robson, J., & Brindle, P. (2010). Derivation, validation, and evaluation of a new QRISK model to estimate lifetime risk of cardiovascular disease: cohort study using QResearch database. BMJ, 341, c6624. doi:10.1136/bmj.c6624

Hippisley-Cox, J., Coupland, C., Robson, J., Sheikh, A., & Brindle, P. (2009). Predicting risk of type 2 diabetes in England and Wales: prospective derivation and validation of QDScore. BMJ, 338, b880. doi:10.1136/bmj.b880

Hippisley-Cox, J., Coupland, C., Vinogradova, Y., Robson, J., May, M., & Brindle, P. (2007). Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. BMJ, 335(7611), 136. doi:10.1136/bmj.39261.471806.55

Hippisley-Cox, J., Coupland, C., Vinogradova, Y., Robson, J., Minhas, R., Sheikh, A., & Brindle, P. (2008). Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. BMJ, 336(7659), 1475-1482. doi:10.1136/bmj.39609.449676.25

Hossain, R., Mahmus, S.M.H., Hossin, M,A., Noori, S.R.H., Jahan, H. (2018). PRMT: Predicting Risk Factor of Obesity among Middle-Aged People Using Data Mining Techniques. Procedia Computer Science, 132, 1068-1076.

Inouye, M., Abraham, G., Nelson, C. P., Wood, A. M., Sweeting, M. J., Dudbridge, F., . . . Group, U. B. C. C. C. W. (2018). Genomic Risk Prediction of Coronary Artery Disease in 480,000 Adults:

Implications for Primary Prevention. J Am Coll Cardiol, 72(16), 1883-1893. doi:10.1016/j.jacc.2018.07.079

Kang, H. M., & Kim, D. J. (2012). Metabolic Syndrome versus Framingham Risk Score for Association of Self-Reported Coronary Heart Disease: The 2005 Korean Health and Nutrition Examination Survey. Diabetes Metab J, 36(3), 237-244. doi:10.4093/dmj.2012.36.3.237

Katulanda, P., Hill, N. R., Stratton, I., Sheriff, R., De Silva, S. D., & Matthews, D. R. (2016). Development and validation of a Diabetes Risk Score for screening undiagnosed diabetes in Sri Lanka (SLDRISK). BMC Endocr Disord, 16(1), 42. doi:10.1186/s12902-016-0124-8

Knuiman, M. W., Vu, H. T., & Bartholomew, H. C. (1998). Multivariate risk estimation for coronary heart disease: the Busselton Health Study. Aust N Z J Public Health, 22(7), 747-753. doi:10.1111/j.1467-842x.1998.tb01487.x

Ko, G., So, W., Tong, P., Ma, R., Kong, A., Ozaki, R., . . . Chan, J. (2010). A simple risk score to identify Southern Chinese at high risk for diabetes. Diabet Med, 27(6), 644-649. doi:10.1111/j.1464-5491.2010.02993.x

L'Italien, G., Ford, I., Norrie, J., LaPuerta, P., Ehreth, J., Jackson, J., & Shepherd, J. (2000). The cardiovascular event reduction tool (CERT)--a simplified cardiac risk prediction model developed from the West of Scotland Coronary Prevention Study (WOSCOPS). Am J Cardiol, 85(6), 720-724. doi:10.1016/s0002-9149(99)00847-4

Laurier, D., Nguyen, P. C., Cazelles, B., & Segond, P. (1994). Estimation of CHD risk in a French working population using a modified Framingham model. The PCV-METRA Group. J Clin Epidemiol, 47(12), 1353-1364. doi:10.1016/0895-4356(94)90079-5

Lees, J. S., Welsh, C. E., Celis-Morales, C. A., Mackay, D., Lewsey, J., Gray, S. R., . . . Mark, P. B. (2019). Glomerular filtration rate by differing measures, albuminuria and prediction of cardiovascular disease, mortality and end-stage kidney disease. Nat Med, 25(11), 1753-1760. doi:10.1038/s41591-019-0627-8

Liao, X., Kerr, D., Morales, J., & Duncan, I. (2019). Application of Machine Learning to Identify Clustering of Cardiometabolic Risk Factors in U.S. Adults. Diabetes Technol Ther, 21(5), 245-253. doi:10.1089/dia.2018.0390

Lindstrom, J., & Tuomilehto, J. (2003). The diabetes risk score: a practical tool to predict type 2 diabetes risk. Diabetes Care, 26(3), 725-731. doi:10.2337/diacare.26.3.725

Menotti, A., Lanti, M., Agabiti-Rosei, E., Carratelli, L., Cavera, G., Dormi, A., . . . Zanchetti, A. (2005). Riskard 2005. New tools for prediction of cardiovascular disease risk derived from Italian population studies. Nutr Metab Cardiovasc Dis, 15(6), 426-440. doi:10.1016/j.numecd.2005.07.007

Merry, A. H., Boer, J. M., Schouten, L. J., Ambergen, T., Steyerberg, E. W., Feskens, E. J., . . . van den Brandt, P. A. (2012). Risk prediction of incident coronary heart disease in The Netherlands: re-estimation and improvement of the SCORE risk function. Eur J Prev Cardiol, 19(4), 840-848. doi:10.1177/1741826711410256

Mohammadreza, B., Farzad, H., Davoud, K., & Fereidoun Prof, A. F. (2012). Prognostic significance of the complex "Visceral Adiposity Index" vs. simple anthropometric measures: Tehran lipid and glucose study. Cardiovasc Diabetol, 11, 20. doi:10.1186/1475-2840-11-20

Moons, K. G., Bots, M. L., Salonen, J. T., Elwood, P. C., Freire de Concalves, A., Nikitin, Y., . . . Grobbee, D. E. (2002). Prediction of stroke in the general population in Europe (EUROSTROKE):

Is there a role for fibrinogen and electrocardiography? J Epidemiol Community Health, 56 Suppl 1, i30-36. doi:10.1136/jech.56.suppl_1.i30

Muhlenbruch, K., Ludwig, T., Jeppesen, C., Joost, H. G., Rathmann, W., Meisinger, C., . . . Schulze, M. B. (2014). Update of the German Diabetes Risk Score and external validation in the German MONICA/KORA study. Diabetes Res Clin Pract, 104(3), 459-466. doi:10.1016/j.diabres.2014.03.013

Muhlenbruch, K., Paprott, R., Joost, H. G., Boeing, H., Heidemann, C., & Schulze, M. B. (2018). Derivation and external validation of a clinical version of the German Diabetes Risk Score (GDRS) including measures of HbA1c. BMJ Open Diabetes Res Care, 6(1), e000524. doi:10.1136/bmjdrc-2018-000524

Nanri, A., Nakagawa, T., Kuwahara, K., Yamamoto, S., Honda, T., Okazaki, H., . . . Japan Epidemiology Collaboration on Occupational Health Study, G. (2015). Development of Risk Score for Predicting 3-Year Incidence of Type 2 Diabetes: Japan Epidemiology Collaboration on Occupational Health Study. PLoS One, 10(11), e0142779. doi:10.1371/journal.pone.0142779

Noda, H., Maruyama, K., Iso, H., Dohi, S., Terai, T., Fujioka, S., . . . "San-yu-kai", M. S. P. C. o. t. J. A. o. O. P. (2010). Prediction of myocardial infarction using coronary risk scores among Japanese male workers: 3M Study. J Atheroscler Thromb, 17(5), 452-459. doi:10.5551/jat.3277

Osborn, D. P., Hardoon, S., Omar, R. Z., Holt, R. I., King, M., Larsen, J., . . . Petersen, I. (2015). Cardiovascular risk prediction models for people with severe mental illness: results from the prediction and management of cardiovascular risk in people with severe mental illnesses (PRIMROSE) research program. JAMA Psychiatry, 72(2), 143-151. doi:10.1001/jamapsychiatry.2014.2133

Panagiotakos, D. B., Georgousopoulou, E. N., Fitzgerald, A. P., Pitsavos, C., & Stefanadis, C. (2015). Validation of the HellenicSCORE (a Calibration of the ESC SCORE Project) Regarding 10-Year Risk of Fatal Cardiovascular Disease in Greece. Hellenic J Cardiol, 56(4), 302-308. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/26233769

Park, Y., Lim, J., Lee, J., & Kim, S. G. (2009). Erythrocyte fatty acid profiles can predict acute non-fatal myocardial infarction. Br J Nutr, 102(9), 1355-1361. doi:10.1017/S0007114509990298

Paynter, N. P., Chasman, D. I., Buring, J. E., Shiffman, D., Cook, N. R., & Ridker, P. M. (2009). Cardiovascular disease risk prediction with and without knowledge of genetic variation at chromosome 9p21.3. Ann Intern Med, 150(2), 65-72. doi:10.7326/0003-4819-150-2-200901200-00003

Paynter, N. P., Mazer, N. A., Pradhan, A. D., Gaziano, J. M., Ridker, P. M., & Cook, N. R. (2011). Cardiovascular risk prediction in diabetic men and women using hemoglobin A1c vs diabetes as a high-risk equivalent. Arch Intern Med, 171(19), 1712-1718. doi:10.1001/archinternmed.2011.351

Pencina, M. J., D'Agostino, R. B., Larson, M. G., Massaro, J. M., & Vasan, R. S. (2009). Predicting the 30-year risk of cardiovascular disease: the framingham heart study. Circulation, 119(24), 3078-3084. doi:10.1161/CIRCULATIONAHA.108.816694

Pocock, S. J., McCormack, V., Gueyffier, F., Boutitie, F., Fagard, R. H., & Boissel, J. P. (2001). A score for predicting risk of death from cardiovascular disease in adults with raised blood pressure, based on individual patient data from randomised controlled trials. BMJ, 323(7304), 75-81. doi:10.1136/bmj.323.7304.75

Pylypchuk, R., Wells, S., Kerr, A., Poppe, K., Riddell, T., Harwood, M., . . . Jackson, R. (2018). Cardiovascular disease risk prediction equations in 400 000 primary care patients in New Zealand: a derivation and validation study. Lancet, 391(10133), 1897-1907. doi:10.1016/S0140-6736(18)30664-0

Rana, J. S., Cote, M., Després, J. P., Sandhu, M. S., Talmud, P. J., Ninio, E., . . . Boekholdt, S. M. (2009). Inflammatory biomarkers and the prediction of coronary events among people at intermediate risk: the EPIC-Norfolk prospective population study. Heart, 95(20), 1682-1687. doi:10.1136/hrt.2009.170134

Ridker, P. M., Buring, J. E., Rifai, N., & Cook, N. R. (2007). Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score. JAMA, 297(6), 611-619. doi:10.1001/jama.297.6.611

Ridker, P. M., Paynter, N. P., Rifai, N., Gaziano, J. M., & Cook, N. R. (2008). C-reactive protein and parental history improve global cardiovascular risk prediction: the Reynolds Risk Score for men. Circulation, 118(22), 2243-2251, 2244p following 2251. doi:10.1161/CIRCULATIONAHA.108.814251

Robinson, C. A., Agarwal, G., & Nerenberg, K. (2011). Validating the CANRISK prognostic model for assessing diabetes risk in Canada's multi-ethnic population. Chronic Dis Inj Can, 32(1), 19-31. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/22153173

Rosella, L. C., Manuel, D. G., Burchill, C., Stukel, T. A., & team, P.-D. (2011). A population-based risk algorithm for the development of diabetes: development and validation of the Diabetes Population Risk Tool (DPoRT). J Epidemiol Community Health, 65(7), 613-620. doi:10.1136/jech.2009.102244

Sarrafzadegan, N., Hassannejad, R., Marateb, H. R., Talaei, M., Sadeghi, M., Roohafza, H. R., . . . Mananas, M. A. (2017). PARS risk charts: A 10-year study of risk assessment for cardiovascular diseases in Eastern Mediterranean Region. PLoS One, 12(12), e0189389. doi:10.1371/journal.pone.0189389

Schulze, M. B., Hoffmann, K., Boeing, H., Linseisen, J., Rohrmann, S., Mohlig, M., . . . Joost, H. G. (2007). An accurate risk score based on anthropometric, dietary, and lifestyle factors to predict the development of type 2 diabetes. Diabetes Care, 30(3), 510-515. doi:10.2337/dc06-2089

Selmer, R., Igland, J., Ariansen, I., Tverdal, A., Njolstad, I., Furu, K., . . . Klemsdal, T. O. (2017). NORRISK 2: A Norwegian risk model for acute cerebral stroke and myocardial infarction. Eur J Prev Cardiol, 24(7), 773-782. doi:10.1177/2047487317693949

Stern, M. P., Williams, K., González-Villalpando, C., Hunt, K. J., & Haffner, S. M. (2004). Does the metabolic syndrome improve identification of individuals at risk of type 2 diabetes and/or cardiovascular disease? Diabetes Care, 27(11), 2676-2681. doi:10.2337/diacare.27.11.2676

Sun, F., Tao, Q., & Zhan, S. (2009). An accurate risk score for estimation 5-year risk of type 2 diabetes based on a health screening population in Taiwan. Diabetes Res Clin Pract, 85(2), 228-234. doi:10.1016/j.diabres.2009.05.005

Tabaei, B. P., Engelgau, M. M., & Herman, W. H. (2005). A multivariate logistic regression equation to screen for dysglycaemia: development and validation. Diabet Med, 22(5), 599-605. doi:10.1111/j.1464-5491.2005.01467.x

Tanabe, N., Iso, H., Okada, K., Nakamura, Y., Harada, A., Ohashi, Y., . . . Group, J. A. L. S. (2010). Serum total and non-high-density lipoprotein cholesterol and the risk prediction of cardiovascular events - the JALS-ECC -. Circ J, 74(7), 1346-1356. doi:10.1253/circj.cj-09-0861

Tohidi, M., Hadaegh, F., Harati, H., & Azizi, F. (2009). C-reactive protein in risk prediction of cardiovascular outcomes: Tehran Lipid and Glucose Study. Int J Cardiol, 132(3), 369-374. doi:10.1016/j.ijcard.2007.11.085

Voss, R., Cullen, P., Schulte, H., & Assmann, G. (2002). Prediction of risk of coronary events in middle-aged men in the Prospective Cardiovascular Münster Study (PROCAM) using neural networks. Int J Epidemiol, 31(6), 1253-1262; discussion 1262-1264. doi:10.1093/ije/31.6.1253

Wen, J., Hao, J., Liang, Y., Li, S., Cao, K., Lu, X., . . . Wang, N. (2017). A non-invasive risk score for predicting incident diabetes among rural Chinese people: A village-based cohort study. PLoS One, 12(11), e0186172. doi:10.1371/journal.pone.0186172

Wickramasinghe, C. D., Ayers, C. R., Das, S., de Lemos, J. A., Willis, B. L., & Berry, J. D. (2014). Prediction of 30-year risk for cardiovascular mortality by fitness and risk factor levels: the Cooper Center Longitudinal Study. Circ Cardiovasc Qual Outcomes, 7(4), 597-602. doi:10.1161/CIRCOUTCOMES.113.000531

Wong, C. K., Siu, S. C., Wan, E. Y., Jiao, F. F., Yu, E. Y., Fung, C. S., . . . Lam, C. L. (2016). Simple non-laboratory- and laboratory-based risk assessment algorithms and nomogram for detecting undiagnosed diabetes mellitus. J Diabetes, 8(3), 414-421. doi:10.1111/1753-0407.12310

Woodward, M., Brindle, P., Tunstall-Pedoe, H., & estimation, S. g. o. r. (2007). Adding social deprivation and family history to cardiovascular risk assessment: the ASSIGN score from the Scottish Heart Health Extended Cohort (SHHEC). Heart, 93(2), 172-176. doi:10.1136/hrt.2006.108167

Wu, Y., Liu, X., Li, X., Li, Y., Zhao, L., Chen, Z., . . . Group, C. M. C. S. o. C. E. R. (2006). Estimation of 10-year risk of fatal and nonfatal ischemic cardiovascular diseases in Chinese adults. Circulation, 114(21), 2217-2225. doi:10.1161/CIRCULATIONAHA.105.607499

Yatsuya, H., Iso, H., Li, Y., Yamagishi, K., Kokubo, Y., Saito, I., . . . Tsugane, S. (2016). Development of a Risk Equation for the Incidence of Coronary Artery Disease and Ischemic Stroke for Middle-Aged Japanese- Japan Public Health Center-Based Prospective Study. Circ J, 80(6), 1386-1395. doi:10.1253/circj.CJ-16-0081

Yatsuya, H., Iso, H., Yamagishi, K., Kokubo, Y., Saito, I., Suzuki, K., . . . Tsugane, S. (2013). Development of a point-based prediction model for the incidence of total stroke: Japan public health center study. Stroke, 44(5), 1295-1302. doi:10.1161/STROKEAHA.111.677534

Zhang, X. F., Attia, J., D'Este, C., Yu, X. H., & Wu, X. G. (2005). A risk score predicted coronary heart disease and stroke in a Chinese cohort. J Clin Epidemiol, 58(9), 951-958. doi:10.1016/j.jclinepi.2005.01.013

Zhou, X., Qiao, Q., Ji, L., Ning, F., Yang, W., Weng, J., . . . Sun, J. (2013). Nonlaboratory-based risk assessment algorithm for undiagnosed type 2 diabetes developed on a nation-wide diabetes survey. Diabetes Care, 36(12), 3944-3952. doi:10.2337/dc13-0593

Appendix D Tables

Appendix D Table 1: Predictors included in QDiabetes, QRISK3 and PRIMROSE

QDiabetes	QRISK3	PRIMROSE
Age (fractional polynomial)	Age (fractional polynomial)	Age (Log)
Townsend Score	Townsend Score	Townsend Score
Smoking	Smoking	Smoking
Atypical Antipsychotics	Atypical Antipsychotics	Second Generation Antipsychotics
Hypertension	Hypertension	Antihypertensives
Schizophrenia / bipolar	Schizophrenia, Depression, bipolar	Schizophrenia / bipolar
BMI (fractional polynomial)	BMI (fractional polynomial)	Systolic Blood Pressure
Ethnicity[1]	Ethnicity[1]	Height
Steroids	Steroids	Weight
FHx CVD	FHx MI	HDL Cholesterol
FHx Type 2 Diabetes	Chol:HDL Ratio	Total Cholesterol
FPG / HbA1C	Type 1 / Type 2 Diabetes	Type 2 Diabetes
Age x Antipsychotics	Age x Smoking	Antidepressants
Age x BMI	Age x BMI	Alcohol Use
Age x Cardiovascular FHx	Age x Cardiovascular FHx	Current Year
Age x HbA1C / FPG	Age x Diabetes	Sex
~~PCOS~~	Age x Steroids	
~~Gestational Diabetes~~	Age x Hypertension	
~~Learning Disabilities~~	Age x Systolic Blood Pressure	
~~Statins~~	Age x Townsend score	
~~Age x Learning Disabilities~~	~~CKD Stages 3-5~~	
~~Age x Statins~~	~~Rheumatoid Arthritis~~	
	~~SLE~~	
	~~Migraines~~	
	~~Erectile Dysfunction~~	
	~~Atrial Fibrillation~~	
	~~Age x Atrial Fibrillation~~	
	~~Age x Migraine~~	
	~~Age x CKD~~	
	~~Age x SLE~~	

FHx=Family History; CVD=Cardiovascular Diseases; MI=Myocardial Infarction; Chol=Cholesterol; HDL=High-Density Lipoprotein; FPG=Fasting Plasma Glucose; HbA1C = Glycated Haemoglobin; BMI=Body Mass Index; PCOS=Polycystic Ovarian Syndrome; CKD=Chronic Kidney Disease; SLE=Systemic Lupus Erythematosus.

Cells with strikethrough text indicate predictors I could not include due to data availability

[1]White/NS, Indian, Pakistani, Bangladeshi, Other Asian, Black African, Black Caribbean, Chinese, Other

Appendix D Table 2: Systematic Review Risk of Bias Assessment Using PROBAST

Author (Year)	ROB				Overall
	Participants	Predictors	Outcome	Analysis	ROB
Abd El-Wahab et al (2019)	-	+	-	-	-
Abdul-Ghani et al (2011)	+	+	+	-	-
Addoh et al (2016)	?	-	-	?	-
Adegbija et al (2015)	+	+	-	-	-
Aekplakorn et al (2006)	+	+	-	-	-
Alaa et al (2018)	+	+	+	-	-
Alghwiri t al (2014)	-	+	-	-	-
Alssema et al (2011)	-	+	-	-	-
Alssema et al (2012)	-	+	-	-	-
Anderson et al (2016)	-	+	-	-	-
Arima et al (2009)	+	+	-	?	-
Artero et al (2015)	+	+	+	-	-
Artigao-Rodenas et al (2013)	+	+	+	-	-
Aslibekyan et al (2011)	+	-	+	-	-
Assmann et al (2002)	+	+	+	?	?
Backholer et al (2017)	-	+	+	?	?
Balkau et al (2004)	+	+	+	-	-
Barazzoni et al (2019)	+	+	+	-	-
Bell et al (2011)	+	+	-	-	-
Boland et al (2005)	?	+	?	-	-
Boucher et al (2019)	+	-	+	?	-
Brand et al (1976)	+	+	-	-	-
Brautbar et al (2009)	-	+	+	-	-
Chambless et al (2003)	+	-	+	-	-
Chen et al (2009)	-	+	+	-	-
Chen et al (2010)	+	+	?	-	-
Chen et al (2017)	+	+	-	-	-
Chien et al (2008)	-	+	+	-	-
Chien et al (2012)	+	+	-	-	-
Choe et al (2018)	+	+	-	?	-
Conroy et al (2003)	+	?	+	-	-

Cross et al (2012)	+	+	+	-	-
D'Agostino et al (2001)	+	+	?	?	?
D'Agostino et al (2008)	+	+	+	?	?
Davies et al (2010)	+	+	+	-	-
De Bacquer et al (2010)	-	+	+	-	-
Dimopoulos et al (2018)	+	+	+	-	-
Dugee et al (2015)	-	+	+	-	-
Dunder et al (2004)	-	+	+	-	-
Ferrario et al (2012)	+	+	+	+	+
Friedland et al (2009)	+	+	?	-	-
Gabriel et al (2015)	+	+	+	-	-
Gao et al (2009)	-	+	+	-	-
Gao et al (2010)	?	?	?	-	-
Gaziano et al (2008)	+	+	+	-	-
Glumer et al (2004)	?	?	?	-	-
Griffin et al (2000)	-	?	?	-	-
Gupta et al (2019)	-	+	?	?	-
Hamer et al (2019)	+	+	+	?	?
Hippisley-Cox & Coupland (2017)	?	+	+	+	?
Hippisley-Cox et al (2013)	?	+	+	+	?
Hippisley-Cox et al (2017)	?	+	+	+	?
Hippisley-Cox et al (2007)	?	+	+	+	?
Hippisley-Cox et al (2008)	?	+	+	+	?
Hippisley-Cox et al (2009)	?	+	+	+	?
Hippisley-Cox et al (2010)	?	+	+	+	?
Heianza et al (2013)	+	+	?	-	-
Heikes et al (2008)	-	?	-	-	-
Hossain et al (2019)	-	+	-	-	-
Hwa Ha et al (2018)	+	+	-	-	-
Inouye et al (2018)	+	+	+	-	-
Kang et al (2012)	?	+	+	?	?
Katulanda et al (2016)	+	?	?	-	-
Kniuman et al (1998)	+	+	+	?	?
Ko et al (2010)	-	?	?	-	-

L'Italien et al (2000)	?	+	+	?	?
Laurier et al (1994)	+	+	?	?	?
Lees et al (2019)	+	+	+	-	-
Liao et al (2019)	+	+	-	-	-
Lindstrom et al (2003)	-	-	-	-	-
Menotti et al (2005)	+	-	+	+	-
Merry et al (2011)	+	+	+	-	-
Mohammadreza et al (2012)	+	+	+	-	-
Moons et al (2002)	+	+	+	-	-
Muehlenbruch et al (2014)	+	-	-	-	-
Muehlenbruch et al (2018)	+	-	-	-	-
Nanri et al (2015)	+	+	-	-	-
Noda et al (2010)	-	-	+	-	-
Osborn et al (2015)	?	+	+	+	?
Panagiotakos et al (2015)	+	+	?	-	-
Park et al (2009)	+	-	+	-	-
Paynter et al (2009)	+	+	-	?	-
Paynter et al (2011)	+	+	+	-	-
Pencina et al (2009)	+	+	+	-	-
Pocock et al (2001)	+	+	+	-	-
Pylypchuk et al (2019)	?	+	+	+	?
Rana et al (2009)	+	+	+	-	-
Ridker et al (2008)	+	+	-	-	-
Ridker et al (2007)	+	+	-	-	-
Robinson et al (2011)	?	+	-	-	-
Rosella et al (2011)	-	+	+	-	-
Sarrafzadegan et al (2017)	+	+	+	-	-
Schulze et al (2007)	?	-	-	-	-
Selmer et al (2017)	+	+	+	?	?
Solares et al (2019)	+	+	+	-	-
Stern et al (2004)	+	+	+	-	-
Sun et al (2009)	+	+	+	-	-
Tabaei et al (2005)	?	+	+	?	?
Tanabe et al (2010)	+	+	-	-	-

Tohidi et al (2008)	+	+	-	-	-
Voss et al (2002)	+	+	-	-	-
Wen et al (2017)	-	+	+	-	-
WHO CVD Risk Working Group (2019)	+	+	+	+	+
Wickramsinghe et al (2014)	+	+	+	?	?
Wong et al (2015)	?	+	-	-	-
Woodward et al (2007)	+	+	+	-	-
Wu et al (2006)	+	+	+	-	-
Yatsuya et al (2012)	+	+	+	-	-
Yatsuya et al (2016)	+	+	+	-	-
Ye et al (2014)	+	+	+	-	-
Zhang et al (2005)	+	+	-	-	-
Zhou et al (2013)	-	+	+	-	-

+ indicates low ROB; - indicates high risk of bias; ? indicates unclear risk of bias

Appendix D Table 3: Participant Characteristics of Studies Included in Systematic Review

		Derivation Cohort					Validation Cohort				
First Author (Year)	**Country**	**Population**	**Sample Size**	**Age, mean (SD)**	**Ethnicity, %**	**Male Sex, %**	**Method**	**Sample Size**	**Age, mean (SD)**	**Ethnicity, %**	**Male Sex, %**
Abd El-Wahab et al (2019)	Egypt	General	270	42.7 (12.7)	-	23.0	-	-	-	-	-
Abdul-Ghani et al (2011)	USA	General	1,562	43 (1)	Caucasian, 35	43.2	External	2,395	46 (1)	Caucasian, 100%	45.9
Addoh et al (2016)	USA	Validation	43,456	Range 20-84	Caucasian, 100	79	External	9,974	Range 20-85	-	-
Aekplakorn et al (2006)	Thailand	General	2,677	42.2 (4.7)	Thai, 100	74.3	External	2,420	44.7 (4.8)	Thai, 100	71.5
Alaa et al (2019)	UK	General	423,604	56.4 (8.1)	-	44.5	Internal, 10-fold cross validation	-	-	-	-
Alghwiri et al (2014)	Jordan, S.Arabia	-	-	-	-	-	External	538	Largest group 40-44 (59%)	Arab, 100	65.5
Alssema et al (2011)	Netherlands, Finland, Sweden, UK, Aus	General	20,564	51.2 (12.7)	-	44.4	Internal, Bootstrap	-	-	-	-
Alssema et al (2012)	Netherlands	General	6,480	58.7 (8.6)	Caucasian, 100	44.8	Internal, Bootstrap	-	-	-	-
Anderson et al (2016)	USA	General	24,331	Largest group 46-60 (30%)	Caucasian, 61	37	External	189,082	-	-	-
Arima et al (2009)	Japan	General	1,756	59 (12)	-	43	Internal, Split Sample	878	59 (12)	-	40
Artero et al (2015)	USA	General	43,356	Range 20-84	Caucasian, 100	79	-	-	-	-	-
Artigao-Rodenas et al (2013)	Spain	General	759	51.1 (12.9)	Spanish, 100	46.7	-	-	-	-	-
Aslibekyan et al (2011)	Costa Rica	General	1,678	-	-	-	Internal, participants excluded from development	1,984	-	-	-
Assmann et al (2002)	Germany	General	5,389	46.7 (7.5)	-	100	Internal, Bootstrap	-	-	-	-
Balkau et al (2004)	European (multi-sample)	General	M=16,506 F=8,907	30-74y	-	65	-	-	-	-	-
Backholer et al (2017)	Australia	General	54,829	55.6 (8.9)	-	41	External	-	-	-	-
Barazzoni et al (2019)	Italy	General	1,965	49 (13)	-	46.4	Internal	263	-	-	-

Bell et al (2011)	USA	General	M=1,351 F=1,484	M=48.6 (9.5) F=48.3 (9.4)	-	-	-	-	-	-	-
Boland et al (2005)	-	General	280	51.1 (11.5)	-	51	External	962	-	-	-
Boucher et al (2019)*	USA	General	M=3,731 F=3,792	-	-	49.6	Internal, Split Sample	-	-	-	-
Brand et al (1976)	USA	General	3,154	39-59y	-	100	-	-	-	-	-
Brautbar et al (2009)	USA	General	9,998	54.1 (5.7)	Caucasian 100%	45.3	-	-	-	-	-
Chambless et al (2003)	USA	General	M=6,071 F=7,983	45-64y	M=77% Caucasian F=71% Caucasian	-	-	-	-	-	-
Chen et al (2009)	Australia	General	67,076	Range 25-74	-	43.5	External	1,998	49-74	-	40.4
Chen et al (2010)	Australia	General	6,060	Largest group 45-54 (30.7%)	-	49.3	External	2 cohorts: 1,993+1,465	-	-	-
Chen et al (2017)	China	General	26,352	57.4 (14.7)	Chinese, 100	44.3	Internal, Subsample	2,015	57.4 (14.7)	Chinese, 100	44.3
Chien et al (2009)	China	General	2,960	54 (12.3)	Chinese, 100	-	Internal, Bootstrap	-	-	-	-
Chien et al (2012)	Taiwan	General	3,430	54.5 (12.3)	-	47.0	External	22,193			
Choe et al (2018)	Korea	General	5,251	49.3 (10.5)	-	39.6	Internal, Split Sample	2,251	52.1 (9.9)	-	74.5
Cross et al (2012)	USA	General	1,084	56.4	99% Caucasian	-	External	623	62.0	42% Caucasian	-
D'Agostino et al (2001)	USA	General	5,251	Range 30-74	Caucasian, 100	46.4	External	23,424	30-81	-	74.4
D'Agostino et al (2008)	USA	General	8,491	49	-	46.7	-	-	-	-	-
Davies et al (2010)	Canada	General	5,642	48.6 (7.2)	-	75.9	External	-	-	-	-
De Bacquer et al (2010)	Belgium	-	-	-	-	-	External	6,212	55 (9.3)	-	51.2
Dimopoulos et al (2018)	Greece	General	2,020	46 (13)	-	-	Internal, 10 fold cross-validation	-	-	-	-
Dugee et al (2015)	Mongolia	General	1,018	46.4 (8.1)	Mongolian, 100	38.4	Internal, Bootstrap	-	-	-	-
Dunder et al (2004)	Sweden	General	574	50	-	100.0	Internal, Subset	534	50	-	100.0

Ferrario et al (2012)	Italy	General	6,865	35-69y	-	100	Internal, Bootstrap	-	-	-	-
Friedland et al (2009)	USA	General	1,168	67.5 (12.7)	-	42.7	External	90	69 (range 44-92)	-	46.7
Gabriel et al (2015)	Spain	General	11,800	57.6 (15.4)	Spanish, 100	45.9	-	-	-	-	-
Gao et al (2009)	India	General	1,544	42.2 (2.1)	Mauritius Indian, 100	36.9	Internal, Split Sample	1,550	-	-	-
Gao et al (2010)	China	General	1,986	53 (11.9)	Chinese, 100	37.3	External	4.336	50 (10.5)	Chinese, 100	38.9
Gaziano et al (2008)	USA	General	M=2,837 F=3,349	M=48.3 (14.0) F=47.4 (14.1)	-	-	-	-	-	-	-
Glumer et al (2004)	Denmark	General	3,250	46 (7.9)	-	49.8	External	1,028	50.7 (7.8)	-	46.0
Griffin et al (2000)	UK	General	650	52.6 (7.4)	Caucasian, 100	43.7	External	528	52.6 (7.4)	Caucasian, 100	43.7
Gupta et al (2019)	India	General	-	-	-	-	Recalibration	M=6,240 F=7,568	M=44.1 (13.9) F=42.0 (12.8)	-	54.7
Hamer et al (2009)	Scotland	General	5,994	53.6 (12.4)	-	44.5	-	-	-	-	-
Heianza et al (2013)	Japan	General	33,335	40-49	Japanese,100	71.4	External	7,477	49.7 (12.1)	Japanese, 100	60.2
Heikes et al (2008)	USA	General	7,092	-	-	-	External	-	-	-	-
Hippisley-Cox & Coupland (2017)	UK	General	8,136,705	44.9 (15.3)	Caucasian, 87.2	49.6	Internal	2,629,940	45.6 (15.5)	Caucasian, 88.4	49.7
Hippisley-Cox et al (2007)	UK	General	1,283,174	48, IQR 40-57	-	48.6	Internal	614,553	47, IQR 40-57	-	49.7
Hippisley-Cox et al (2008)	UK	General	1,535,583	49, IQR 41-60	Caucasian, 97.4	49.6	Internal	750,232	49, IQR 41-59	Caucasian, 96.7	49.9
Hippisley-Cox et al (2009)	UK	General	2,540,753	41, IQR 31-56	Caucasian, 96.7	49.5	Internal	1,232,832	42, IQR 32-56	Caucasian, 96.5	49.5
Hippisley-Cox et al (2010)	UK	General	2,343,759	48.1 (14.3)	Caucasian, 95.1	49.2	Internal	1.267,159	48.0 (14.2)	Caucasian, 96.3	49.1
Hippisley-Cox et al (2013)	UK	General	3,549,478	45.0 (15.4)	Caucasian, 93.9	49.2	Internal	1,897,168	44.9 (15.3)	Caucasian, 92.8	49.3
Hippisley-Cox et al (2017)	UK	General	7,889,803	42.9 (14.7)	Caucasian, 88.8	49.0	Internal	2,671,298	42.9 (14.4)	Caucasian, 89.5	49.1
Hossain et al (2018)	Bangladesh	General	259	-	-	52.9	Internal, 10-fold cross-validation	-	-	-	-

Hwa Ha et al (2018)	S. Korea	General	359,349	51.9 (9.4)	-	53.2	External	6,660	50.8 (8.6)	Caucasian, 47.5	47.5
Inouye et al (2018)	UK	General	482,629	56.5 (8.1)	-	45.6	-	-	-	-	-
Kang et al (2012)	S. Korea	General	M=2,257 F=3,014	M=46 F=44	-	-	-	-	-	-	-
Katulanda et al (2016)	Sri Lanka	General	2,826	45.3 (15.1)	-	39.6	External	1,450	45.1 (14.9)	-	40.5
Ko et al (2010)	Hong Kong	General	2,448	37.2 (8.9)	-	34.7	External	3,734	38.4 (12.8)	-	19.4
Knuiman et al (1998)	Australia	General	M=1,036 F=1,222	M=59.6 (10.6) F=58.4 (10.3)	-	-	-	-	-	-	-
L'Italien et al (2000)	Scotland	General	6,595	45-65y	-	100.0	External	"a subset"	-	-	-
Laurier et al (1994)	France	-	-	-	-	-	Recalibration	M=4,131 F=1635	43-53y	-	71.2
Lees et al (2019)	UK	General	440,526	37-73	-	-	-	-	-	-	-
Liao et al (2019)	USA	General	217,254	43 (12)	-	45.3	Internal, 5-fold cross-validation	-	-	-	-
Lindstrom et al (2003)	Finland	General	-	-	-	-	External	4,586	-	-	-
Menotti et al (2005)	Italy	General	M=11,039 F=4,777	M=51.6 (8.7) F=52.8 (10.5)	-	69.8	-	-	-	-	-
Merry et al (2011)	Netherlands	General	20,055	41.0 (10.9)	-	46.8	-	-	-	-	-
Mohammadreza et al (2012)	Iran	General	M=2,778 F=3,629	>30	-	43.4	Internal, bootstrapping	-	-	-	-
Moons et al (2002)	European (multi-sample)	General	8,309	62 (7)	-	-	-	-	-	-	-
Muehlenbruch et al (2014)	Germany	General	21,846	49.3 (8.8)	-	37.4	External	7,797	49.8 (8.9)	-	48.9
Muehlenbruch et al (2018)	Germany	General	25,392	49.6 (8.9)	-	38.7	External	3,717	42.6 (25.4)	-	49.1
Nanri et al (2015)	Japan	General	24,950	45.5 (7.9)	-	85.6	Internal, Subgroup	12,466	45.5 (7.8)		85.6
Noda et al (2010)	Japan	General	612	50.4 (5.5)	-	100.0	-	-	-	-	-
Osborn et al (2015)	UK	Psychiatric	38,824	49.5 (15.6)	-	47.4	Internal	-	-	-	-

Panagiotakos et al (2015)	Greece	General	3,042	-	-	-	Internal, Subgroup	2,583	-	-	50
Park et al (2009)	Korea	General	100	58.1 (1.7)	-	37.0	-	-	-	-	-
Paynter et al (2009)	USA	General	22,129	53 (48-59y)	-	0.0	-	-	-	-	-
Paynter et al (2011)	USA	General	M=11,280 F=24,674	55(50-62)	-	31.0	Internal, 10-fold cross validation	-	-	-	-
Pencina et al (2009)	USA	General	M=2,173 F=2,333	M=37.3 (9.2) F=26.3 (9.3)	-	48.2	Internal, 5-fold cross validation	-	-	-	-
Pocock et al (2001)	Europe & USA (multi-sample)	General	47,088	-	-	-	-	-	-	-	-
Pylypchuk et al (2019)	New Zealand	General	M=226,053 F=175,699	M=51.8 (9.9) F=56.0 (8.9)	European: M=57% F=55%	56.0	Internal validation – Split Sample	-	-	-	-
Rana et al (2009)	UK	General	2,550	65 (8)	-	63.6	-	-	-	-	-
Ridker et al (2007)	USA	General	16,400	52 (48-58)	White 95.2%	0.0	Internal – Split Sample	8,158	52 (49-59)	White 95.3%	0.0
Ridker et al (2008)	USA	General (Male only)	10,724	63 (IQR 57-70)	-	100	-	-	-	-	-
Robinson et al (2011)	Canada	General	4,366	52.6 (12.5)	Caucasian, 65.7	36.4	Internal, Subgroup	1,857	52.6 (12.5)	Caucasian, 65.7	36.4
Rosella et al (2011)	Canada	General	19.861	45 (-)	Caucasian, 88	46.2	2 x External	9,899 & 26,465	46 (-)	Caucasian, 88.2	46.1
Sarrafzadegan et al (2017)	Iran	General	4,588	51.2 (11.9)	-	48.7	Internal, Bootstrap	-	-	-	-
Schulze et al (2007)	Germany	General	25,167	-	-	38.6	2 x External	23,398 & 657	37 (-)	-	28.6
Selmer et al (2017)	-	General	66,712	52.0 (10.9)	-	44.5	External	39,289	57.8 (12.0)	-	50.8
Solares et al (2019)	UK	General	64,772	50	-	29.9	Internal – Split Sample	16,192	-	-	-
Stern et al (2004)	USA	General	1,709	25-64	-	-	External	1,353	35-64y	-	-
Sun et al (2009)	Taiwan	General	35,972	47.5 (9.9)	-	-	Internal	36,989	-	-	-
Tabaei et al (2005)	Egypt	General	516	44 (15)	-	42	External	516	45 (14)	-	43
Tanabe et al (2010)	Japan	General	22,430	57.8 (11.1)	-	40.0	-	-	-	-	-
Tohidi et al (2008)	Iran	General	385	59 (11)	-	67.0	-	-	-	-	-

Voss et al (2002)	Germany	General	5,159	51.9 (6.7)	-	100.0	Internal, 4-fold cross validation	-	-	-	-
Wen et al (2017)	China	General	2,845	51 (11)	-	44.4	External	1,287	51 (11)	-	44.4
WHO CVD Risk Chart Working Group (2019)	Worldwide (Multi-study)	General	M=202,962 F=173,215	M=53 (48-60) F=55 (49-63)	-	64.2	External	1,096,061	-	-	-
Wickramsinghe et al (2014)	USA	General	16,533	42.2 (9.8)	Caucasian, 100	82.4	Internal, Cross-validation	-	-	-	-
Wong et al (2015)	China	General	2,518	-	-	92.7	External	839	-	-	92.5
Woodward et al (2007)	Scotland	General	M=6,450 F=6,757	M=48.9 (0.1) F=48.8 (0.1)	-	48.8	-	-	-	-	-
Wu et al (2006)	China	General	M=4,890 F=5,013	46 (6)	-	49.3	External	17,329	-	-	-
Yatsuya et al (2013)	Japan	General	15.672	-	-	33.9	External	3,454	-	-	33.9
Yatsuya et al (2016)	Japan	General	15,672	57.4	-	33.9	External	11,598	50.2 (5.8)	-	37.0
Ye et al (2014)	China	General	1,912	58.3 (5.9)	-	41.9	Internal, Cross-validation	-	-	-	-
Zhang et al (2005)	China	General	3,000	45 (8)	-	100.0	Internal, Split sample	1,400	-	-	-
Zhou et al (2013)	China	General	41,809	44 (14)	Chinese, 100	39.5	2 x External	1,162 & 1,693	50 (10) & 52 (11)	Chinese, 100	34.9 & 39.6

- indicates no data available

Appendix D Table 4: Algorithm Characteristics of Studies Included in Systematic Review

Author (Year)	Outcome			Algorithm				
	Risk Predicted	**Split by Sex**	**Time-frame of Risk**	**Predictors in Model**	***n* Events**	**EPV Ratio**	**Variable Selection Method**	**Designed For**
Abd El-Wahab et al (2019)	Metabolic Syndrome	No	Cross-Sectional	Obesity, Morbid Obesity, Employment, FHx Chronic Illness	156	7.8	Forward Stepwise	Clinical Practice
Abdul-Ghani et al (2011)	T2DM	No	8 Years	Age, Sex, Ethnicity, BMI, FPG, HDL, SBP, 1hrGlucose	174	21.8	Previous Research	Prediction
Addoh et al (2016)	CVD mortality	No	11 Years	Sex, Age, BMI, WC, HR, Physical Activity, Smoking	79	11.3	-	Prediction
Aekplakorn et al (2006)	T2DM	No	12 Years	Age Sex, BMI, WC, HTN, FHx T2DM, IFG	125	27.8	Univariable	Prediction
Alaa et al (2019)	CVD	No	5 years	473 variables	4.801	10.2	Machine Learning	Research
Alghwiri et al (2014)	T2DM	No	Cross-Sectional	Age, Sex, WC, BMI, Physical Activity, Fruit/veg intake, HTN, FPG, FHx T2DM, Ethnicity, Education	-	-	Previous Research	Screening
Alssema et al (2011)	T2DM	No	5 Years	Age, BMI, WC, BP meds, IFG, Sex, Smoking, FHx T2DM	844	84.4	Previous Research	Prediction
Alssema et al (2012)	T2DM, CVD	Yes	7 Years	Age, BMI, WC, BP meds, Smoking, FHx CVD, FHx T2DM	839	69.6	Backward Elimination	Prediction
Anderson et al (2016)	T2DM, Prediabetes	No	-	T2DM, FPG, HTN, Tri, Lipid Disorder, Insurance Status, Ethnicity, Income Prediabetes = Age, Insurance, BMI, Temperature, FPG, Tri, ALT, CRP, HDL	3,765	8.51	Machine learning	Prediction, Research
Arima et al (2009)	CVD	No	14 years	Age, Sex, SBP., T2DM, LDL, HDL, smoking	216	8.1	Previous Research	Clinical Practice
Artero et al (2015)	CVD mortality	No	15 Years	Age, BMI, WC, HR, Physical Activity, Smoking, Sex	M=577 F=50	M=52.4 F=4.54	Previous Research	Prediction
Artigao-Rodenas et al (2013)	Cardiovascular Events	Yes	10 Years	Age, Chol, HDL, SBP, HTN, T2DM, Smoking	M=26 F=47	M=3.7 F=6.7	Previous Research	Prediction
Aslibekyan et al (2011)	MI	No	Cross-Sectional	Dietary trans fats, Dietary saturated fats, Dietary polyunsaturated fats, Dietary cholesterol, Dietary fibre, Dietary folate, "other components", physical activity, smoking, alcohol, socioeconomic status, waist:hip ratio.	839	64.5	Previous Research	Clinical Practice
Assmann et al (2002)	Acute Coronary Event	No	10 Years	Age, LDL, Smoking, HDL, SBP, FHx MI, Tri, T2DM	325	5.7	Previous Research	Prediction
Backholer et al (2017)	CVD mortality	No	5 Years	Age, Sex, Smoking, T2DM, SBP, HDL, Social Deprivation Score, eGFR, Sex*T2DM, Sex*HDL, Sex*Social Deprivation, Age*SBP, Age*Smoking	1,375	80.8	Univariable	Prediction
Balkau et al (2004)	CVD	Yes	10 years	Age, study centre, FPG, smoking, SBP, cholesterol, BMI	632	27.5	Previous Research	Clinical Practice

Barazzoni et al (2019)	Prediabetes	No	5 years	Age, Sex, T2DM, HTN, Cholesterol	497	24.9	Forward Stepwise	Clinical Practice
Bell et al (2011)	CVD	Yes	10 years	Averaged 2 measures (SBP, Cholesterol, HDL), Smoking, T2DM, HTN meds, Cholesterol meds,	M=77 F=207	M=11.0 F=29.6	Previous Research	Clinical Practice
Boland et al (2005)	CVD	No	10 Years	Age, SBP, Smoking, Cholesterol, Previous CVD event, FHx CVD, T2DM	77	9.6	Previous Research	Clinical Practice
Boucher et al (2019)*	Metabolic Syndrome	No	Cross-Sectional	Mid-Upper Arm Circumference Cut-offs	M=598 F=683	M=598 F=683	Machine Learning	Clinical Practice
Brand et al (1976)	CVD	No	8.5 years	Age, Cholesterol, SBP, Smoking, Weight, ECG, haematocrit	257	28.6	Previous Research	Research
Brautbar et al (2009)	CVD	No	15 years	Age, Sex, Smoking, T2DM, SBP, HTN meds, cholesterol, HDL, 9p21 genotype	1,349	134.9	Previous Research	Research
Chambless et al (2003)	CVD	Yes	10 Years	Cholesterol, HDL, SBP, BP meds, smoking, T2DM, BMI, WHR, Keys Score, Albumin, WCC, FEV, fibrinogen, factor VII, vWF, HR, pack-years, exercise, creatinine	M=599 F=345	M=31.5 F=18.2	Previous Research	Research
Chen et al (2009)	CVD mortality	Yes	10 Years	Age, Chol, SBP, Smoking	62	12.3	Previous Research	Prediction
Chen et al (2010)	T2DM	No	5 Years	Age, Sex, Ethnicity, FHx T2DM, Raised FPG, BP meds, Smoking, Physical activity, WC	445 (33)	13.5	Stepwise Backward	Prediction
Chen et al (2017)	T2DM	No	-	Age, FHx T2DM, Diet, BMI, HTN, FPG	387	29.8	Univariable	Prediction
Chien et al (2009)	T2DM	No	10 Years	Age, FPG, BMI, Tri, WCC, HDL	548	36.5	Univariable	Prediction
Chien et al (2012)	CVD	No	10 Years	Age, Sex, BMI, SBP, FHx CVD, Smoking	171	15.5	Subset Stepwise	Clinical Practice
Choe et al (2018)	Metabolic Syndrome	No	Cross-Sectional	Age, Sex. BMI, Smoking, Alcohol, Exercise, 10xSNPs	223	13.9	Machine Learning	Clinical Practice
Cross et al (2012)	Non-fatal CVD	No	5 Years	Age, Sex, T2DM, FHx MI, CTACK, Eotaxin, Fas Ligand, HGF, IL-16, MCP-3, sFas	385	8.75	Forward Selection	Clinical Practice
D'Agostino et al (2001)	CHD	Yes	5 Years	BP, Chol, Age, Smoking, T2DM	Males ARIC – 195 PHS – 182 HHP – 77 PR – 107 SHS – 46 CHS – 71 Females ARIC – 80 SHS – 23 CHS 44	Males ARIC – 39 PHS – 36.4 HHP – 15.4 PR – 21.3 SHS – 9.2 CHS 14.2 Females ARIC – 9 SHS – 4.6 CHS – 8.8	Previous Research	Prediction
D'Agostino et al (2008)	CVD	Yes	12 Years	Age, Chol, HDL, SBP, BP meds, Smoking, T2DM	M=718 F=456	M=65.3 F=41.5	Univariable	Prediction

Davies et al (2010)	CVD	No	-	Age, T2DM, HTN, Smoking, Cholesterol, HDL, Sex, +12 SNPs	3,323	174.9	Machine Learning	Research
De Bacquer et al (2010)	CVD Mortality	No	10 Years	Age, Sex, Smoking, SBP, Chol	274	45.7	Previous Research	Prediction
Dimopoulos et al (2018)	CVD	No	10 Year	Age, Sex, Smoking, School Years, Diet Score, BMR, BMI, DBP, SBP, Hx HTN, FPG, T2DM, Chol, TG, Hx Hyperchol, IL-6	317	3.17	Univariable, Machine Learning	Clinical Practice
Dugee et al (2015)	T2DM	No	Cross-Sectional	Sex, WC, HTN, FPG, Physical Activity, Sitting time	59	3.7	Univariable	Screening
Dunder et al (2004)	CVD	No	28.7 Years	Apo B/Apo A Ratio, Proinsulin. SBP. MI, Smoking	135	7.5	Univariable	Clinical Practice
Ferrario et al (2012)	CVD	No	10 Years	Age, Cholesterol, SBP, Smoking, HDL, T2DM, BP meds, FHx CVD	312	10.8	Forward Selection	Clinical Practice
Friedland et al (2009)	CVD	No	Cross-Sectional	Age, HTN, T2DM, Cholesterol, Smoking, Audiogram Patterns	316	28.7	Enter	Clinical Practice
Gabriel et al (2015)	CVD	Yes	10 Year	Age, Smoking, T2DM, SBP, Chol	1,214	71.4	Univariable	Prediction
Gao et al (2009)	T2DM	Yes	11 Years	Age, BMI, WC, FHx T2DM	511	21.3	Univariable	Prediction
Gao et al (2010)	T2DM	No	Cross-Sectional	Age, Sex, WC, FHx T2DM	194	11.4	Stepwise Backward	Screening
Gaziano et al (2008)	CVD	Yes	21 Years	Age, SBP, Smoking, Cholesterol, T2DM. BP meds	1529	191.1	Previous Research	Clinical Practice
Glumer et al (2004)	T2DM	No	Cross-Sectional	Age, Sex, BMI, HTN, Physical Activity, FHx T2DM	135	9.0	Stepwise Backward	Screening
Griffin et al (2000)	T2DM	No	Cross-Sectional	Age, Sex, BMI, Steroids, BP Meds, FHx T2DM, Smoking	25	2.1	Forward Selection	Screening
Gupta et al (2019)*	CVD	Yes	10 Years	Age, BMI, SBP, BP Meds, Smoking, T2DM	-	-	Previous Research	Clinical Practice
Hamer et al (2009)	CVD	No	7 Years	Age, Cholesterol, HDL, SBP, BP Meds, Smoking, T2DM, CRP	308	34.2	Previous Research	Research
Heianza et al (2013)	T2DM	No	4 Years	Age, Sex, FHx T2DM, Smoking, BMI, HTN	965	53.6	Backward Elimination	Prediction
Heikes et al (2008)	T2DM or Prediabetes	No	Cross-Sectional	Age, Sex, Weight, Height, WHR, BMI, HTN, FHx T2DM	326	18.1	Machine Learning	Screening
Hippisley-Cox & Coupland (2017)	T2DM	Yes	10 Years	Age, Ethnicity, Social Deprivation, BMI, Smoking, FHx T2DM, CVD, BP Meds, Steroids, Atypical Antipsychotics, Statins, SMI, LD, Gestational Diabetes, PCOS, HbA1C	M=100,419 F=77,895	M=4,781.9 F=2,434.2	Previous Research; Best Subset Selection	Prediction
Hippisley-Cox et al (2007)	CVD	Yes	10 Years	Age, Chol:HDL Ratio, SBP, BMI, FHx CVD, Smoking, Deprivation, BP	M=37,843 F=27,828	M=3,153.6 F=2,319.0	Previous Research; Best Subset Selection	Prediction

Hippisley-Cox et al (2008)	CVD	Yes	10 Years	Ethnicity, Age, BMI, Townsend Score, SBP, HDL:Chol Ratio, FHx CVD, Smoking, HTN, T2DM, RA, AF, Renal Disease	M=55,667 F=41,042	M=1,855.6 F=1,368.1	Previous Research; Best Subset Selection	Prediction
Hippisley-Cox et al (2009)	T2DM	Yes	10 Years	Age, Ethnicity, Deprivation, Smoking, FHx T2DM, CVD, BP Meds, Steroids, BMI	M=43,165 F=34,916	M=1,541.6 F=1,247.0	Previous Research; Best Subset Selection	Prediction
Hippisley-Cox et al (2010)	CVD	Yes	10 Years	BMI, SBP, HDL:Chol Ratio, Townsend Score, Smoking, Ethnic Group, FHx CVD, T2DM, BP Meds, RA, AF, Renal Disease	121,623	5,288	Previous Research; Best Subset Selection	Prediction
Hippisley-Cox et al (2013)	Stroke/TIA	Yes	10 Years	Age, BMI, Smoking, Ethnicity, SBP, Chol:HDL ratio, BMI, FHx CVD, Deprivation, BP Meds, RA, Renal Disease, T1DM, T2DM, AF, CCF, Valve Disease	M=38,074 F=39,504	M=1,057.6 F=1,039.6	Previous Research; Best Subset Selection	Prediction
Hippisley-Cox et al (2017)	CVD	Yes	10 Years	Deprivation, Ethnicity, Smoking, FHx CVD, T1DM, T2DM, BP Meds, RA, AF, CKD>3, Migraine, Steroids, SLE, Atypical Antipsychotics, SMI, Erectile Dysfunction, Chol:HDL ratio, SBP	M=203,016 F=160,549	M=5,075.4 F=4,013.7	Previous Research; Best Subset Selection	Prediction
Hossain et al (2018)	Obesity	No	Cross-Sectional	Age, Height, Weight, Marital Status, Healthy Lifestyle, BMI, Sleep	-	-	Machine Learning	Research
Hwa Ha et al (2018)	T2DM	Yes	10 Years	Age, FHx T2DM, ETOH, Smoking, Physical Activity, BP Meds, Statins, BMI, SBP, Chol, FPG	37,678	1,569.9	Previous Research	Prediction
Inouye et al (2018)	CVD	No	6 Years	Genome Score, T2DM, BMI, Smoking, HTN, FHx CVD, Chol	22,242	3,177.4	Enter	Research
Kang et al (2012)	CVD	Yes	Cross-Sectional	Age, Smoking, HTN, TG, HDL, FPG, Weight	M=39 F=63	M=5.6 F=9.0	Previous Research	Clinical Practice
Katulanda et al (2016)	T2DM	No	Cross-Sectional	Age, WC, BMI, HTN, Balanitis/Vulvitis, FHx T2DM, Gestational DM, Physical Activity, Osmotic Symptoms	128	8.5	Univariable	Screening
Knuiman et al (1998)	CVD	Yes	15 Years	Age, SBP, BP Meds, HDL:Chol Ratio, Smoking, ECG, CVD, T2DM	M=243 F=172	M=18.7 F=13.2	Univariable	Clinical Practice
Ko et al (2010)	T2DM	No	Cross-Sectional	Age, Sex, BMI, HTN, Lipids, FHx T2DM, Gestational DM	1,270	181.4	Univariable	Screening
L'Italien et al (2000)	MI	No	5 Years	Age DBP, HDL:Chol Ratio, Smoking, T2DM, FHx CVD, Angina, Statin	-	-	Previous Research	Clinical Practice
Laurier et al (1994)	CVD	Yes	10 Years	Age, Chol, SBP, T2DM, Smoking	-	-	Previous Research	Clinical Practice
Lees et al (2019)	CVD	No	8 Years	Age, Sex, Ethnicity, Smoking, SBP, DBP, BP Meds, Statins, Chol, HDL, eGFR	2,552	159.5	Previous Research	Clinical Practice
Liao et al (2019)	T2DM	No	Cross-Sectional	Age, WC, BMI, TG, SBP, HDL, LDL, Chol, DBP, Sedentariness, Stress Score, Alcohol, Fruit & Veg, Sleep, Sex, Smoking, Education, Depression, Hx CVD.	17,598	606.8	Machine Learning	Research

Lindstrom et al (2003)	T2DM	No	5 Years	Age, BMI, WC, BP Meds, FPG, Physical Activity, Fruit/Veg Consumption	182	18.2	Previous Research	Prediction
Menotti et al (2005)	CVD	Yes	15 Years	Age. BP. BMI, HDL Non-HDL Chol, T2DM, Smoking, HR	M=397 F=111	M=14.7 F=4.1	Previous Research	Clinical Practice
Merry et al (2011)	CVD	No	10 Years	Age, Sex, Smoking, SBP, HDL:Chol ratio	783	97.9	Previous Research	Clinical Practice
Mohammadreza et al (2012)	CVD	Yes	9 Years	Age, SBP, BP meds, HDL, Chol, T2DM, Smoking, VAI	534	33.38	Previous Research	Clinical Practice
Moons et al (2002)	Stroke	No	7.3 Years	Age, Hx Stroke, T2DM, Smoking, Hx, HTN, DBP	219	13.7	Backward Elimination	Clinical Practice
Muehlenbruch et al (2014)	T2DM	No	5 Years	Age, Height, Weight, WC, HTN, ETOH, Exercise, Smoking, Ex-Smoker, Wholegrain intake, Coffee intake, Red meat intake, FHx T2DM(x3)	492	49.2	Previous Research	Prediction
Muehlenbruch et al (2018)	T2DM	No	5 Years	Age, Height, WC, HTN, Physical Activity, Ex-Smoker x2, Smoking x 2, Wholegrain intake, coffee intake, red meat intake, FHx T2DM x 3, HbA1C	857	57.1	Previous Research	Prediction
Nanri et al (2015)	T2DM	No	3 Years	Sex, Age, BMI, WC, Smoking, HTN, Dyslipidaemia, FPG, HbA1C	1,122	74.8	Backward Elimination	Prediction
Noda et al (2010)	MI	No	1 Year	BP, LDL, HDL, TG, FPG, Smoking	204	12.8	Enter	Clinical Practice
Osborn et al (2015)	CVD	No	10 Years	Sex, Age, BMI, SBP, Weight, Height, T2DM, Smoking, Year, Deprivation, Antidepressants, Alcohol, SMI, FGA, SGA	2,324	154.93	Backward Elimination	Prediction
Panagiotakos et al (2015)	CVD	Yes	10 Years	Smoking, BMI, Hypercholesterolaemia, HTN, T2DM	47	4.27	Previous Research	Prediction
Park et al (2009)	MI	No	Cross-Sectional	EPA, ALA, trans-oleic acid, arachidonic acid	50	3.1	Univariable	Research
Paynter et al (2009)	CVD	No	10 Years	Age, SBP, Chol, HDL, Smoking, BP Meds, T2DM, HbA1C, hsCRP, FHx MI, rs1075724 genotype AG/GG	715	59.6	Previous Research	Clinical Practice
Paynter et al (2011)	CVD	Yes	10 Years	Age, SBP, Chol, HDL, Smoking, CRP, FHx MI, HbA1C	M=170 F=125	M=21.3 F=15.7	Previous Research	Clinical Practice
Pencina et al (2009)	CVD	No	30 Years	Age, Sex, SBP, BP Meds, Smoking, T2DM, Chol, HDL, BMI	671	33.6	Enter	Clinical Practice
Pocock et al (2001)	CVD mortality	No	5 Years	Age, Sex, SBP, Chol, Height, Creatinine, Smoking, T2DM, LVH, Hx Stroke, Hx MI.	1,639	102.4	Enter	Clinical Practice
Pylypchuk et al (2019)	CVD	Yes	5 Years	Deprivation, Sex, Age, Ethnicity, FHx CVD, Smoking, T2DM, SBP, HDL:Chol Ratio, AF, BP Meds, Chol Meds, Anti-thromotic Meds.	M=9,736 F=5,650	M=463.6 F=269.1	Previous Research	Clinical Practice
Rana et al (2009)	CVD	No	6 Years	Age, Sex, SBP, Chol, HDL, TG, CRP, MPO, Type 2 PhosA2, Lipoprotein Phos A2, MCP-1, Adiponectin	921	70.8	Previous Research	Clinical Practice
Ridker et al (2007)	CVD	No	10 Years	Age, HbA1C, SBP, Smoking, ApoB, hsCRP, ApoA-1, FHx MI	504	14.4	Forward Stepwise	Clinical Practice
Ridker et al (2008)	CVD	Yes	10 Years	Age, BP, Smoking, Chol, HDL, CRP, FHx MI <60	1072	107.2	Previous Research	Prediction

Robinson et al (2011)	Prediabetes/T2DM	No	Cross-Sectional	Sex, Age, BMI, WC, Physical Activity, Daily fruit/veg, HTN, FHx T2DM, Maternal Ethnicity, Paternal Ethnicity, Education, Self-reported health status, Smoking, Gestational DM, Macrosomia	852	37.1	Previous Research	Screening
Rosella et al (2011)	T2DM	Yes	9 Years	Age, Height, Weight, Chronic Disease, Ethnicity, Immigration Status, Smoking, Educational Achievement, Household Income, Alcohol, Physical Activity	M=714 F=651	M=47.6 F=43.4	Previous Research	Prediction
Sarrafzadegan et al (2017)	CVD	No	10 Years	Age, Sex, WHR, SBP, Tri, T2DM, Smoking, FHx CVD	705	41.5	Forward Selection	Prediction
Schulze et al (2007)	T2DM	No	7 Years	Age, Sex, Weight, Height, BMI, WC, HTN, Alcohol, Physical Activity, Occupational Activity, Education, Smoking, Food Intake, lgAge, lgWC, Sex*Height, Sex*WC	849	28.3	Forward Selection	Prediction
Selmer et al (2017)	CVD	Yes	10 Year	Chol, HDL, Smoking, SBP, BP meds, FHx CVD	M=3,658 F=2,459	M=609.7 F=409.8	Previous Research	Prediction
Solares et al (2019)	CVD	No	10 Years	Sex, SBP, Smoking, Deprivation, T2DM, Chol, HDL, LDL, Past SBP	3,222	161.1	Machine Learning	Clinical Practice
Stern et al (2004)	T2DM	No	5 Years	Age, Sex, Ethnicity, FPG, SBP, HDL, FHx T2DM, Met Syndrome	195	21.7	Previous Research	Clinical Practice
Sun et al (2009)	T2DM	No	5 Year	Sex, Age, Education, Smoking, BMI, WC, HTN, FPG, Tri, ALT, eGFR, HDL	1,770	104.1	Previous Research	Prediction
Tabaei et al (2005)	Prediabetes/T2DM	No	Cross-Sectional	Age, Sex, BMI, Post-prandial time, SBP, Glu, HDL	77	5.5	Forward Selection	Screening
Tanabe et al (2010)	MI	No	5 Years	Sex, Age, HDL, Chol, BMI, HTN, T2DM, Smoking	104	8.0	Univariable	Clinical Practice
Tohidi et al (2008)	CVD	No	3 Years	Smoking, WHR, HTN, T2DM, FHx CVD, Chol, HDL, CRP	207	25.9	Forward Selection	Clinical Practice
Voss et al (2002)	MI	No	10 Years	Age, SBP, LDL, HDL, TG, Gamma-GT, BMI, Height, Smoking, T2DM, HTN, FHx MI, FHx HTN	325	5.70	Machine Learning	Clinical Practice
Wen et al (2017)	T2DM	No	-	Age, BMI, WC, FHx T2DM	218	18.2	Forward Stepwise	Prediction
WHO CVD Risk Chart Working Group (2019)	CVD	Yes	10 Years	Age, Smoking, SBP, T2DM, Chol	19,333	3,866.6	Previous Research	Clinical Practice
Wickramsinghe et al (2014)	CVD Death	No	30 Years	Age, BMI, SBP, Fitness, T2DM, Chol, Smoking, Sex	M=1,027 F=96	M=51.35 F=4.8	Previous Research	Prediction
Wong et al (2015)	T2DM	No	XS	Age, Sex, Smoking, BMI, FHx T2DM, Exercise, BP, WC, Tri, Chol, HDL, LDL	209	13.1	Forward Stepwise	Screening
Woodward et al (2007)	CVD	Yes	10 Years	Deprivation, Chol, HDL. SBP, Smoking, Fhx CVD, T2DM	M=743 F=422	M=57.2 F=32.5	Univariable	Clinical Practice
Wu et al (2006)	CVD	Yes	10 Years	Age, SBP, Chol, BMI, Smoking, T2DM	M=224 F=266	M=17.2 F=20.5	Univariable	Clinical Practice
Yatsuya et al (2013)	Stroke	No	14 Years	Age, Sex, Smoking, BMI, HTN, BP Meds, T2DM, Sex*Smoking, HTN*BP Meds	790	30.4	Backward Selection	Prediction

Yatsuya et al (2016)	CVD	No	17 Years	Age, Sex, Smoking, SBP, BP Meds, T2DM, HDL, LDL	744	16.9	Backward Selection	Prediction
Ye et al (2014)	T2DM	No	6 Years	Sex, HTN, BMI, FPG, HbA1C, CRP	924	71.1	Backward Elimination	Prediction
Zhang et al (2005)	CVD	No	10 Years	BP, Age, Chol, BMI, Smoking	55	3.9	Univariable	Clinical Practice
Zhou et al (2013)	T2DM	No	Cross-Sectional	Sex, Age, BMI, WC, SBP, FHx T2DM	2,520	172.8	Forward Stepwise	Screening

- indicates no data available

CVD = Cardiovascular Diseases; T2DM = Type 2 Diabetes Mellitus; EPV = Events Per Variable; M/F = Male or Female; SysBP = Systolic Blood Pressure; T2DM = Type 2 Diabetes Mellitus; Tot Chol = Total Cholesterol; HTN = Diagnosis of Hypertension; HDL = High-Density Lipoprotein; FHx T2DM = Family history of Type 2 Diabetes Mellitus; BP Meds = Prescribed Antihypertensive Medication; WC = Waist Circumference; FHx CVD = Family History Cardiovascular Diseases; Phys Act = Physical Activity; FPG = Fasting Plasma Glucose; Tri = Triglycerides; Chol:HDL = Cholesterol:HDL Ratio; LDL = Low-Density Lipoprotein; ETOH = Alcohol Use; ECG = Electrocardiogram Findings; CVD Event = Personal History of Cardiovascular Diseases; HbA1C = Glycated Haemoglobin; WHR = Waist:Hip Ratio; Genetic = Genotype Data; DBP = Diastolic Blood Pressure; Gest DM = Gestational Diabetes Mellitus; RA = Rheumatoid Arthritis; Renal Dis = Renal Disorders; HR = Heart Rate; SMI = Diagnosis of Serious Mental Illness; eGFR = Glomerular Filtration Rate; IFG = Impaired Fasting Glucose; ALT = Alanine Aminotransferase; Atyp Antipsych = Prescribed Antipsychotic Medication; T1DM = Type 1 Diabetes Mellitus; WCC = White Cell Count; Chron Dis = Personal History of Chronic Disease; ApoA/ApoB = Apolipoprotein A/B Levels.

Appendix D Table 5: Algorithm Performance of Studies Included in Systematic Review

	Model Performance (Derivation Cohort)					Model Performance (Validation Cohort)					Other	
Author (Year)	**Discrimination**	**Calibration[a]**	**Sens**	**Spec**	**PPV**	**Discrimination**	**Calibration[a]**	**Sens**	**Spec**	**PPV**	**Analysis to assess uptake**	**Economic Analysis**
Abd El-Wahab et al (2019)	C=0.83 (0.80-0.89)	-	-	-	-	-	-	-	-	-	No	No
Abdul-Ghani et al (2011)	-	-	77.8	77.4	44.8	-	-	76	72	11.9	No	Discussed
Addoh et al (2016)	-	-	-	-	-	-	-	-	-	-	Discussed	No
Aekplakorn et al (2006)	C=0.79	p>0.05	77	60	-	C=0.78 (0.72-0.79)	p=0.88	84.4	52.5	-	Discussed	No
Alaa et al (2018)	C=0.76 (0.76-0.76)		69.9%	-	2.6%	-	-	-	-	-	No	No
Alghwiri et al (2014)	-	-	-	-	-	-	-	-	-	-	Discussed	No
Alssema et al (2011)	C=0.74 (0.73-0.76)	10.0, p=0.27	-	-	-	C=0.77 (0.75-0.78)	10.0 p=0.27	76	63	-	Questionnaire	No
Alssema et al (2012)	M= C=0.80 (0.78-0.82) F= C=0.82 (0.81-0.83)	M=6.3,p=0.62 F=7.6, p=0.48	M=75 F=83	M=66 F=62	-	M= C=0.80 (0.78-0.81) F= C=0.82 (0.80-0.83)	-	-	-	-	Questionnaire	No

Anderson et al (2016)	T2DM C=0.78 Prediabetes C=0.72	-	-	-	-	T2DM C=0.78 Prediabetes -	-	-	-	-	No	No
Arima et al (2009)	-	-	-	-	-	C=0.81 (0.77-0.86)	Calibration Plots	-	79	-	No	No
Artero et al (2015)	C= 0.68 (0.66-0.81)	-	-	-	-	-	-	-	-	-	No	No
Artigao-Rodenas et al (2013)	M= C=0.78 (0.71-0.85) F= C=0.79 (0.72-0.86)	M=10.3, p=0.25; F=6.58,p=0.58	79.1	65.0	25.2	-	-	-	-	-	No	No
Aslibekyan et al (2011)	-	-	-	-	-	C=0.63	-	-	-	-	No	No
Assmann et al (2002)	C=0.82	6.5 p=0.30	-	-	-	-	-	-	-	-	Creation of simple points-scoring system	No
Backholer et al (2017)	C=0.91 (0.89-0.93)	p=0.42	-	-	-	C=0.75 (0.71-0.79)	-	-	-	-	GUI under consideration	No
Balkau et al (2004)	-	-	-	-	-	-	-	-	-	-	No	No
Barazzoni et al (2019)	ROC Curves reported only	-	-	-	-	-	-	-	-	-	No	No
Bell et al (2011)	M= C=0.78 F=C=0.78	M= p=0.24 F = p=0.96	-	-	-	-	-	-	-	-	No	No
Boland et al (2005)	-	-	-	-	-	-	-	-	-	-	Feasibility Assessment Included	No
Boucher et al (2019)*	-	-	-	-	-	-	-	87.0	66.6	70.5	No	No
Brand et al (1976)	-	Calibration Slope	-	-	-	-	-	-	-	-	No	No
Brautbar et al (2009)	C=0.79 (0.79-0.79)	-	-	-	-	-	-	-	-	-	No	No
Chambless et al (2003)	M= C=0.71 F= C=0.84	-	-	-	-	-	-	-	-	-	No	No
Chen et al (2009)	M= C=0.76 (0.69-0.84) F= C=0.71 (0.62-0.80)	F=7.43, p=0.11 M=2.32,p=0.68	-	-	-	-	-	-	-	-	Discussed	No
Chen et al (2010)	C=0.79 (0.76-0.81)	15.1, p=0.06	74.0	67.7	12.7	C=0.66 (0.60-0.71) & C=0.79 (0.72-0.86)	9.2 p=0.32 & 29.4 p<0.01	-	-	-	Has been made into online tool	No
Chen et al (2017)	C=0.75	p=0.43	63.1	75.9	-	C=0.67	p=0.73	59.2	65.7	-	Discussed	Discussed

Chien et al (2009)	C=0.70 (0.68-0.73)	p=0.874	62	78	-	C=0.66	-	62	78	-	-	-
Chien et al (2012)	C=0.78 (0.74-0.82)	Calibration Plots	-	-	-	-	-	0.71	0.59	-	No	No
Choe et al (2018)	C=0.69	-	0.12	0.98	-	-	-	0.38	0.79	-	No	No
Cross et al (2012)	C=0.65	-	-	-	-	C=0.66	-	-	-	-	No	No
Davies et al (2010)	C=0.61	-	-	-	-	C=0.60	-	-	-	-	No	No
D'Agostino et al (2001)	M= C=0.79 F= C=0.83	-	-	-	-	M= C=0.71 F= C=0.81	M=10.0, p= - F=5.15, p= -	-	-	-	No	No
D'Agostino et al (2008)	M= C=0.76 (0.75-0.78) F= C=0.79 (0.77-0.81)	M=13.48,p=0.14 F=7.79, p=0.56	-	-	-	-	-	-	-	-	-	-
De Bacquer et al (2010)	-	-	-	-	-	C=0.86	8.31 p=0.14	77	72	-	No	No
Dimopoulos et al (2018)	-	-	-	-	-	-	-	0.83	0.24	0.87	No	No
Dugee et al (2015)	C=0.76 (0.70-0.82)	p=0.44	-	-	-	C=0.72	p=0.83	81	59	11	Discussed	No
Dunder et al (2004)	-	-	-	-	-	C=0.66	4.7, p=0.79	-	-	-	No	No
Ferrario et al (2012)	C=0.75	-	-	-	-	C=0.74 (0.68-0.80)	Intercept	-	-	-	No	No
Friedland et al (2009)	C=0.86 (0.83-0.88)	-	-	-	-	C=0.84	-	-	-	-	No	No
Gabriel et al (2015)	M= C=0.79 F= C=0.82	-	-	-	-	-	-	-	-	-	No	No
Gao et al (2009)	-	-	-	-	-	M= C=0.71 (0.66-0.77) F= C=0.71 (0.66-0.77)	-	M=72 F=77	M=47 F=50	-	Discussed	No
Gao et al (2010)	-	Female 10.36 p=0.24 Male 6.23 p=0.62	-	-	-	Female C=0.69 (0.63-0.72) Male C=0.64 (0.59-0.68)	-	Female 80.1 Male 64.1	Female 47.5 Male 56.7	-	Discussed	No
Gaziano et al (2008)	M=C=0.78 (0.77-0.80)	M=6.70, p=0.570 F=6.62, p=0.579	M=85.0 F=76.8	M=57.8 F=73.4	M=47.1 F=44.2	-	-	-	-	-	No	No

	F=C=0.83 (0.81-0.85)											
Glumer et al (2004)	C=0.80 (0.77-0.84)	-	73.3	74.3	11.0	C=0.80	-	75.9	72.2	7.3	No	No
Griffin et al (2000)	-	-	-	-	-	C=0.80	-	77.3	72.0	11.13	Discussed	No
Gupta et al (2019)	-	-	-	-	-	-	-	-	-	-	No	No
Hamer et al (2009)	C=0.78 (0.76-0.80)	-	-	-	-	-	-	-	-	-	No	No
Heianza et al (2013)	C=0.77 (0.76-0.78)	-	72.7	68.1	6.4	C=0.73	-	74.2	70.6	9.6	Discussed	Discussed
Heikes et al (2008)	C=0.85	-	88.2	74.9	13.7	C=0.70	-	77.7	51.4	0.40	In development	Discussed
Hippisley-Cox & Coupland (2017)	-	-	-	-	-	M= C=0.86 (0.85-0.86) F= C=0.88 (0.87-0.88)	Calibration Slopes M=0.99 F=0.99	45.9	90.8		Online tool available	No
Hippisley-Cox et al (2007)	-	-	-	-	-	M=C=0.77 F=C=0.79	Calibration Plots	-	-	-	Online tool available	No
Hippisley-Cox et al (2008)	-	-	-	-	-	M= C=0.79 (0.79-0.79) F= C=0.82 (0.81-0.82)	Calibration Plots	-	-	-	Online tool available	No
Hippisley-Cox et al (2009)	-	-	-	-	-	M= C=0.83 (0.83-0.84) F=C=0.85 (0.85-0.86)	Calibration Plots	-	-	-	Online tool available	No
Hippisley-Cox et al (2010)	-	-	-	-	-	M= C=0.83 (0.83-0.83) F=C=0.84 (0.84-0.84)	Calibration Plots	-	-	-	Online tool available	No
Hippisley-Cox et al (2013)	-	-	-	-	-	M= C=0.87 (0.87-0.87) F= C=0.88 (0.88-0.89)	Calibration Plots	-	-	-	Online tool available	No
Hippisley-Cox et al (2017)	-	-	-	-	-	M= C=0.86 (0.86-0.86) F= C=0.88 (0.88-0.88)	Calibration Plots	-	-	-	Online tool available	No
Hossain et al (2018)	-	-	-	-	-	-	-	-	-	-	No	No

Hwa Ha et al (2018)	M= C=0.71 (0.70-0.73) F= C=0.76 (0.75-0.78)	-	-	-	-	M= C=0.63 (0.53-0.73) F= C=0.66 (0.55-0.76)	Calibration Plots	-	-	-	Discussed	No
Inouye et al (2018)	C=0.70 (0.69-0.70)	-	-	-	-	-	-	-	-	-	No	No
Kang et al (2012)	M= C=0.68 (0.54-0.71) F= C=0.73 (00.67-0.80)	-	-	-	-	-	-	-	-	-	No	No
Katulanda et al (2016)	C=0.78	-	-	-	-	C=0.74	-	77.9	63.0	9.4	Discussed	No
Knuiman et al (1998)	-	-	-	-	-	-	-	-	-	-	No	No
Ko et al (2010)	C=0.74 (0.71-0.77)	-	-	-	40	C=0.68 (0.66-0.70) C=0.77 (0.72-0.82)	-	-	-	49.3 & 21.2	Discussed	No
L'Italien et al (2000)	-	-	-	-	-	-	-	-	-	-	No	No
Laurier et al (1994)	-	-	-	-	-	-	-	-	-	-	No	No
Lees et al (2019)	C=0.74 (0.74-0.75)	-	-	-	-	-	-	-	-	-	No	No
Liao et al (2019)	-	-	-	-	-	C=0.82	-	0.72	0.77	-	No	No
Lindstrom et al (2003)	C=0.85	-	78	77	13	C=0.87	-	81	96	5	Discussed	No
Menotti et al (2005)	-	CITL M=10.8, F=2.2	-	-	-	-	-	-	-	-	No, but software produced	No
Merry et al (2011)	C=0.80 (0.78-0.82)	Calibration Plots	-	-	-	-	-	-	-	-	No	No
Mohammadreza et al (2012)	M= C=0.78 (0.80-0.80) F= C=0.84 (0.82-0.86)	M=13.6, p=0.135 F=11.1, p=0.266	-	-	-	-	-	-	-	-	No	No
Moons et al (2002)	C=0.68 (0.63-0.73)	p<0.50	39	84	-	-	-	-	-	-	No	No
Muehlenbruch et al (2014)	C=0.82 (0.80-0.85)	-	74	78	7	C=0.82	Calibration Plots	87	61	6	No	No
Muehlenbruch et al (2018)	C=0.87 (0.81-0.92)	-	63	90	14	C=0.91 (0.88-0.94)	Calibration Plots	81	84	11	Discussed	No

Nanri et al (2015)	C=0.89 (0.88-0.90)	-	84.2	80.3	16.7	C=0.88 (0.87-0.90)	11.7 p=0.17	82.2	80.0	16.4	Discussed	No
Noda et al (2010)	C=0.77	p=0.94	-	-	-	-	-	-	-	-	No	No
Osborn et al (2015)	-	-	-	-	-	M= C=0.80 (0.76-0.83) F= C=0.79 (0.76-0.82)	Calibration Plots	-	-	-	Discussed	Yes
Panagiotakos et al (2015)	-	-	-	-	-	-	Kendall's Tau M=0.91 F=0.89	70	67	35	Discussed	No
Park et al (2009)	C=0.97	-	-	-	-	-	-	-	-	-	No	No
Paynter et al (2009)	C=0.81 (0.79-0.83)	7.43 p=0.49	-	-	-	-	-	-	-	-	No	No
Paynter et al (2011)	M= C=0.61 F= C=0.70	M= 6.5, p=0.011; F=15.9, p=0.001	-	-	-	-	-	-	-	-	No	No
Pencina et al (2009)	C=0.80 (0.79-0.82)	4.25, p=0.894				C=0.80 (0.77-0.83)	3.98, p=0.913				No, but software produced	
Pocock et al (2001)	-	-	-	-	-	-	-	-	-	-	No	No
Pylypchuk et al (2019)	M= C=0.73 (0.72-0.73) F= C=0.73 (0.72-0.73)	Calibration Plots	-	-	-	M= C=0.72 (0.71-0.72) F= C=0.72 (0.71-0.73)	Calibration Plots	-	-	-	No, but patient example provided	No
Rana et al (2009)	C=0.65 (0.59-0.64)	p=0.009	-	-	-	-	-	-	-	-	No	No
Ridker et al (2007)	-	-	-	-	-	C=0.81	p=0.38	-	-	-	No	No, but clinical example provided
Ridker et al (2008)	C=0.71	15.6 p=0.08	-	-	-	-	-	-	-	-	Online tool	Discussed
Robinson et al (2011)	-	-	-	-	-	C=0.75 (0.73-0.78)	-	70	67	35	Discussed	Discussed
Rosella et al (2011)	M= C=0.80 (0.78-0.83) F= C=0.78 (0.76-0.79)	M=4.22 p>0.1 F=5.22 p>0.1	-	-	-	C=0.77 to 0.80	13.04-18.27, p>0.1	-	-	-	Discussed	No
Sarrafzadegan et al (2017)	C=0.74 (0.72-0.76)	Nam D'Agostino 10.8, p=0.29	-	-	-	C=0.74 (0.70-0.78)	-	-	-	-	Online tool	No
Schulze et al (2007)	-	-	-	-	-	C=0.84	-	83.1	68.3	5.9	Online tool	No

Selmer et al (2017)	M= C=0.79 F= C=0.84	Calibration Plots	64	66	-	M= C=0.84 (0.83-0.85) F= C=0.79 (0.79-0.90)	Calibration Plots	-	-	-	Discussed	No
Solares et al (2019)	C=0.71 (0.69-0.75)	Slope=0.98 (0.84-1.14)	-	-	-	C=0.74 (0.73-0.74)	Slope=0.87 (0.62-1.07)	-	-	-	No	No
Stern et al (2004)	C=0.82	-	75.9	-	-	C=0.77	-	74.4	-	-	No	No
Sun et al (2009)	C=0.85 (0.83-0.87)	-	74.9	79.9	15.4		0.57, p=0.45	72.9	80.7	14.4	Discussed	Discussed
Tabaei et al (2005)	C=0.82	7.98, p=0.44	55	90	65	-	-	53	89	63	Calculator Developed	Discussed
Tanabe et al (2010)	C=0.82	-	-	-	-	-	-	-	-	-	No	No
Tohidi et al (2008)	C=0.80 (0.75-0.85)	-	41.8	92.8	76.5	-	-	-	-	-	No	No
Voss et al (2002)	C=0.90 (0.89-0.91)	-	74.5	97.0	64.0	-	-	-	-	-	No	No
Wen et al (2017)	C=0.72 (0.67-0.73)	-	-	-	-	C=0.70 (0.67-0.74)	p=0.81	74.3	58.8	9.3	Discussed	No
WHO CVD Risk Chart Working Group (2019)	M= C=0.67 (0.66-0.67) F= C=0.76 (0.75-0.77)	Calibration Plots	-	-	-	C=0.69 (0.63-0.74) to C=0.83 (0.78-0.88)	Calibration Plots	-	-	-	No, but risk charts designed	No
Wickramsinghe et al (2014)	C=0.81 (0.80-0.82)	Nam D'Agostino 10.9, p=0.29		-	-	C=0.81 (0.79-0.82)	Nam D'Agostino 6.03, p=0.74	-	-	-	Online tool	No
Wong et al (2015)	C=0.70 (0.66-0.73)	p=0.053	66.2	60.2	-	C=0.71 (0.65-0.77)	p=0.48	72.1	57.8	-	Simple Nomogram Designed	No
Woodward et al (2007)	M= C=0.73 F= C=0.77		46.3	82.5	20.3						No,. but online calculator created	No
Wu et al (2006)	M=0.80 (0.76-0.83) F=0.79 (0.75-0.83)					M= C=0.79 (0.76-0.83) F= C=0.78 (0.74-0.82)	-	-	-	-	No	No
Yatsuya et al (2013)	C=0.74	Grønnesby–Borgan, p=0.62	-	-	-	C=0.69	Grønnesby–Borgan, p=0.17	-	-	-	Discussed	No
Yatsuya et al (2016)	C=0.81 (0.78-0.84)	Grønnesby–Borgan, p=0.74	-	-	-	C=0.77	Grønnesby–Borgan, p=0.74	-	-	-	Discussed	No

Ye et al (2014)	C=0.73 (0.71-0.75)	-	-	-	-	C=0.72	p=0.62	56	75	-	Discussed	No
Zhang et al (2005)	C=0.76	-	-	-	-	C=0.76	p=0.71	-	-	-	No	No
Zhou et al (2013)	C=0.75 (0.74-0.76)	7.0, p=0.54	-	-	-	C=0.72 (0.55-0.78) & C=0.70 (0.68-0.72)	Graphical	92.3 & 86.8	35.5 & 38.8	-	Discussed	Discussed

- indicates no available data

[a]Hosmer-Lemeshow statistics (χ^2 and p-value) presented unless otherwise stated.

Sens = Sensitivity; Spec = Specificity; PPV = Positive Predictive Value; M/F = Indicates sex-stratified analysis (M=Male, F=Female

Appendix D Table 6: TRIPOD Checklist: PsyMetRiC Model Development & Validation

Section/Topic			Checklist Item	Section/ Paragraph
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	Title
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	Abstract
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	Introduction Paragraphs 1-2
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	Introduction Paragraph 3
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	Methods – Data Sources – Paragraph 1-3
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	Methods – Data Sources – Paragraph 1-3
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	Methods – Data Sources – Paragraph 1-3
	5b	D;V	Describe eligibility criteria for participants.	Methods – Data Sources – Paragraph 1-3
	5c	D;V	Give details of treatments received, if relevant.	N/A
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	Methods – Outcome – Paragraph 1
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	N/A (retrospective analysis)
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	Methods – Data Sources – Paragraph 1-3; Methods – Predictor Variables – Paragraph 1
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	N/A (retrospective analysis)
Sample size	8	D;V	Explain how the study size was arrived at.	Methods – Data Sources – Paragraph 1-3
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	Methods – Statistical Analysis – Paragraph 1
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	Methods – Statistical Analysis – Paragraph 1
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	Methods – Statistical Analysis – Paragraph 1
	10c	V	For validation, describe how the predictions were calculated.	Methods – Statistical Analysis – Paragraph 2
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	Methods – Statistical Analysis – Paragraph 2
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	N/A
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	N/A
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	Table 1
Results				

Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time.	Methods – Data Sources – Paragraph 1-3; Table 1
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors).	Table 1
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	Table 1
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	Table 1
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	N/A
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	Table 2
	15b	D	Explain how to the use the prediction model.	Methods – Statistical Analysis – Paragraph 1; Online Data Visualisation App
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	Results – Paragraphs 2-5
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	N/A
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	Discussion – Paragraph 11
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	Discussion – Paragraph 1
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	Discussion – Paragraphs 1-11
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	Discussion – Paragraphs 1-11
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	Results – Paragraph 6; Data Availability Statement
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	Funding Statement

Appendix D Table 7: Missing Sample Analysis: Model Development Sample (CAMEO)

Variable		Included	Missing	Test Statistic
Age, mean (SD)		25.42 (4.77)	28.81 (11.69)	t=5.32, p=0.01
Sex, *n* (%)	Male Female	208 (69.57) 91 (30.43)	490 (60) 324 (40)	χ^2=7.81, p=0.01
Ethnicity, *n* (%)	White Black Asian	250 (83.61) 15 (5.01) 34 (11.37)	676 (83.05) 34 (4.18) 88 (10.81)	χ^2=0.19, p=0.54
Smoking, *n* (%)	Yes No	182 (51.70) 117 (48.30)	443 (54.42) 371 (45.58)	χ^2=0.15, p=0.70
Body Mass Index, mean (SD)		20.53 (8.49)	23.4 (8.80)	t=1.96, p=0.20
Metabolically Active Antipsychotics, *n* (%)	Yes No	216 (72.24) 83 (27.76)	465 (57.13) 349 (42.87)	χ^2=21.04, p=0.01

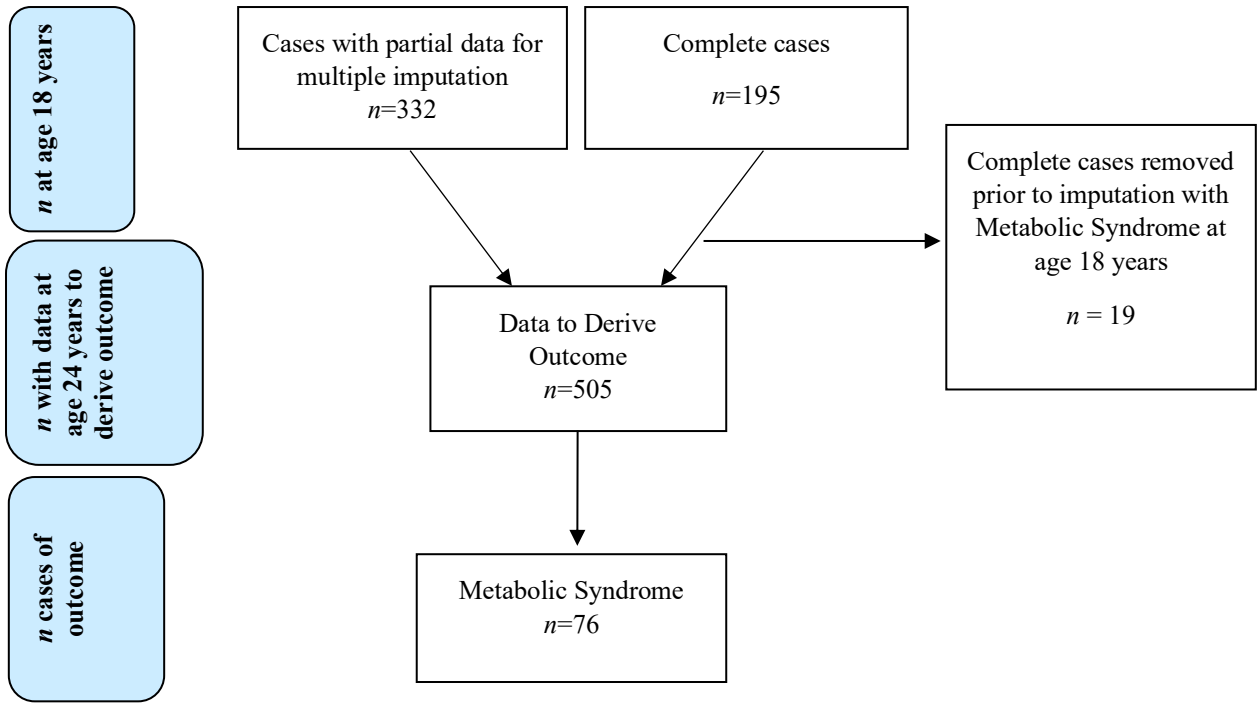
Missing sample analysis was not conducted for the Birmingham sample since there were no participants that were excluded on the basis of missing data on all exposure/outcome variables; cases were excluded only on the basis of having the outcome at baseline.

Appendix D Table 8: Missing Sample Analysis: External Validation Sample (SLAM)

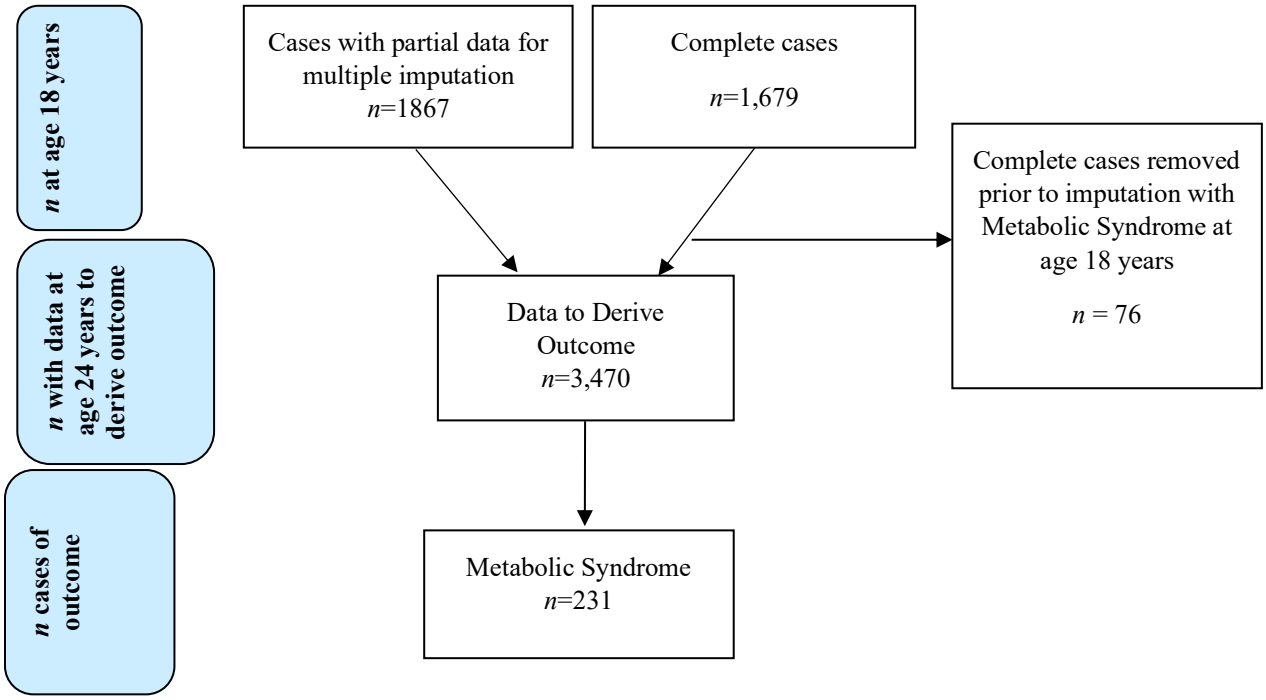
Variable		Included	Missing	Test Statistic
Age, mean (SD)		24.45 (4.75)	29.86 (10.43)	t=18.35, p=0.01
Sex, *n* (%)	Male	440 (67.59)	1472 (59.42)	χ^2=15.46, p=0.01
	Female	211 (32.41)	1002 (40.58)	
Ethnicity, *n* (%)	White	154 (30.20)	1001 (40.46)	χ^2=18.97, p=0.01
	Black	250 (49.02)	1016 (41.07)	
	Asian	106 (20.78)	458 (18.57)	
Smoking, *n* (%)	Yes	469 (91.96)	2029 (81.16)	χ^2=30.81, p=0.01
	No	41 (8.04)	446 (18.84)	
Body Mass Index, mean (SD)		22.96 (6.94)	24.38 (6.72)	t=157.41, p=0.01
Metabolically Active Antipsychotics, *n* (%)	Yes No	472 (92.55) 38 (7.45)	1957 (79.10) 518 (21.90)	χ^2=50.68, p=0.01

Appendix D Figures

Appendix D Figure 1: Flow-Diagram of Participants at Risk of Psychosis at Age 18 or 24 Years Who Were Included in The Exploratory Validation Analysis



Appendix D Figure 2: Flow-Diagram of All Participants Who Were Included in The Exploratory Validation Sensitivity Analysis



Appendix D Published Manuscripts

Acta Psychiatrica Scandinavica

Acta Psychiatr Scand 2020: 142: 215–232
All rights reserved
DOI: 10.1111/acps.13212

ACTA PSYCHIATRICA SCANDINAVICA

Systematic Review or Meta-Analysis

Cardiometabolic risk prediction algorithms for young people with psychosis: a systematic review and exploratory analysis

Perry BI, Upthegrove R, Crawford O, Jang S, Lau E, McGill I, Carver E, Jones PB, Khandaker GM. Cardiometabolic risk prediction algorithms for young people with psychosis: a systematic review and exploratory analysis.

Objective: Cardiometabolic risk prediction algorithms are common in clinical practice. Young people with psychosis are at high risk for developing cardiometabolic disorders. We aimed to examine whether existing cardiometabolic risk prediction algorithms are suitable for young people with psychosis.
Methods: We conducted a systematic review and narrative synthesis of studies reporting the development and validation of cardiometabolic risk prediction algorithms for general or psychiatric populations. Furthermore, we used data from 505 participants with or at risk of psychosis at age 18 years in the ALSPAC birth cohort, to explore the performance of three algorithms (QDiabetes, QRISK3 and PRIMROSE) highlighted as potentially suitable. We repeated analyses after artificially increasing participant age to the mean age of the original algorithm studies to examine the impact of age on predictive performance.
Results: We screened 7820 results, including 110 studies. All algorithms were developed in relatively older participants, and most were at high risk of bias. Three studies (QDiabetes, QRISK3 and PRIMROSE) featured psychiatric predictors. Age was more strongly weighted than other risk factors in each algorithm. In our exploratory analysis, calibration plots for all three algorithms implied a consistent systematic underprediction of cardiometabolic risk in the younger sample. After increasing participant age, calibration plots were markedly improved.
Conclusion: Existing cardiometabolic risk prediction algorithms cannot be recommended for young people with or at risk of psychosis. Existing algorithms may underpredict risk in young people, even in the face of other high-risk features. Recalibration of existing algorithms or a new tailored algorithm for the population is required.

B. I. Perry[1,2], R. Upthegrove[3], O. Crawford[4], S. Jang[4], E. Lau[4], I. McGill[4], E. Carver[4], P. B. Jones[1,2], G. M. Khandaker[1,2]

[1]Department of Psychiatry, University of Cambridge, Cambridge, UK, [2]Cambridgeshire and Peterborough NHS Foundation Trust, Cambridge, UK, [3]Institute for Mental Health, University of Birmingham, Birmingham, UK and [4]University of Cambridge School of Clinical Medicine, Cambridge, UK

Key words: systematic review; cardiometabolic risk; prediction; algorithms; psychosis; ALSPAC

Benjamin I. Perry, Herchel Smith Building, Robinson Way, Cambridge CB2 0SZ, UK. Email: bip20@medschl.cam.ac.uk

Accepted for publication July 6, 2020

Summations

- A large number of cardiometabolic risk prediction algorithms have been developed, but only three algorithms (QRISK3, QDiabetes and PRIMROSE) included psychiatric predictors. All three algorithms were developed and validated in large samples of relatively older participants.
- From our exploratory analysis, we show that all three algorithms may underpredict cardiometabolic risk in young adults with or at risk of developing psychosis, which may be a function of the way age is modelled in the algorithms.
- No existing cardiometabolic risk prediction algorithm can be recommended for use in young adults with or at risk of developing psychosis, yet the population remains at higher risk of cardiometabolic disorders than their age-matched peers.

Limitations

- Due to study heterogeneity, we were unable to follow a meta-analytic approach to the synthesis of systematic review results.
- Our exploratory analysis of QRISK3, QDiabetes and PRIMROSE was limited by relatively small sample size and a related but distinct outcome definition to the original algorithms.

Introduction

Cardiometabolic disorders broadly include cardiovascular diseases (CVD), disorders of adiposity such as obesity and disorders of glucose-insulin homeostasis such as type 2 diabetes mellitus (T2DM) (1). They impose a huge societal burden costing an estimated £30 billion and accounting for over 190 000 deaths each year in the UK alone (2). A particularly high-risk group for the development of cardiometabolic disorders are people with psychotic disorders such as schizophrenia, who make up around 0.8% of the population (3) and have up to a 30% increased incidence of cardiometabolic disorders than the general population (4). Indeed, increased physical comorbidity is a leading cause for significantly increased mortality rates and reduced life expectancy for people with schizophrenia compared with the general population (5–7). We therefore need clinical tools to predict cardiometabolic risk in this group in order to optimize care and improve long-term outcomes. Yet, a recent report of a small sample of people with chronic schizophrenia suggests that some commonly used cardiometabolic risk prediction algorithms return differing risk prediction scores when tested on the same participants. This calls into question the reliability and suitability of such algorithms for relatively older people with chronic schizophrenia, let alone young people with or at risk of psychosis (8).

Recent evidence suggests that the physical comorbidity associated with schizophrenia starts early. Markers of developing cardiometabolic disorders are a feature that distinguish cases of first-episode psychosis from matched general population controls (9, 10) and are associated with young adults at risk of developing psychosis (11). The field of early intervention in psychosis rests on a premise that intervening early could improve longer-term outcomes, and this premise applies equally to the treatment of cardiometabolic disorders. Therefore, cardiometabolic risk prediction algorithms may be a useful tool for healthcare professionals to help tailor treatment plans for young people with psychosis that could help to reduce both long-term physical and psychiatric morbidity. However, such a tool could only be clinically useful if the predictions it makes are accurate. It is unclear as to whether this may or may not be the case.

Aims of the study

We conducted a systematic review to identify and compare existing cardiometabolic risk prediction algorithms developed for the general or psychiatric populations and consider their suitability for young people with psychosis. Next, we performed an exploratory analysis using data from a large UK birth cohort to examine the predictive ability of any algorithms highlighted as potentially suitable by the review, in a sample of young adults with or at risk of psychosis. To explore the impact of age on risk estimates, we reassessed model performance after artificially increasing the age of participants to the mean age of the original algorithm development study, leaving all other predictors unchanged.

Method

Systematic review

Literature search. We conducted a systematic literature search of EMBASE (1947-present), Ovid MEDLINE (1946-present), PsychINFO (1806-present), Web of Science (from inception) and the first twenty pages of Google Scholar (12) to 1 December 2019. We also searched the references of included studies. Our search strategy is presented below. MeSH headings (denoted with *) and text terms were used:

Group 1: metabolism* (OR) metabolic* (OR) diabetes mellitus* (OR) cardiovascular diseases* (OR) obesity* (OR) cardiometabolic

(AND)

Group 2: risk assessment* (OR) risk* (OR) outcome assessment* (OR) patient outcome assessment* (OR) prognosis*

(AND)

Group 3: calculator (OR) computers* (OR) algorithms* (OR) software* (OR) tool.

We applied the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-analyses) guidelines (13). The systematic review was registered on PROSPERO (CRD42019150377).

Study selection. The inclusion criteria were as follows: (i) studies reporting the development and/or validation of cardiometabolic risk algorithms designed for either the general or psychiatric populations; (ii) studies which reported in combination the development and validation (internal or external) of an original algorithm; reported the development but not validation of an algorithm; reported the first validation of a previously developed but not validated algorithm; or reported a new recalibration of a previously developed algorithm; (iii) cardiometabolic risk was defined as CVD (stroke, myocardial infarction, hypertension, unstable angina) and its common predeterminants including T2DM, prediabetes, obesity or dyslipidaemia; (iv) studies reported in any language; (v) published and unpublished research, conference proceedings and academic theses. The exclusion criteria were as follows: (i) algorithms designed specifically for other defined health groups (i.e. postoperative patients or patients with any physical health diagnoses at baseline) and (ii) studies reporting validation without recalibration of previously validated algorithms.

Titles and abstracts were screened independently by four authors (BIP, EL, IM and EC) prior to full-text screening. Any discrepancies were resolved in consultation with a senior author (GMK). Data were extracted by three authors (BIP, OC and SJ) from studies that met the inclusion criteria. Searches were re-run immediately prior to the final analyses, and further studies retrieved for inclusion using the processes outlined above.

Data extraction and synthesis. We extracted data on general characteristics (e.g. population, location, study type, type of risk predicted), the characteristics of included participants (e.g. age, sex, ethnicity) and characteristics of the developed/validated algorithms (e.g. included predictors, algorithm performance). Risk of bias was assessed using the 'Prediction model Risk Of Bias Assessment Tool' (PROBAST) (14), which aims to identify shortcomings in study design, conduct or analysis that could lead to systematically distorted estimates of model predictive performance. PROBAST includes four domains for potential sources of bias in prediction model studies (participants, predictors, outcome and analysis) which are then summarized by an overall judgement, either low risk, high risk or unclear risk of bias (14). We plotted the range and frequency of predictors included in studies. We illustrated the relative weighting of different predictors in one included study which featured psychiatric predictors. Algorithm performance was compared using statistics relating to model discrimination (how well an algorithm discriminates people at higher risk from people at lower risk, e.g. Harrell's C Statistic, where a score of 1.0 indicates perfect discrimination, and a score of 0.5 indicates the model is no better than chance) and model calibration (the accuracy of absolute risk estimates, e.g. calibration plots) (15). We also examined the events-per-variable ratio (EPV) (the ratio of outcome events: predictors considered in algorithm development) of each study to assess the potential risk of model overfitting (16). An EPV of 10 or more had previously been considered satisfactory (17), though higher ratios have more recently been advised (18). Where an EPV ratio was not reported, we calculated it where possible from the information available in the study. Finally, we considered the likely suitability of all included algorithms for young people with psychosis. We summarized and compared studies with a narrative synthesis (19).

Exploratory analysis

Data source. The Avon Longitudinal Study of Parents And Children (ALSPAC) birth cohort initially recruited 14 541 pregnant women resident in a geographically defined region in southwest of England, with expected dates of delivery 1 April 1991 to 31 December 1992, resulting in 14 062 live births (20–22). Following further periods of recruitment over time, 913 additional participants were recruited. See http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/ for a fully searchable data dictionary. Study data were collected and managed using REDCap electronic data capture tools hosted at University of Bristol (23,24). Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and Local Research Ethics Committees. All participants provided informed consent.

Study sample. We included participants who at age 18 or 24 years were identified as experiencing definite psychotic symptoms or psychotic disorder. In ALSPAC, psychotic symptoms were identified through the face-to-face, semi-structured Psychosis-Like Symptom Interview (PLIKS) conducted by trained psychology graduates and coded according to the definitions in the Schedules for Clinical Assessment in Neuropsychiatry, version

2.0. See Supplementary Data for further information. We excluded participants who already met the outcome criteria at age 18 years and participants who had missing data on all included variables. Additionally, we conducted a *post hoc* sensitivity analysis to examine the potential impact of sample size; we performed the analysis again including all participants from the total ALSPAC sample at age 18 years who did not meet the criteria for the outcome at age 18 years and who did not have missing data on all included variables. See Figures S1–S2 for flow charts of included participants.

Outcome. We used the harmonized definition (25) of the metabolic syndrome measured at age 24y as the outcome, in which it is an established precursor of T2DM (26) and CVD (27). Metabolic syndrome is a more appropriate outcome for a sample of relatively young participants. The follow-up period was six years. The binary outcome was coded present for participants meeting ≥3 factors from the following: ethnicity-specific waist circumference (≥94 cm in males and ≥80 cm in females for Caucasians; ≥90 cm in males and ≥80 cm in females for other ethnic groups (25)); elevated triglycerides (≥1.7 mmol/L); reduced high-density lipoprotein (HDL (<1.0 mmol/L in males and <1.3 mmol/L in females); elevated seated blood pressure (systolic ≥ 130 mmHg); and elevated fasting plasma glucose (FPG) (≥5.7 mmol/L). See Supplementary Methods for further detail on biochemical measurements.

Predictors. We included all available predictors from QRISK3, QDiabetes and PRIMROSE, which were the three algorithms highlighted as being potentially the most suitable for young people with psychosis. These included age, Townsend deprivation score, body mass index (BMI), ethnicity, smoking, antipsychotic medication use, antidepressant use, corticosteroid use, psychosis, depression, family history of cardiovascular disease or type 2 diabetes, hypertension, FPG, cholesterol:HDL ratio, systolic blood pressure, total cholesterol, HDL, alcohol intake and year of assessment. For a full list of predictors for each algorithm and details on how they were measured, see Table S1 and Methods S1.

Statistical analysis. Estimated six-year risk estimates for metabolic syndrome were calculated for QDiabetes (28), QRISK3 (29) and PRIMROSE (30), by applying the published fully specified algorithms to our sample. QDiabetes and PRIMROSE comprise different models depending on the availability of blood test results; thus, we used the model which performed best in the original model development studies (28, 30). For QDiabetes, the best performing model included FPG; for PRIMROSE, the best performing model included lipids. QDiabetes and QRISK3 estimate risk separately for males and females. We used multiple imputation using chained equations (31) to address the impact of missing predictor data. See Methods S1 for further details. Algorithm performance was assessed using measures of discrimination (Harrell's C statistic and R^2) and a measure of calibration (calibration plots). Calibration plots included grouped observations, which were split at each 0.2 of predicted risk. First, we calculated model performance using actual participant age (18 years). To assess the impact of age on model performance, we artificially substituted every participants' age in ALSPAC to the mean age from the original algorithm development study (QDiabetes = 44.9 years; QRISK3 = 42.9 years; and PRIMROSE = 49.5 years), leaving all other predictors unchanged. We re-ran each algorithm and compared the model performance statistics described above. Statistical analysis was carried out in R version 3.6.0 (32).

Results

Systematic review

Study selection and quality assessment. The literature search returned 7744 results after removing duplicates. We reviewed 362 full texts, of which 110 studies met inclusion criteria (28-30, 33-138). See Fig. 1 for the PRISMA diagram. Three studies were not contained within peer-reviewed journals and were published either as conference proceedings (108), a thesis (93) or a preprint (106). Reporting quality was relatively poor across the majority of studies, with 108 studies (98%) either at unclear or high risk of bias following assessment with the PROBAST tool (14). See Table S2.

Study characteristics. Table S3 reports the characteristics of included studies. All studies were conducted on general population samples of healthy adults, except one which was conducted on patients with severe mental illness, defined as either schizophrenia, other psychotic disorder or bipolar disorder (30). The majority of included studies were conducted in high-income or upper-middle-income countries, with the UK, USA and China best represented. Eleven studies were conducted in lower- or middle-income countries. Sample sizes were highly variable in both development (n = 100

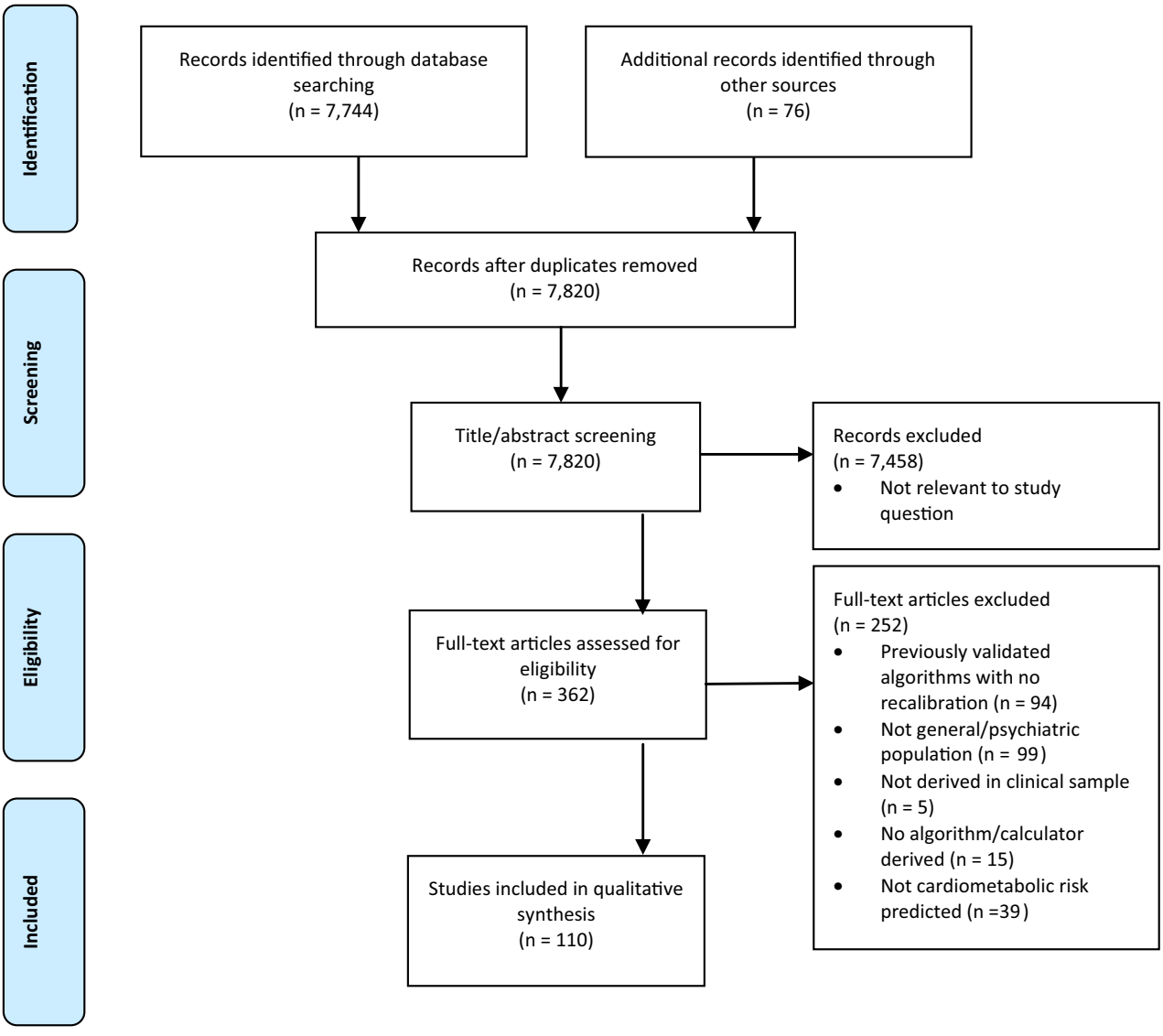


Fig. 1. PRISMA diagram. [Colour figure can be viewed at wileyonlinelibrary.com]

participants (120) to $n = 8\ 136\ 705$ participants (28)) and validation cohorts ($n = 90$ participants (104) to $n = 2\ 671\ 298$ participants (29)). Sixty-one studies (55%) assessed the risk of fatal or non-fatal CVD; 31 studies (28%) assessed the risk of T2DM; five studies (5%) assessed the risk of either prediabetes or T2DM; three studies (3%) assessed the risk of metabolic syndrome or obesity; and three studies (3%) assessed the risk of stroke or transient ischaemic attack.

Lengths of predicted risks ranged from one (119) to 30 (80, 123) years. The most common risk prediction timeframes were either ten-year risk (38 studies, 35%) or five-year risk (14 studies, 13%). Thirty-nine studies (35%) performed external validation of an original algorithm. Forty studies (36%) performed internal validation, either by subsetting the initial cohort or by bootstrap methods. All algorithms were designed using either Cox proportional hazards or logistic regression analysis. Most studies selected variables for inclusion from previous research or clinical importance (50 studies, 45%), or using statistical methods, that is forward or backward selection (31 studies, 28%). Seventeen studies (15%) used simple univariable analysis of each considered predictor, which is least preferable since it cannot assess interactions between two or more variables. Eleven studies (10%) used machine learning techniques.

Participant characteristics. All studies were conducted in adults. The mean age of participants based on the 76 studies that reported mean age was 50.50 years (SD 9.31). No studies included a mean age of participants below 35 years. Eighty-nine studies (81%) reported the sex distribution of

the derivation cohort (mean 55.29% male (SD 17.27)), and 42 studies (38%) reported for the validation cohort (mean 52.25% male (SD 14.44)). The majority of studies included roughly equal sex distribution, apart from nine studies which included only (121, 127) or mostly females (82, 83, 85, 98, 120, 122, 128) and 12 studies which included only (41, 71, 94, 102,103,112,119,132,136) or mostly males (69,80,81). Thirty-three studies (30%) reported the ethnic makeup of their sample, where samples ranged from being ethnically completely homogenous in 18 studies (16%) to relatively heterogeneous, with less than 66% of participants falling into the most common ethnic group (63,72,84,125). See Table S3.

Algorithm characteristics. *Predictors included in existing algorithms*. Figure 2 shows the frequency of different predictors included in studies. The most common predictors were age (98 studies, 89%), smoking (83 studies, 75%) and systolic blood pressure (55 studies, 50%). Inflammatory markers such as CRP or IL-6 were included as predictors in 15 studies (14%). The number of predictors considered for each algorithm varied between four (44, 52, 53, 79) and 473 predictors (86). EPV varied between 2.1 (55) and 5,075.4 (29). Twenty studies featured EPV ratios that were likely < 10. See Table S4.

Performance of existing algorithms. Discrimination statistics were presented in 93 studies (85%), and calibration statistics were presented in 62 studies (56%). From the 80 studies that included both model development and validation analysis, 35 (44%) reported performance statistics from both development and validation cohorts, 27 (34%) reported only validation cohort statistics, and ten (13%) reported development only statistics. Most commonly overall, studies reported both discrimination and calibration statistics (35 studies, 32%). Next most commonly, studies reported measures for discrimination, calibration and sensitivity/specificity (23 studies, 21%). Eleven studies (10%) reported no model performance statistics. Discrimination was mostly assessed with area under the curve (AUC/Harrell's C statistics). AUC ranged between 0.61 (100) and 0.97 (120) though notably the latter was at risk of model overfit, with a sample size of $n = 100$ and an EPV ratio of 3.1. The

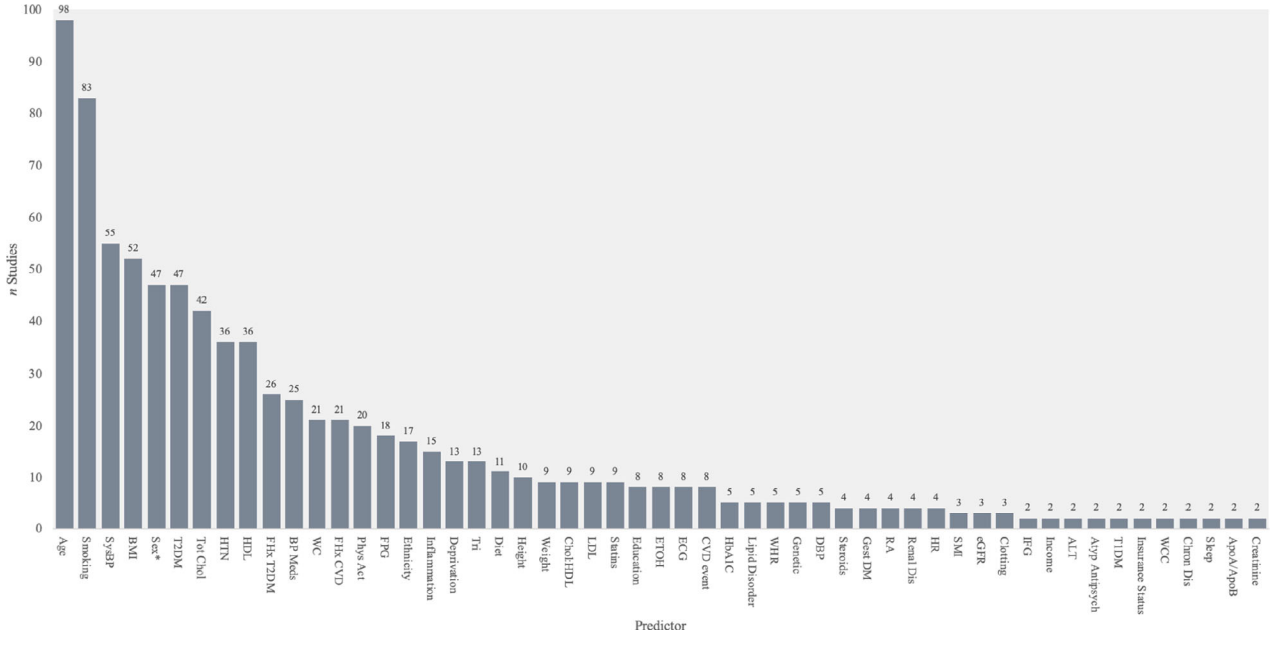


Fig. 2. Range and frequency of different predictors used in current algorithms. ALT, Alanine Aminotransferase; ApoA/ApoB, Apolipoprotein A/B Levels; Atyp Antipsych, Prescribed Antipsychotic Medication; BP Meds, Prescribed Antihypertensive Medication; Chol:HDL, Cholesterol:HDL Ratio; Chron Dis, Personal History of Chronic Disease; CVD Event, Personal History of Cardiovascular Diseases; DBP, Diastolic Blood Pressure; ECG, Electrocardiogram Findings; eGFR, Glomerular Filtration Rate; ETOH, Alcohol Use; FHx CVD, Family History Cardiovascular Diseases; FHx T2DM, Family history of Type 2 Diabetes Mellitus; FPG, Fasting Plasma Glucose; Genetic, Genotype Data; Gest DM, Gestational Diabetes Mellitus; HbA1C, Glycated Haemoglobin; HDL, High-Density Lipoprotein; HR, Heart Rate; HTN, Diagnosis of Hypertension; IFG, Impaired Fasting Glucose; LDL, Low-Density Lipoprotein; Phys Act, Physical Activity; RA, Rheumatoid Arthritis; Renal Dis, Renal Disorders; SMI, Diagnosis of Serious Mental Illness; SysBP, Systolic Blood Pressure; T1DM, Type 1 Diabetes Mellitus; T2DM, Type 2 Diabetes Mellitus; Tot Chol, Total Cholesterol; Tri, Triglycerides; WC, Waist Circumference; WCC, White Cell Count; WHR, Waist:Hip Ratio. *not counted as a predictor in studies that developed sex-specific algorithms. [Colour figure can be viewed at wileyonlinelibrary.com]

mean AUC across all included studies was 0.77, with 54 studies (49%) scoring above 0.75, suggestive of 'good' discrimination. The majority of studies that reported calibration statistics used the Hosmer–Lemeshow goodness-of-fit chi-squared test. Seventeen studies (15%) used the preferred (139) method of calibration plots. See Table S5.

Potential applicability of existing cardiometabolic risk algorithms for young people with psychosis. Psychiatric illness and treatment were taken into account in three studies (28–30) predicting risk of CVD (29, 30) or T2DM (28). Two of these studies (QRISK3 and QDiabetes (28, 29)) were conducted on large general population samples, and one (PRIMROSE) was conducted in people with severe mental illness (30). QRISK3 and QDiabetes (28, 29) included diagnosis of severe mental illness as a single predictor, whereas PRIMROSE included separate predictors for bipolar disorder and psychosis (30). QRISK3 and QDiabetes included the presence of any atypical antipsychotic as a predictor (28, 29); PRIMOSE included first- or second-generation antipsychotics as separate predictors, along with antidepressants as another predictor (30). All three studies were conducted on middle-aged adults (mean ages QDiabetes: 42.9 years (28), QRISK3: 44.9 years (29), PRIMROSE: 49.5 years (30)). In PRIMROSE, age was applied as a non-linear term with a log transformation and was weighted heavily in comparison to other risk factors. See Figure S3. In both QRISK3 and QDiabetes, age was applied as a fractional polynomial, also implying a non-linear impact on risk. QRISK3 and QDiabetes both included a number of interactions between age and other predictors, further amplifying the relative importance of age in the algorithms.

QRISK3, QDiabetes and PRIMROSE were taken forward for the exploratory analysis, on the basis of the following: large samples used in development and validation; strong performance statistics; low risk of bias in three domains; and inclusion of psychiatric predictors/development in a psychiatric sample.

Exploratory analysis

Baseline characteristics. The six-year observed risk of metabolic syndrome at age 24 years in our sample of participants with or at risk of psychosis was 14.21% in females and 11.88% in males. In our sensitivity analysis (all available ALSPAC participants), the six-year observed risk was 7.54% for females and 5.76% for males. In our primary analysis, we included 3030 person-years of observation. In our sensitivity analysis, we included 19 020 person-years of observation. Characteristics of included participants for both the primary and sensitivity analyses are presented in Table 1 and Table S6 respectively. Associations between algorithm predictors and outcome are reported in Table S7.

Primary analysis – psychosis sample. Discrimination. At age 18 years, Harrell's C Statistics were as follows: QDiabetes males $C = 0.75$ (95% CI, 0.72–0.78) and females $C = 0.78$ (95% CI, 0.73–0.84); QRISK3 males $C = 0.58$ (95% CI, 0.52–0.65) and females $C = 0.61$ (95% CI, 0.55–0.66); and PRIMROSE $C = 0.73$ (95% CI, 0.70–0.78). After substituting participant ages to the mean age of the original studies, Harrell's C statistics mildly improved for each algorithm. Similarly, at age 18 years, R^2 statistics were marginally higher in females than males in QDiabetes and QRISK3 and improved mildly after substituting participant ages to the mean age of the original studies. See Table 2.

Calibration. At age 18 years, calibration was poor across all three algorithms, with observed risk estimates consistently higher than predicted risk estimates. After substituting participant ages to the mean age of the original studies, calibration improved markedly in all three algorithms. See Figure 3.

Sensitivity analysis – whole ALSPAC sample. Discrimination. QDiabetes and QRISK3 performed better in the overall sample than the psychosis sample. PRIMROSE performed better in the psychosis sample. Harrell's C Statistics were as

Table 1. Characteristics of ALSPAC participants with or at risk of psychosis included in exploratory analysis

Characteristic (*N*, % unless stated)	Females	Males
Number of participants	323 (63.9)	182 (36.1)
Total person-years of observation	1938	1092
Ethnicity – White/Not-recorded	315 (97.5)	176 (96.7)
Systolic BP (mmHG), Mean (SD)	109.88 (8.28)	118.90 (9.67)
HDL (mmol/L), Mean (SD)	1.29 (0.36)	1.18 (0.33)
FPG (mmol/L), Mean (SD)	4.88 (0.36)	5.19 (0.66)
Total cholesterol (mmol/L), Mean (SD)	3.86 (0.68)	3.55 (0.63)
Chol:HDL ratio, ratio SD	3.04 (0.85)	3.08 (0.85)
BMI (kg/m^2), Mean (SD)	23.75 (3.55)	23.62 (4.50)
Family history cardiometabolic/Cardiovascular disorders	194 (60.1)	117 (64.3)
Smoking (≥1 cigarette daily)	173 (53.6)	100 (54.9)
Depression	90 (27.9)	28 (15.4)
Alcohol use	47 (15.4)	31 (16.7)
Antidepressant medication	45 (14.7)	16 (8.6)
Antipsychotic medication	12 (3.7)	6 (2.1)

Table 2. Discrimination statistics for algorithms tested on psychosis risk group at age 18 years and mean age of original study

	Harrell's C statistic (95% CI); R^2 statistic			
	Age 18 years		Mean age original study	
Algorithm	Male	Female	Male	Female
QDiabetes FPG	C = 0.70 (0.65–0.74) R^2 = 0.13 (0.09–0.19)	C = 0.78 (0.73–0.84) R^2 = 0.16 (0.10–0.24)	C = 0.78 (0.75–0.80) R^2 = 0.21 (0.14–0.27)	C = 0.83 (0.80–0.87) R^2 = 0.25 (0.19–0.31)
QRISK3	C = 0.58 (0.52–0.65) R^2 = 0.09 (0.05–0.16)	C = 0.61 (0.55–0.66) R^2 = 0.10 (0.03–0.18)	C = 0.63 (0.58–0.69) R^2 = 0.11 (0.07–0.16)	C = 0.66 (0.59–0.72) R^2 = 0.13 (0.05–0.20)
PRIMROSE Lipid	0.73 (0.70–0.78) R^2 = 0.13 (0.10–0.0.17)		0.75 (0.69–0.79) R^2 = 0.16 (0.12–0.22)	

follows: QDiabetes males C = 0.72 (95% C.I., 0.70–0.73) and females C = 0.82 (95% CI, 0.79–0.84); QRISK3 males C = 0.64 (95% CI, 0.62–0.66) and females C = 0.62 (95% CI, 0.59–0.65); and PRIMROSE C = 0.68 (95% CI, 0.67–0.70). Similarly, at age 18 years, R^2 statistics were marginally higher in females than males in QDiabetes, but marginally higher in males in QRISK3. After substituting age to the mean age of the original studies, Harrell's C statistics and R^2 improved in all three algorithms. See Table S8.

Calibration. In a similar pattern to the psychosis sample, at age 18 years, calibration was poor across all three algorithms, with observed risk estimates consistently higher than predicted risk estimates. After substituting participant ages to the mean age of the original studies, calibration improved markedly in all three algorithms. See Figure S4.

Discussion

Main findings

We performed a systematic review of cardiometabolic risk prediction algorithms developed either for the general or psychiatric populations and considered their potential suitability for young people with psychosis. We also used data from a sample of relatively young adults to first explore whether existing cardiometabolic risk prediction algorithms may be suitable for young people with or at risk of psychosis and second to explore the impact of the manner in which age is weighted in existing cardiometabolic risk prediction algorithms. We do not present the results of our exploratory analysis as an *external validation* of the three algorithms, since the algorithms we tested were not developed to predict metabolic syndrome. Rather, we present our findings as a means to explore the likely suitability of these algorithms for a population of individuals who may be at higher cardiometabolic risk compared with the general population. It should be made clear from the outset that the three algorithms we tested, as we show in the results of our systematic review, were developed and validated on large samples and perform well in the populations they were designed for.

Systematic review

We identified a substantial number of cardiometabolic risk prediction algorithms, yet most have not been integrated into clinical practice. Predicted outcomes ranged from prediabetes and T2DM, CVD or transient ischaemic attack and stroke. The five most commonly included predictors across all algorithms were age, smoking, systolic blood pressure, sex and BMI. One included algorithm (PRIMROSE) was developed in a population of people with severe mental illness (30), which predicted risk of CVD. Two (QRISK3 and QDiabetes) were developed in the general population and included psychiatric predictors (28, 29) such as a diagnosis of schizophrenia.

All included algorithms were developed in samples of middle- to older-age adults. One might traditionally consider this proportionate, since cardiometabolic disorders are traditionally considered diseases of advancing age. Yet, cardiometabolic risk still exists in the absence of advancing age; even in the general population, there is an increasing prevalence of early-onset T2DM (140) and childhood obesity (141), likely related to the shift towards a more sedentary lifestyle and unhealthy diet in recent decades. The absence of an algorithm developed for younger populations is an important finding, since early intervention may reduce the risk of young people forming part of a future generation of patients with chronic cardiovascular diseases (142). This finding suggests the need for either new or recalibrated versions of currently existing cardiometabolic risk algorithms tailored to the younger generations.

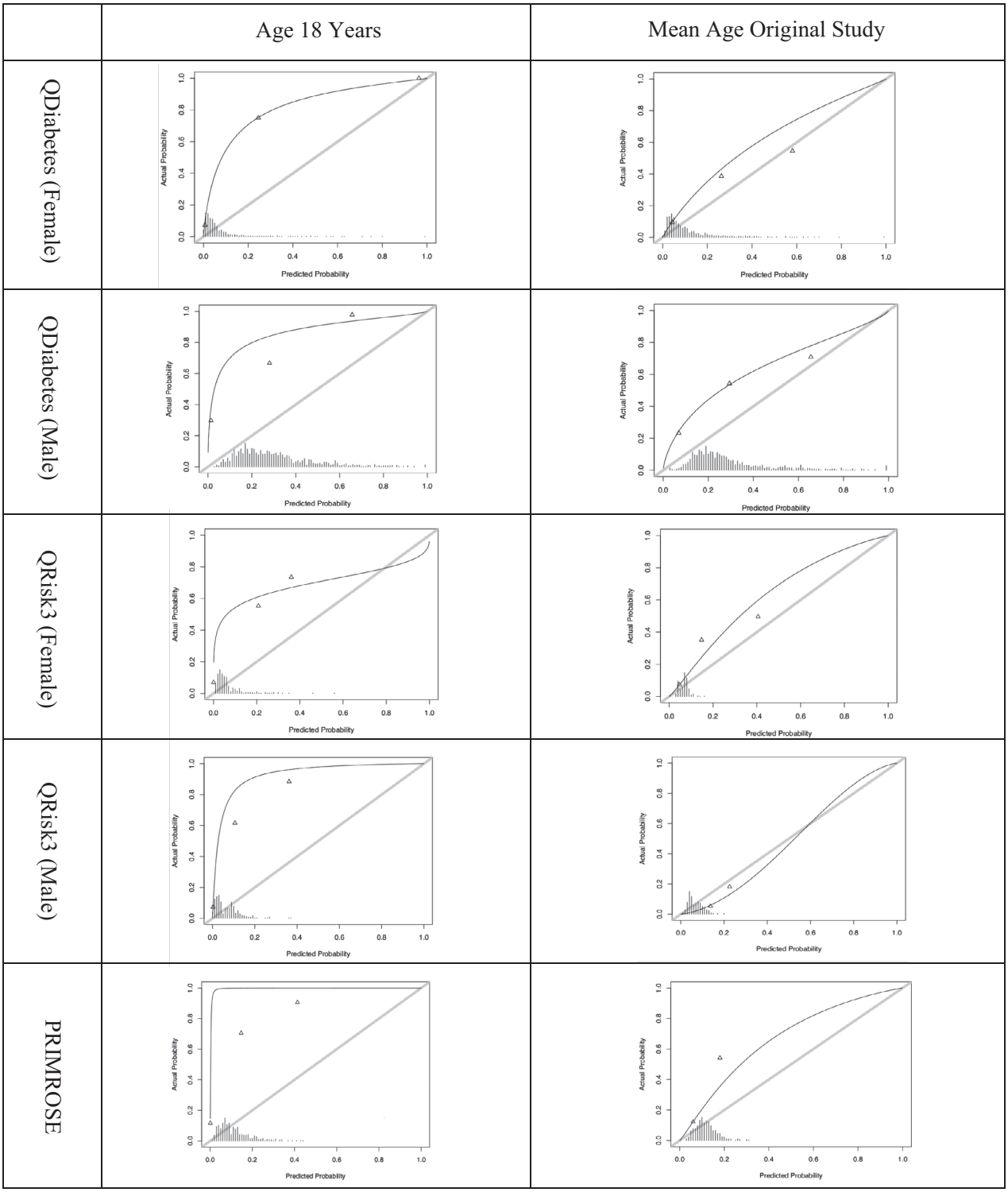


Fig. 3. Calibration plots of algorithms tested on ALSPAC psychosis risk group at age 18 years and at mean age of original study. Perfect calibration (dark grey) would follow the diagonal (light grey) line, indicating perfect agreement between observed/expected risk. Grouped observations were split at each 0.2 of predicted risk.

Primary prevention is the best means with which to address the personal and societal burden attributed to T2DM, CVD and its complications (143). Whilst this message is important for the general population, it is particularly important for young people with/at risk of psychosis, who are at a

higher risk of precipitant cardiometabolic disorders. This population may be more likely to smoke (144), exercise less (145) and eat a more unhealthy diet (145) than their peers and yet may also be prescribed medication that in itself can adversely and severely impact cardiometabolic indices (146). Further, they may be faced with inappropriate barriers to accessing healthcare (147), diagnostic overshadowing (148) and may have an intrinsic biological propensity for altered cardiometabolic function (149). Meta-analyses featuring mostly antipsychotic-naïve young people with first-episode psychosis have consistently reported an increased incidence of insulin resistance, impaired glucose tolerance (9, 10) and dyslipidaemia (9, 150, 151) compared with matched controls from the general population, after adjusting for anthropometric and sociodemographic factors. Each is predeterminants of cardiometabolic disorders such as T2DM and obesity. These factors may not be adequately captured by currently existing algorithms. Additionally, meta-analyses of cross-sectional studies suggest that psychosis is associated with higher levels of circulating inflammatory markers (152–155), and evidence from some longitudinal studies suggests an association between inflammatory markers at baseline and psychosis at follow-up (156–158), although other longitudinal studies have reported negative findings (159). Inflammatory states are also associated with cardiometabolic disorders (160–163). Whilst 15 relatively newer algorithms from our systematic review did include inflammatory predictors, none also included psychiatric predictors.

Each of the three algorithms that did include psychiatric factors featured an antipsychotic-related predictor. Antipsychotic-associated weight gain can occur relatively quickly after initiation (164) and is associated with altered eating behaviours (165) and sedentariness (166). However, whilst there are some efficacy differences between antipsychotics, these are gradual rather than discrete (167). Differences in side-effects are more marked, and each has an inherently different impact upon cardiometabolic risk (168). This may be explained by differing affinities to receptors other than the dopamine-2 (D2) receptor, for example the histamine-1 (H1) receptor, serotonin-2c (5-HT2c) and adrenergic receptors (a2 and b3) (169), which may have a role in the regulation of food intake (170). The varied impact upon cardiometabolic risk by different antipsychotics does not abide by the traditional distinctions of either typical/atypical or first/second generation, which were the binary distinctions of the included algorithms. A more appropriate antipsychotic predictor may instead model antipsychotics based on their relative cardiometabolic risk.

We used the PROBAST tool (14) to examine the risk of bias of included studies in our systematic review. Only two studies were rated as low risk of bias, with all others rated as either unclear or high risk of bias. This may be a reflection of the relatively recent introduction of the 'Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis' (TRIPOD) guidelines for prediction model studies (139). Nevertheless, the results suggest that the reported performance statistics and therefore clinical validity of the majority of included studies should be accepted with extreme caution.

The EPV ratio also varied widely between studies. A low EPV ratio can be an indicator of model overfit (17) which can bias results. We identified 20 studies with an EPV ratio of likely < 10, and therefore, the performance reported in those studies should be interpreted with caution. Finally, it is striking that whilst many included studies promoted the use of their algorithms in clinical practice, there appears to have been relatively little follow-up to assess either clinical or economic impact. A notable exception was PRIMROSE (30), which was the only algorithm developed and validated on a sample of people with mental illness. A cost-effectiveness analysis (171) found it improved quality of life and reduced healthcare-related costs in comparison with using no algorithm.

A previously published systematic review (172) examining cardiovascular risk prediction algorithms in the general population also identified a very large number of studies. The review similarly concluded the methodological shortcomings of most risk prediction algorithms likely limit their suitability for clinical practice. The previous review differs from our own since we were interested in identifying original or recalibrated algorithms and assessing their suitability for young people with psychosis. Therefore, we did not include studies reporting new validations in a similar population to already validated algorithms. The previous review also presented sex-stratified algorithms as distinct entities, increasing the apparent number of algorithms they reported. For ease of simplicity and in consideration of our overarching research question, we did not take this step. Finally, a large number of new algorithms have been developed since the previous review, which we were able to include in our own.

Exploratory analysis

We considered three algorithms for exploratory analysis: QRISK3, QDiabetes and PRIMROSE.

These were selected due to the large sample sizes in model development and validation, model performance statistics, relatively low risk of bias and the inclusion of psychiatric predictors/development in a psychiatric population.

We found that discrimination statistics were relatively good at age 18 years for QDiabetes and PRIMROSE and improved further when substituting to the mean age of original studies. This means that QDiabetes and PRIMROSE were able to predict higher risks in 'cases' than 'non-cases', even in relatively young adults. This did not apply to QRISK3, particularly in males, where the algorithm was little better than chance at discriminating higher and lower cardiometabolic risk in young adults with or at risk of psychosis.

For all three algorithms, however, the discriminative ability in our sample was attenuated compared with the original published studies (28–30). This may be because our sample included younger participants than the original studies. For example, both QRISK3 and QDiabetes were developed and validated in participants aged 25 and over, and PRIMROSE was developed and validated in participants aged 30 and over. QRISK3 and QDiabetes set a minimum age of 25 when using their online calculators, although PRIMROSE sets a minimum of age 18 years. Additionally, in our primary analysis, we tested a sample of participants with or at risk of psychosis, whereas QDiabetes and QRISK3 were designed for use in the general population. Furthermore, we tested a different outcome compared with the original algorithms. We tested metabolic syndrome since it is an established precursor of both T2DM and CVD (26, 27) and is a more suitable outcome for younger populations. The improvement in discrimination statistics after substituting age provides some face validity to our choice of outcome.

However, discriminative ability is only half the story, since discrimination statistics cannot assess the accuracy of the amount of risk apportioned by a model; this represents a test of *absolute risk* estimates and is examined with a measure of calibration. Our calibration plots at 18 years showed that observed risk was systematically greater than predicted risk in all models, suggesting a notable underprediction of risk in younger participants. Calibration plots improved markedly in all algorithms when we artificially substituted age to the mean age of the original studies. This suggests that the manner with which age is modelled in current algorithms is a major limiting factor in applying them to younger populations. This is likely because many cardiometabolic risk factors are cumulative over time (173); thus, age becomes increasingly important with regard to cardiometabolic risk as one gets older. This notion is elegantly painted by all three algorithms, which modelled age as either a non-linear function, included interactions between age and other predictors, or both.

Strengths and limitations

Strengths of this systematic review include following PRISMA reporting guidelines (13), as would be expected for a high-quality review. Alongside the review, we were able to complement our findings with an exploratory analysis using data from a large birth cohort of young adults. We were able to test three validated cardiometabolic risk prediction algorithms which are commonly used in clinical medicine in the UK, on a different population who are in clear and crucial need of a suitable tool.

Limitations of the study first and foremost relate to the exploratory analysis. The three algorithms we tested were not designed for use in young adults, though this in itself should not be a barrier to explore potential suitability in a different population. Nevertheless, our results should not be seen to cast doubt on the predictive ability of such algorithms when applied to the populations intended by the authors. We were unable to include every predictor from the algorithms we tested, which may have impacted upon performance statistics. That said, the impact of this limitation on our results may not have been uniform for each predictor we could not include. For example, even if we had the data, it is unlikely that many participants in our relatively young cohort would have diagnosed CVD or chronic kidney disease, a history of gestational diabetes or be prescribed statins. Also, our measured outcome differed from the outcome of the algorithms we tested. Whilst three algorithms included in the systematic review did aim to predict risk of metabolic syndrome, we did not consider them for our exploratory analysis since they did not include psychiatric predictors, were at relatively high risk of bias, and study authors did not publish their fully specified algorithm equations. Nevertheless, metabolic syndrome is a precursor of T2DM (26) and CVD (27), and the relatively good performance of the algorithm when we artificially substituted age to the mean age of the original study suggests face validity to our chosen outcome. Our sample size was relatively small compared with the original studies. However, by testing a more encompassing outcome, we were able to include a greater number of cases and reduce the impact of model overfit.

Other limitations relate to the systematic review. We were unable to follow a meta-analytic approach to the synthesis of results due to study

heterogeneity. The lack of meta-analytic approach meant we were unable to examine the risk of publication bias, which may have played a part in the configuration of studies we included in our synthesis, since only three included studies were not published in peer-reviewed journals.

In conclusion, young people who are at higher risk than the general population of developing psychosis are also at higher risk of developing cardiometabolic disorders. A suitable cardiometabolic risk prediction algorithm for this population would be highly beneficial to general and psychiatric practitioners to help them to tailor treatment plans with the aim of reducing long-term physical and psychiatric morbidity. Existing cardiometabolic risk algorithms cannot be recommended for this purpose since they likely underestimate the cardiometabolic risk of all young people, let alone a group already at significantly higher risk than the general population. Existing algorithms require recalibration to suit younger populations, and, better still, a new cardiometabolic risk prediction algorithm is required which is specifically developed for young people with psychosis. A well-designed algorithm may include a more appropriate distinction of metabolically active antipsychotics; should more appropriately weight the predictors for the specific characteristics of young people with psychosis; and may include a more age-appropriate outcome, such as metabolic syndrome. Further, particular attention should be paid to patient acceptability, to ensure the algorithm is actually used in clinical practice rather than simply buried in a research database. In lieu of a suitable algorithm, simple lifestyle interventions such as smoking cessation, encouraging a healthy diet and increasing physical activity must be offered to all young people with or at risk of psychosis. Indeed, encouraging results are emerging from studies of primary prevention in this population (174, 175), who may not have yet developed chronic and pervasive lifestyle behaviours which are associated with chronic illness.

Acknowledgements

This report is independent research supported by the National Institute for Health Research (NIHR Doctoral Research Fellowship, Dr Benjamin Ian Perry, DRF-2018-11-ST2-018). The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health and Social Care. GMK acknowledges funding support from the Wellcome Trust (Intermediate Clinical Fellowship; grant code: 201486/Z/16/Z); the BMA Foundation (J Moulton Grant 2019); the MQ: Transforming Mental Health (grant code: MQDS17/40 [with PBJ]); the Medical Research Council (grant code: MC_PC_17213) and the Medical Research Council (grant code: MR/S037675/1 [with RU]). PBJ receives grant support from the NIHR Applied Research Collaboration East of England. This publication is the work of the authors who will serve as guarantors for the contents of this paper. The authors report no conflicts of interest.

Conflict of Interest

The authors declare no conflicts of interest.

Peer Review

The peer review history for this article is available at https://publons.com/publon/10.1111/acps.13212.

Data availability statement

Access to anonymized data from ALSPAC is subjected to formal application processes. Please see http://www.bristol.ac.uk/alspac/researchers/ for more information.

References

1. Mahmood SS, Levy D, Vasan RS, Wang TJ. The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective. Lancet 2014;**383**:999–1008.
2. Scarborough P, Bhatnagar P, Wickramasinghe K, Smolina K, Mitchell C, Rayner M eds. Heart Disease Statistics. British Heart Foundation, London; 2010.
3. Simeone JC, Ward AJ, Rotella P, Collins J, Windisch R. An evaluation of variation in published estimates of schizophrenia prevalence from 1990 horizontal line 2013: a systematic literature review. BMC Psychiatry. 2015;**12**:193.
4. Holt RI, Peveler RC, Byrne CD. Schizophrenia, the metabolic syndrome and diabetes. Diabet Med. 2004;**21**:515–523.
5. Laursen TM, Plana-Ripoll O, Andersen PK et al. Cause-specific life years lost among persons diagnosed with schizophrenia: Is it getting better or worse? Schizophr Res. 2019;**206**:284–290.
6. Plana-Ripoll O, Pedersen CB, Agerbo E et al. A comprehensive analysis of mortality-related health metrics associated with mental disorders: a nationwide, register-based cohort study. Lancet 2019;**394**:1827–1835.
7. Plana-Ripoll O, Weye N, Momen NC et al. Changes Over Time in the Differential Mortality Gap in Individuals With Mental Disorders. JAMA. Psychiatry 2020.
8. Berry A, Drake RJ, Webb RT, Ashcroft DM, Carr MJ, Yung AR. Investigating the agreement between cardiovascular disease risk calculators among people diagnosed with Schizophrenia. Front Psychiatry 2018;**9**:685.
9. Perry BI, McIntosh G, Weich S, Singh S, Rees K. The association between first-episode psychosis and abnormal glycaemic control: systematic review and meta-analysis. Lancet Psychiatry. 2016;**3**:1049–1058.
10. Pillinger T, Beck K, Gobjila C, Donocik JG, Jauhar S, Howes OD. Impaired glucose homeostasis in first-episode schizophrenia: a systematic review and meta-analysis. JAMA Psychiatry 2017;**74**:261–269.
11. Perry BI, Upthegrove R, Thompson A, Marwaha S, Zammit S, Khandaker G. Relationships between dysglycaemia,

immune activation and psychotic experiences: findings from the U.K. ALSPAC Birth Cohort, In preparation.
12. Haddaway NR, Collins AM, Coughlin D, Kirk S. The role of google scholar in evidence reviews and its applicability to grey literature searching. PLoS One 2015;**10**:e0138237.
13. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Medicine 2009;**6**: e1000097.
14. Wolff RF, Moons KGM, Riley RD et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med. 2019;**170**:51–58.
15. Alba AC, Agoritsas T, Walsh M et al. Discrimination and calibration of clinical prediction models: Users' guides to the medical literature. JAMA 2017;**318**:1377–1384.
16. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol. 1996;**49**:1373–1379.
17. Pavlou M, Ambler G, Seaman SR et al. How to develop a more accurate risk prediction model when there are few events. BMJ 2015;**351**:h3868.
18. Ogundimu EO, Altman DG, Collins GS. Adequate sample size for developing prediction models is not simply related to events per variable. J Clin Epidemiol. 2016;**76**:175–182.
19. Rodgers M, Sowden A, Petticrew M et al. Testing methodological guidance on the conduct of narrative synthesis in systematic reviews: effectiveness of interventions to promote smoke alarm ownership and function. Evaluation 2009;**15**:47–71.
20. Boyd A, Golding J, Macleod J et al. Cohort Profile: The 'Children of the 90s'–the index offspring of the Avon Longitudinal Study of Parents and Children. Int J Epidemiol 2013;**42**:111–127.
21. Fraser A, Macdonald-Wallis C, Tilling K et al. Cohort profile: the avon longitudinal study of parents and children: ALSPAC mothers cohort. Int J Epidemiol 2013;**42**:97–110.
22. Northstone K, Lewcock M, Groom A et al. The Avon Longitudinal Study of Parents and Children (ALSPAC): an update on the enrolled sample of index children in 2019. Wellcome Open Res. 2019;**4**:51.
23. Harris PA, Taylor R, Minor BL et al. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform. 2019;**95**:103208.
24. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)–a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;**42**:377–381.
25. Alberti KG, Eckel RH, Grundy SM et al. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. Circulation 2009;**120**:1640–1645.
26. Shin JA, Lee JH, Lim SY et al. Metabolic syndrome as a predictor of type 2 diabetes, and its clinical interpretations and usefulness. J Diabetes Investig 2013;**4**:334–343.
27. Wilson PW, D'Agostino RB, Parise H, Sullivan L, Meigs JB. Metabolic syndrome as a precursor of cardiovascular disease and type 2 diabetes mellitus. Circulation 2005;**112**:3066–3072.
28. Hippisley-Cox J, Coupland C. Development and validation of QDiabetes-2018 risk prediction algorithm to estimate future risk of type 2 diabetes: cohort study. BMJ 2017;**359**:j5019.
29. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. BMJ 2017;**357**:j2099.
30. Osborn DP, Hardoon S, Omar RZ et al. Cardiovascular risk prediction models for people with severe mental illness: results from the prediction and management of cardiovascular risk in people with severe mental illnesses (PRIMROSE) research program. JAMA Psychiatry. 2015;**72**:143–151.
31. Buuren SV, Groothuis-Oudshoorn K. MICE: multivariate imputation by chained equations in R. J Stat Softw 2011;**45**:1–67.
32. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2017.
33. Abdul-Ghani MA, Abdul-Ghani T, Stern MP et al. Two-step approach for the prediction of future type 2 diabetes risk. Diabetes Care 2011;**34**:2108–2112.
34. Addoh O, Edwards MK, Loprinzi PD. Predictive validity of a medical-related cardiorespiratory fitness algorithm in predicting cardiovascular disease- and all-cause mortality: implications for integration into clinical practice. Mayo Clin Proc. 2016;**91**:1320–1321.
35. Aekplakorn W, Bunnag P, Woodward M et al. A risk score for predicting incident diabetes in the Thai population. Diabetes Care 2006;**29**:1872–1877.
36. Alghwiri A, Alghadir A, Awad H. The Arab Risk (ARABRISK): Translation and Validation. Biomed Res 2014;**25**:271–275.
37. Alssema M, Newson RS, Bakker SJL et al. One risk assessment tool for cardiovascular disease, type 2 diabetes, and chronic kidney disease. Diabetes Care 2012;**35**:741–748.
38. Alssema M, Vistisen D, Heymans MW et al. The Evaluation of Screening and Early Detection Strategies for Type 2 Diabetes and Impaired Glucose Tolerance (DETECT-2) update of the Finnish diabetes risk score for prediction of incident type 2 diabetes. Diabetologia 2011;**54**:1004–1012.
39. Artero EG, Jackson AS, Sui X et al. Longitudinal algorithms to estimate cardiorespiratory fitness: associations with nonfatal cardiovascular disease and disease-specific mortality. J Am Coll Cardiol 2014;**63**:2289–2296.
40. Artigao-Rodenas LM, Carbayo-Herencia JA, Divison-Garrote JA et al. Framingham risk score for prediction of cardiovascular diseases: a population-based study from southern Europe. PLoS One 2013;**8**:e73529.
41. Assmann G, Cullen P, Schulte H. Simple scoring scheme for calculating the risk of acute coronary events based on the 10-year follow-up of the prospective cardiovascular Munster (PROCAM) study. Circulation 2002;**105**:310–315.
42. Backholer K, Hirakawa Y, Tonkin A et al. Development of an Australian cardiovascular disease mortality risk score using multiple imputation and recalibration from national statistics. BMC Cardiovasc Disord. 2017;**17**:17.
43. Chen L, Magliano DJ, Balkau B et al. AUSDRISK: an Australian Type 2 Diabetes Risk Assessment Tool based on demographic, lifestyle and simple anthropometric measures. Med J Aust. 2010;**192**:197–202.
44. Chen L, Tonkin AM, Moon L et al. Recalibration and validation of the SCORE risk chart in the Australian population: the AusSCORE chart. Eur J Cardiovasc Prev Rehabil 2009;**16**:562–570.

45. Chen X, Wu Z, Chen Y et al. Risk score model of type 2 diabetes prediction for rural Chinese adults: the Rural Deqing Cohort Study. J Endocrinol Invest 2017;**40**:1115–1123.
46. Chien K, Cai T, Hsu H et al. A prediction model for type 2 diabetes risk among Chinese people. Diabetologia 2009;**52**:443–450.
47. D'Agostino RB, Grundy S, Sullivan LM, Wilson P. Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. JAMA 2001;**286**:180–187.
48. D'Agostino RB, Vasan RS, Pencina MJ et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation 2008;**117**:743–753.
49. de Bacquer D, de Backer G. Predictive ability of the SCORE Belgium risk chart for cardiovascular mortality. Int J Cardiol 2010;**143**:385–390.
50. Dugee O, Janchiv O, Jousilahti P et al. Adapting existing diabetes risk scores for an Asian population: a risk score for detecting undiagnosed diabetes in the Mongolian population. BMC Public Health 2015;**15**:938.
51. Gabriel R, Brotons C, Tormo MJ et al. The ERICE-score: the new native cardiovascular score for the low-risk and aged Mediterranean population of Spain. Rev Esp Cardiol (Engl Ed) 2015;**68**:205–215.
52. Gao WG, Dong YH, Pang ZC et al. A simple Chinese risk score for undiagnosed diabetes. Diabet Med 2010;**27**:274–281.
53. Gao WG, Qiao Q, Pitkaniemi J et al. Risk prediction models for the development of diabetes in Mauritian Indians. Diabet Med 2009;**26**:996–1002.
54. Glumer C, Carstensen B, Sandbaek A et al. A Danish diabetes risk score for targeted screening: the Inter99 study. Diabetes Care 2004;**27**:727–733.
55. Griffin SJ, Little PS, Hales CN, Kinmonth AL, Wareham NJ. Diabetes risk score: towards earlier detection of type 2 diabetes in general practice. Diabetes Metab Res Rev 2000;**16**:164–171.
56. Heianza Y, Arase Y, Saito K et al. Development of a screening score for undiagnosed diabetes and its application in estimating absolute risk of future type 2 diabetes in Japan: Toranomon Hospital Health Management Center Study 10 (TOPICS 10). J Clin Endocrinol Metab. 2013;**98**:1051–1060.
57. Heikes KE, Eddy DM, Arondekar B, Schlessinger L. Diabetes Risk Calculator: a simple tool for detecting undiagnosed diabetes and pre-diabetes. Diabetes Care 2008;**31**:1040–1045.
58. Hippisley-Cox J, Coupland C, Brindle P. Derivation and validation of QStroke score for predicting risk of ischaemic stroke in primary care and comparison with other risk scores: a prospective open cohort study. BMJ 2013;**346**:f2573.
59. Hippisley-Cox J, Coupland C, Robson J, Brindle P. Derivation, validation, and evaluation of a new QRISK model to estimate lifetime risk of cardiovascular disease: cohort study using QResearch database. BMJ 2010;**341**:c6624.
60. Hippisley-Cox J, Coupland C, Robson J, Sheikh A, Brindle P. Predicting risk of type 2 diabetes in England and Wales: prospective derivation and validation of QDScore. BMJ 2009;**338**:b880.
61. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M, Brindle P. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. BMJ 2007;**335**:136.
62. Hippisley-Cox J, Coupland C, Vinogradova Y et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. BMJ 2008;**336**:1475–1482.
63. Ha KH, Lee YH, Song SO et al. Development and validation of the Korean diabetes risk score: a 10-year national cohort study. Diabetes Metab J 2018;**42**:402–414.
64. Katulanda P, Hill NR, Stratton I, Sheriff R, De Silva SDN, Matthews DR. Development and validation of a Diabetes Risk Score for screening undiagnosed diabetes in Sri Lanka (SLDRISK). BMC Endocr Disord. 2016;**16**:42.
65. Ko G, So W, Tong P et al. A simple risk score to identify Southern Chinese at high risk for diabetes. Diabet Med. 2010;**27**:644–649.
66. Lindstrom J, Tuomilehto J. The diabetes risk score: a practical tool to predict type 2 diabetes risk. Diabetes Care 2003;**26**:725–731.
67. Muhlenbruch K, Ludwig T, Jeppesen C et al. Update of the German Diabetes Risk Score and external validation in the German MONICA/KORA study. Diabetes Res Clin Pract 2014;**104**:459–466.
68. Muhlenbruch K, Paprott R, Joost HG, Boeing H, Heidemann C, Schulze MB . Derivation and external validation of a clinical version of the German Diabetes Risk Score (GDRS) including measures of HbA1c. BMJ Open Diabetes Res Care. 2018;**6**:e000524.
69. Nanri A, Nakagawa T, Kuwahara K et al. Development of risk score for predicting 3-year incidence of type 2 diabetes: japan epidemiology collaboration on occupational health study. PLoS One 2015;**10**:e0142779.
70. Panagiotakos DB, Georgousopoulou EN, Fitzgerald AP, Pitsavos C, Stefanadis C. Validation of the HellenicSCORE (a Calibration of the ESC SCORE Project) regarding 10-year risk of fatal cardiovascular disease in Greece. Hellenic J Cardiol 2015;**56**:302–328.
71. Ridker PM, Paynter NP, Rifai N, Gaziano JM, Cook NR. C-reactive protein and parental history improve global cardiovascular risk prediction: the Reynolds Risk Score for men. Circulation 2008;**118**:2243–2251.
72. Robinson CA, Agarwal G, Nerenberg K. Validating the CANRISK prognostic model for assessing diabetes risk in Canada's multi-ethnic population. Chronic Dis Inj Can 2011;**32**:19–31.
73. Rosella LC, Manuel DG, Burchill C, Stukel TA. A population-based risk algorithm for the development of diabetes: development and validation of the Diabetes Population Risk Tool (DPoRT). J Epidemiol Community Health 2011;**65**:613–620.
74. Sarrafzadegan N, Hassannejad R, Marateb HR et al. PARS risk charts: A 10-year study of risk assessment for cardiovascular diseases in Eastern Mediterranean Region. PLoS One 2017;**12**:e0189389.
75. Schulze MB, Hoffmann K, Boeing H et al. An accurate risk score based on anthropometric, dietary, and lifestyle factors to predict the development of type 2 diabetes. Diabetes Care 2007;**30**:510–515.
76. Selmer R, Igland J, Ariansen I et al. NORRISK 2: A Norwegian risk model for acute cerebral stroke and myocardial infarction. Eur J Prev Cardiol 2017;**24**:773–782.
77. Sun F, Tao Q, Zhan S. An accurate risk score for estimation 5-year risk of type 2 diabetes based on a health screening population in Taiwan. Diabetes Res Clin Pract. 2009;**85**:228–234.
78. Tabaei BP, Engelgau MM, Herman WH. A multivariate logistic regression equation to screen for dysglycaemia: development and validation. Diabet Med. 2005;**22**:599–605.
79. Wen J, Hao J, Liang Y et al. A non-invasive risk score for predicting incident diabetes among rural Chinese people:

A village-based cohort study. PLoS One 2017;**12**: e0186172.
80. Wickramasinghe CD, Ayers CR, Das S, de Lemos JA, Willis BL, Berry JD. Prediction of 30-year risk for cardiovascular mortality by fitness and risk factor levels: the Cooper Center Longitudinal Study. Circ Cardiovasc Qual Outcomes 2014;**7**:597–602.
81. Wong CK, Siu SC, Wan EY et al. Simple non-laboratory- and laboratory-based risk assessment algorithms and nomogram for detecting undiagnosed diabetes mellitus. J Diabetes. 2016;**8**:414–421.
82. Yatsuya H, Iso H, Li Y et al. Development of a risk equation for the incidence of coronary artery disease and ischemic stroke for middle-aged Japanese- japan public health center-based prospective study. Circ J 2016;**80**:1386–1395.
83. Yatsuya H, Iso H, Yamagishi K et al. Development of a point-based prediction model for the incidence of total stroke: Japan public health center study. Stroke 2013;**44**:1295–1302.
84. Anderson JP, Parikh JR, Shenfeld DK et al. Reverse engineering and evaluation of prediction models for progression to type 2 diabetes: an application of machine learning using electronic health records. J Diabetes Sci Technol 2015;**10**:6–18.
85. AbdEl-Wahab EW, Shatat HZ, Charl F. Adapting a prediction rule for metabolic syndrome risk assessment suitable for developing countries. J Prim Care Community Health. 2019;**2019**:2150132719882760.
86. Alaa AM, Bolton T, di Angelantonio E, Rudd JHF, van der Schaar M. Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. PLoS One 2019;**14**: e0213653.
87. Arima H, Yonemoto K, Doi Y et al. Development and validation of a cardiovascular risk prediction model for Japanese: the Hisayama study. Hypertens Res 2009;**32**:1119–1122.
88. Aslibekyan S, Campos H, Loucks EB et al. Development of a cardiovascular risk score for use in low- and middle-income countries. J Nutr. 2011;**141**:1375–1380.
89. Balkau B, Hu G, Qiao Q et al. Prediction of the risk of cardiovascular mortality using a score that includes glucose as a risk factor, The DECODE Study. Diabetologia 2004;**47**:2118–2128.
90. Barazzoni R, Gortan Cappellari G, Semolic Aet al. Central adiposity markers, plasma lipid profile and cardiometabolic risk prediction in overweight-obese individuals. Clin Nutr. 2019;**38**:1171–1179.
91. Bell K, Hayen A, McGeechan K, Neal B, Irwig L. Effects of additional blood pressure and lipid measurements on the prediction of cardiovascular risk. Eur J Prev Cardiol 2012;**19**:1474–1485.
92. Boland B, de Muylder R, Goderis G et al. Cardiovascular prevention in general practice: development and validation of an algorithm. Acta Cardiol 2004;**59**:598–605.
93. Boucher H. Exploring the Utility of MUAC in Classifying Adult Metabolic Syndrome Using NHANES 2015–16. William Honors College: Honors Research Projects, University of Akron; 2019.
94. Brand RJ, Rosenman RH, Sholtz RI, Friedman M. Multivariate prediction of coronary heart disease in the western collaborative group study compared to the findings of the Framingham study. Circulation 1976;**53**:348–355.
95. Brautbar A, Ballantyne CM, Lawson K et al. Impact of adding a single allele in the 9p21 locus to traditional risk factors on reclassification of coronary heart disease risk and implications for lipid-modifying therapy in the Atherosclerosis Risk in Communities study. Circ Cardiovasc Genet 2009;**2**:279–285.
96. Chambless LE, Folsom AR, Sharrett AR et al. Coronary heart disease risk prediction in the Atherosclerosis Risk in Communities (ARIC) study. J Clin Epidemiol 2003;**56**:880–890.
97. Chien KL, Hsu HC, Su TC et al. Constructing a point-based prediction model for the risk of coronary artery disease in a Chinese community: a report from a cohort study in Taiwan. Int J Cardiol. 2012;**157**:263–268.
98. Choe EK, Rhee H, Lee S et al. Metabolic syndrome prediction using machine learning models with genetic and clinical information from a nonobese healthy population. Genomics Inform. 2018;**16**:e31.
99. Cross DS, McCarty CA, Hytopoulos E et al. Coronary risk assessment among intermediate risk patients using a clinical and biomarker based algorithm developed and validated in two population cohorts. Curr Med Res Opin 2012;**28**:1819–1830.
100. Davies RW, Dandona S, Stewart AF et al. Improved prediction of cardiovascular disease based on a panel of single nucleotide polymorphisms identified through genome-wide association studies. Circ Cardiovasc Genet 2010;**3**:468–474.
101. Dimopoulos AC, Nikolaidou M, Caballero FF et al. Machine learning methodologies versus cardiovascular risk scores, in predicting disease risk. BMC Med Res Methodol 2018;**12**:179.
102. Dunder K, Lind L, Zethelius B, Berglund L, Lithell H. Evaluation of a scoring scheme, including proinsulin and the apolipoprotein B/apolipoprotein A1 ratio, for the risk of acute coronary events in middle-aged men: Uppsala Longitudinal Study of Adult Men (ULSAM). Am Heart J. 2004;**148**:596–601.
103. Ferrario M, Chiodini P, Chambless LE et al. Prediction of coronary events in a low incidence population. Assessing accuracy of the CUORE Cohort Study prediction equation. Int J Epidemiol 2005;**34**:413–421.
104. Friedland DR, Cederberg C, Tarima S. Audiometric pattern as a predictor of cardiovascular status: development of a model for assessment of risk. Laryngoscope. 2009;**119**:473–486.
105. Gaziano TA, Young CR, Fitzmaurice G, Atwood S, Gaziano JM. Laboratory-based versus non-laboratory-based method for assessment of cardiovascular disease risk: the NHANES I Follow-up Study cohort. Lancet 2008;**371**:923–931.
106. Gupta P, Prieto-Merino D, Ajay VS. Cardiovascular risk prediction in India: comparison of the original and recalibrated Framingham prognostic models in Urban populations. [Version 1; peer review : 2 approved with reservations]. Wellcome Open Research 2019;**4**:1–14.
107. Hamer M, Chida Y, Stamatakis E. Utility of C-reactive protein for cardiovascular risk stratification across three age groups in subjects without existing cardiovascular diseases. Am J Cardiol 2009;**104**:538–542.
108. Hossain R, Mahmud SMH, Hossin MA. Predicting risk factor of obesity among middle-aged people using data mining techniques. Procedia Computer Sci 2018;**132**:1068–1076.
109. Inouye M, Abraham G, Nelson CP et al. Genomic risk prediction of coronary artery disease in 480,000 adults: implications for primary prevention. J Am Coll Cardiol 2018;**10**:1883–1893.
110. Kang HM, Kim DJ. Metabolic syndrome versus framingham risk score for association of self-reported coronary

heart disease: The 2005 Korean Health and Nutrition Examination Survey. Diabetes Metab J 2012;**36**:237–244.
111. Knuiman MW, Vu HT, Bartholomew HC. Multivariate risk estimation for coronary heart disease: the Busselton Health Study. Aust N Z J Public Health 1998;**22**:747–753.
112. L'Italien G, Ford I, Norrie J et al. The cardiovascular event reduction tool (CERT)–a simplified cardiac risk prediction model developed from the West of Scotland Coronary Prevention Study (WOSCOPS). Am J Cardiol. 2000;**85**:720–724.
113. Laurier D, Chau NP, Cazelles B, Segond P. Estimation of CHD risk in a French working population using a modified Framingham model. The PCV-METRA Group. J Clin Epidemiol 1994;**47**:1353–1364.
114. Lees JS, Welsh CE, Celis-Morales CA et al. Glomerular filtration rate by differing measures, albuminuria and prediction of cardiovascular disease, mortality and end-stage kidney disease. Nat Med 2019;**25**:1753–1760.
115. Liao X, Kerr D, Morales J, Duncan I. Application of machine learning to identify clustering of cardiometabolic risk factors in U.S. adults. Diabetes Technol Ther 2019;**05**:245–253.
116. Menotti A, Lanti M, Agabiti-Rosei E et al. New tools for prediction of cardiovascular disease risk derived from Italian population studies. Nutr Metab Cardiovasc Dis 2005;**2005**:426–440.
117. Merry AH, Boer JM, Schouten LJ et al. Risk prediction of incident coronary heart disease in The Netherlands: re-estimation and improvement of the SCORE risk function. Eur J Prev Cardiol 2012;**19**:840–848.
118. Mohammadreza B, Farzad H. Davoud K, FEREIDOUN PROF AF. Prognostic significance of the complex "Visceral Adiposity Index" vs. simple anthropometric measures: Tehran lipid and glucose study. Cardiovasc Diabetol 2012;**11**:20.
119. Noda H, Maruyama K, Iso H et al. Prediction of myocardial infarction using coronary risk scores among Japanese male workers: 3M Study. J Atheroscler Thromb 2010;**17**:452–459.
120. Park Y, Lim J, Lee J, Kim SG. Erythrocyte fatty acid profiles can predict acute non-fatal myocardial infarction. Br J Nutr 2009;**102**:1355–1361.
121. Paynter NP, Chasman DI, Buring JE, Shiffman D, Cook NR, Ridker PM. Cardiovascular disease risk prediction with and without knowledge of genetic variation at chromosome 9p21.3. Ann Intern Med 2009;**150**:65–72.
122. Paynter NP, Mazer NA, Pradhan AD, Gaziano JM, Ridker PM, Cook NR. Cardiovascular risk prediction in diabetic men and women using hemoglobin A1c vs diabetes as a high-risk equivalent. Arch Intern Med 2011;**171**:1712–1718.
123. Pencina MJ, D'Agostino RB, Larson MG, Massaro JM, Vasan RS. Predicting the 30-year risk of cardiovascular disease: the Framingham heart study. Circulation 2009;**119**:3078–3084.
124. Pocock SJ, McCormack V, Gueyffier F, Boutitie F, Fagard RH, Boissel JP. A score for predicting risk of death from cardiovascular disease in adults with raised blood pressure, based on individual patient data from randomised controlled trials. BMJ 2001;**323**:75–81.
125. Pylypchuk R, Wells S, Kerr A et al. Cardiovascular disease risk prediction equations in 400 000 primary care patients in New Zealand: a derivation and validation study. Lancet 2018;**05**:1897–1907.
126. Rana JS, Cote M, Després JP et al. Inflammatory biomarkers and the prediction of coronary events among people at intermediate risk: the EPIC-Norfolk prospective population study. Heart 2009;**95**:1682–1687.
127. Ridker PM, Buring JE, Rifai N, Cook NR. Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score. JAMA 2007;**297**:611–619.
128. Ayala Solares JR, Canoy D, Raimondi FED et al. Long-term exposure to elevated systolic blood pressure in predicting incident cardiovascular disease: evidence from large-scale routine electronic health records. J Am Heart Assoc 2019;**8**:e012129.
129. Stern MP, Williams K, González-Villalpando C, Hunt KJ, Haffner SM. Does the metabolic syndrome improve identification of individuals at risk of type 2 diabetes and/or cardiovascular disease? Diabetes Care 2004;**27**:2676–2681.
130. Tanabe N, Iso H, Okada K et al. Serum total and non-high-density lipoprotein cholesterol and the risk prediction of cardiovascular events – the JALS-ECC -. Circ J. 2010;**74**:1346–1356.
131. Tohidi M, Hadaegh F, Harati H, Azizi F. C-reactive protein in risk prediction of cardiovascular outcomes: Tehran Lipid and Glucose Study. Int J Cardiol 2009;**132**:369–374.
132. Voss R, Cullen P, Schulte H, Assmann G. Prediction of risk of coronary events in middle-aged men in the Prospective Cardiovascular Münster Study (PROCAM) using neural networks. Int J Epidemiol 2002;**31**:1253–1262; discussion 62–64.
133. GROUP WCRCW. World Health Organization cardiovascular disease risk charts: revised models to estimate risk in 21 global regions. Lancet Glob Health 2019;**7**:e1332–e1345.
134. Woodward M, Brindle P, Tunstall-Pedoe H. Adding social deprivation and family history to cardiovascular risk assessment: the ASSIGN score from the Scottish Heart Health Extended Cohort (SHHEC). Heart 2007;**93**:172–176.
135. Wu Y, Liu X, Li X et al. Estimation of 10-year risk of fatal and nonfatal ischemic cardiovascular diseases in Chinese adults. Circulation 2006;**114**:2217–25.
136. Zhang XF, Attia J, D'Este C, Yu XH, Wu XG. A risk score predicted coronary heart disease and stroke in a Chinese cohort. J Clin Epidemiol 2005;**58**:951–958.
137. Zhou X, Qiao Q, Ji L et al. Nonlaboratory-based risk assessment algorithm for undiagnosed type 2 diabetes developed on a nation-wide diabetes survey. Diabetes Care 2013;**36**:3944–3952.
138. Moons KG, Bots ML, Salonen JT et al. Prediction of stroke in the general population in Europe (EUROSTROKE): Is there a role for fibrinogen and electrocardiography? J Epidemiol Community Health. 2002;**56** (Suppl 1):i30–i36.
139. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ 2015;**350**:g7594.
140. Wilmot E, Idris I. Early onset type 2 diabetes: risk factors, clinical impact and management. Ther Adv Chronic Dis 2014;**5**:234–244.
141. Skinner AC, Perrin EM, Skelton JA. Prevalence of obesity and severe obesity in US children, 1999–2014. Obesity (Silver Spring) 2016;**24**:1116–1123.
142. Chrysant SG. A new paradigm in the treatment of the cardiovascular disease continuum: focus on prevention. Hippokratia 2011;**15**:7–11.
143. Weintraub WS, Daniels SR, Burke LE et al. Value of primordial and primary prevention for cardiovascular

disease: a policy statement from the American Heart Association. Circulation 2011;**124**:967–990.
144. Sagud M, Vuksan-Cusa B, Jaksic N, Mihaljevic-Peles A, Rojnic Kuzman M, Pivac N. Smoking in Schizophrenia: an updated. Review. Psychiatr Danub 2018;**30**:216–223.
145. Heald A, Pendlebury J, Anderson S et al. Lifestyle factors and the metabolic syndrome in Schizophrenia: a cross-sectional study. Ann Gen Psychiatry 2017;**16**:12.
146. Leucht S, Cipriani A, Spineli L et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. Lancet 2013;**382**:951–962.
147. Lawrence D, Kisely S. Inequalities in healthcare provision for people with severe mental illness. J Psychopharmacol 2010;**24**:61–68.
148. Jones S, Howard L, Thornicroft G. 'Diagnostic overshadowing': worse physical health care for people with mental illness. Acta Psychiatr Scand 2008;**118**:169–171.
149. Perry BI, Upthegrove R, Thompson A et al. Dysglycaemia, inflammation and psychosis: findings from the UK ALSPAC Birth Cohort. Schizophr Bull 2019;**45**:330–338.
150. Pillinger T, Beck K, Stubbs B, Howes OD. Cholesterol and triglyceride levels in first-episode psychosis: systematic review and meta-analysis. Br J Psychiatry 2017;**211**:339–349.
151. Misiak B, Stanczykiewicz B, Laczmanski L, Frydecka D. Lipid profile disturbances in antipsychotic-naive patients with first-episode non-affective psychosis: a systematic review and meta-analysis. Schizophr Res 2017;**190**:18–27.
152. Upthegrove R, Manzanares-Teson N, Barnes NM. Cytokine function in medication-naive first episode psychosis: a systematic review and meta-analysis. Schizophr Res 2014;**155**:101–108.
153. Miller BJ, Buckley P, Seabolt W, Mellor A, Kirkpatrick B. Meta-analysis of cytokine alterations in schizophrenia: clinical status and antipsychotic effects. Biol Psychiatry 2011;**70**:663–671.
154. Miller BJ, Culpepper N, Rapaport MH. C-reactive protein levels in schizophrenia: a review and meta-analysis. Clin Schizophr Relat Psychoses 2014;**7**:223–230.
155. Fernandes BS, Steiner J, Bernstein HG et al. C-reactive protein is increased in schizophrenia but is not altered by antipsychotics: meta-analysis and implications. Mol Psychiatry. 2016;**21**:554–564.
156. Khandaker GM, Pearson RM, Zammit S, Lewis G, Jones PB. Association of serum interleukin 6 and c-reactive protein in childhood with depression and psychosis in young adult life. JAMA Psychiatry 2014;**71**:1121.
157. Metcalf SA, Jones PB, Nordstrom T et al. Serum C-reactive protein in adolescence and risk of schizophrenia in adulthood: A prospective birth cohort study. Brain Behav Immun. 2017;**59**:253–259.
158. Goldsmith DR, Haroon E, Miller AH et al. Association of baseline inflammatory markers and the development of negative symptoms in individuals at clinical high risk for psychosis. Brain Behav Immun 2019;**76**:268–274.
159. Nielsen PR, Agerbo E, Skogstrand K, Hougaard DM, Meyer U, Mortensen PB. Neonatal levels of inflammatory markers and later risk of schizophrenia. Biol Psychiatry. 2015;**77**:548–555.
160. Rethorst CD, Bernstein I, Trivedi MH. Inflammation, obesity, and metabolic syndrome in depression: analysis of the 2009–2010 National Health and Nutrition Examination Survey (NHANES). J Clin Psychiatry 2014;**75**: e1428–e1432.
161. Monteiro R, Azevedo I. Chronic inflammation in obesity and the metabolic syndrome. Mediators Inflamm 2010;**2010**:1–10.
162. Hermsdorff HH, Zulet MA, Puchau B, Martinez JA. Central adiposity rather than total adiposity measurements are specifically involved in the inflammatory status from healthy young adults. Inflammation 2011;**34**:161–170.
163. Calabro P, Yeh ET. Intra-abdominal adiposity, inflammation, and cardiovascular risk: new insight into global cardiometabolic risk. Curr Hypertens Rep 2008;**10**:32–38.
164. Spertus J, Horvitz-Lennon M, Abing H, Normand S-L. Risk of weight gain for specific antipsychotic drugs: a meta-analysis. NPJ Schizophr 2018;**4**:12.
165. Sentissi O, Viala A, Bourdel MC et al. Impact of antipsychotic treatments on the motivation to eat: preliminary results in 153 schizophrenic patients. Int Clin Psychopharmacol 2009;**24**:257–264.
166. Vancampfort D, Firth J, Schuch FB et al. Sedentary behavior and physical activity levels in people with schizophrenia, bipolar disorder and major depressive disorder: a global systematic review and meta-analysis. World Psychiatry 2017;**16**:308–315.
167. Huhn M, Nikolakopoulou A, Schneider-Thoma J et al. Comparative efficacy and tolerability of 32 oral antipsychotics for the acute treatment of adults with multi-episode schizophrenia: a systematic review and network meta-analysis. Lancet 2019;**394**:939–951.
168. Vancampfort D, Stubbs B, Mitchell AJ et al. Risk of metabolic syndrome and its components in people with schizophrenia and related psychotic disorders, bipolar disorder and major depressive disorder: a systematic review and meta-analysis. World Psychiatry. 2015;**14**:339–347.
169. Starrenburg FC, Bogers JP. How can antipsychotics cause Diabetes Mellitus? Insights based on receptor-binding profiles, humoral factors and transporter proteins. Eur Psychiatry. 2009;**24**:164–170.
170. Kroeze WK, Hufeisen SJ, Popadak BA et al. H1-histamine receptor affinity predicts short-term weight gain for typical and atypical antipsychotic drugs. Neuropsychopharmacology 2003;**28**:519–526.
171. Zomer E, Osborn D, Nazareth I et al. Effectiveness and cost-effectiveness of a cardiovascular risk prediction algorithm for people with severe mental illness (PRIMROSE). BMJ Open. 2017;**7**:e018181.
172. Damen JA, Hooft L, Schuit E et al. Prediction models for cardiovascular disease risk in the general population: systematic review. BMJ 2016;**353**:i2416.
173. Reinikainen J, Laatikainen T, Karvanen J, Tolonen H. Lifetime cumulative risk factors predict cardiovascular disease mortality in a 50-year follow-up study in Finland. Int J Epidemiol. 2015;**44**:108–116.
174. Teasdale SB, Curtis J, Ward PB et al. The effectiveness of the Keeping the Body in Mind Xtend pilot lifestyle program on dietary intake in first-episode psychosis: Two-year outcomes. Obes Res Clin Pract. 2019;**13**:214–216.
175. Curtis J, Watkins A, Rosenbaum S et al. Evaluating an individualized lifestyle and life skills intervention to prevent antipsychotic-induced weight gain in first-episode psychosis. Early Interv Psychiatry. 2016;**10**:267–276.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Flow-diagram of included participants at risk of psychosis at age 18 or 24 years.

Development and external validation of the Psychosis Metabolic Risk Calculator (PsyMetRiC): a cardiometabolic risk prediction algorithm for young people with psychosis

Benjamin I Perry, Emanuele F Osimo, Rachel Upthegrove, Pavan K Mallikarjun, Jessica Yorke, Jan Stochl, Jesus Perez, Stan Zammit, Oliver Howes, Peter B Jones, Golam M Khandaker

Summary

Background Young people with psychosis are at high risk of developing cardiometabolic disorders; however, there is no suitable cardiometabolic risk prediction algorithm for this group. We aimed to develop and externally validate a cardiometabolic risk prediction algorithm for young people with psychosis.

Methods We developed the Psychosis Metabolic Risk Calculator (PsyMetRiC) to predict up to 6-year risk of incident metabolic syndrome in young people (aged 16–35 years) with psychosis from commonly recorded information at baseline. We developed two PsyMetRiC versions using the forced entry method: a full model (including age, sex, ethnicity, body-mass index, smoking status, prescription of a metabolically active antipsychotic medication, HDL concentration, and triglyceride concentration) and a partial model excluding biochemical results. PsyMetRiC was developed using data from two UK psychosis early intervention services (Jan 1, 2013, to Nov 4, 2020) and externally validated in another UK early intervention service (Jan 1, 2012, to June 3, 2020). A sensitivity analysis was done in UK birth cohort participants (aged 18 years) who were at risk of developing psychosis. Algorithm performance was assessed primarily via discrimination (C statistic) and calibration (calibration plots). We did a decision curve analysis and produced an online data-visualisation app.

Findings 651 patients were included in the development samples, 510 in the validation sample, and 505 in the sensitivity analysis sample. PsyMetRiC performed well at internal (full model: C 0·80, 95% CI 0·74–0·86; partial model: 0·79, 0·73–0·84) and external validation (full model: 0·75, 0·69–0·80; and partial model: 0·74, 0·67–0·79). Calibration of the full model was good, but there was evidence of slight miscalibration of the partial model. At a cutoff score of 0·18, in the full model PsyMetRiC improved net benefit by 7·95% (sensitivity 75%, 95% CI 66–82; specificity 74%, 71–78), equivalent to detecting an additional 47% of metabolic syndrome cases.

Interpretation We have developed an age-appropriate algorithm to predict the risk of incident metabolic syndrome, a precursor of cardiometabolic morbidity and mortality, in young people with psychosis. PsyMetRiC has the potential to become a valuable resource for early intervention service clinicians and could enable personalised, informed health-care decisions regarding choice of antipsychotic medication and lifestyle interventions.

Funding National Institute for Health Research and Wellcome Trust.

Lancet Psychiatry 2021; 8: 589–98

Published Online June 1, 2021 https://doi.org/10.1016/S2215-0366(21)00114-0

Department of Psychiatry, University of Cambridge, Cambridge, UK (B I Perry MRCPsych, E F Osimo MRCPsych, J Stochl PhD, Prof J Perez PhD, Prof P B Jones PhD, Prof G M Khandaker PhD); **Cambridgeshire and Peterborough NHS Foundation Trust, Cambridge, UK** (B I Perry, E F Osimo, Prof J Perez, Prof P B Jones, Prof G M Khandaker); **MRC London Institute of Medical Sciences, Institute of Clinical Sciences, Imperial College, London, UK** (E F Osimo, Prof O Howes PhD); **Institute for Mental Health, University of Birmingham, Birmingham, UK** (Prof R Upthegrove PhD, P K Mallikarjun PhD); **Birmingham Women's and Children's NHS Trust Early Intervention Service, Birmingham, UK** (J Yorke MBBS); **Department of Kinanthropology, Charles University, Prague, Czech Republic** (J Stochl); **Centre for Academic Mental Health** (Prof S Zammit PhD, Prof G M Khandaker) **and MRC Integrative Epidemiology Unit** (Prof G M Khandaker), **Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK; MRC Centre for Neuropsychiatric Genetics and Genomics, Cardiff University, Cardiff, UK** (Prof S Zammit); **Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK** (Prof O Howes)

Introduction

People with psychotic disorders such as schizophrenia have a life expectancy shortened by 10–15 years compared with the general population,[1] predominantly owing to a higher prevalence of physical conditions such as type 2 diabetes, obesity, and cardiovascular disease (CVD).[2] These comorbidities lead to a reduced quality of life and substantial health economic burden[3] and usually develop early in the course of the psychotic disorder. For example, insulin resistance and dyslipidaemia are detectable from the onset of psychosis in adults in the second or third decades of life,[4,5] probably due to a combination of genetic, lifestyle, and other environmental influences.[6] Since some treatments for psychosis can exacerbate cardiometabolic risk (eg, certain antipsychotic medications), identification of young adults at the highest risk of adverse cardiometabolic outcomes as soon as possible after diagnosis of a psychotic disorder is crucial, so that interventions can be tailored to reduce the risk of longer-term cardiovascular morbidity and mortality.

Prognostic risk prediction algorithms are a valuable means to encourage personalised, informed health-care decisions. In the general population, cardiometabolic risk prediction algorithms such as QRISK3[7] are commonly used to predict CVD risk from baseline demographic, lifestyle, and clinical information, to identify higher-risk individuals for tailored interventions. A recent systematic review[8] explored the suitability of existing cardiometabolic risk prediction algorithms for young people with psychosis. However, all algorithms

Correspondence to:
Dr Benjamin I Perry,
Department of Psychiatry,
University of Cambridge,
Cambridge, CB2 0SZ, UK
bip20@medschl.cam.ac.uk

Research in context

Evidence before this study

Cardiometabolic risk prediction algorithms are commonly used in the general population as tools to encourage informed, personalised treatment decisions with the aim of primary prevention of longer-term cardiometabolic outcomes. In a recent systematic review of cardiometabolic risk prediction algorithms developed either for general or psychiatric populations, we searched Embase (1947 to Dec 1, 2019), Ovid MEDLINE (1946 to Dec 1, 2019), PsychINFO (1806 to Dec 1, 2019), Web of Science (from inception to Dec 1, 2019), and the first 20 pages of Google Scholar (to Dec 1, 2019). Search terms related to cardiometabolic (metabolism, metabolic, diabetes mellitus, cardiovascular disease, obesity, cardiometabolic); risk prediction (risk assessment, risk, outcome assessment, prediction, prognosis); and algorithm (calculator, computers, algorithms, software, tool) were included. Over 100 studies were included in the review. Yet, few were validated externally, only one was developed in a sample of people with mental illness, none were done in young populations, most were rated as being at high risk of bias, and most did not include relevant predictors such as antipsychotic medication. Additionally, existing algorithms substantially underpredict cardiometabolic risk in young people with or at risk of developing psychosis. Therefore, existing algorithms are unlikely to be suitable for young people with psychosis.

Added value of this study

We have developed and externally validated, to our knowledge, the first clinically useful and age-appropriate cardiometabolic risk prediction algorithm tailored for young people with psychosis—the Psychosis Metabolic Risk Calculator (PsyMetRiC)—using patient data from three geographically distinct UK National Health Service psychosis early intervention services. PsyMetRiC can reliably predict the risk of incident metabolic syndrome in young people with psychosis and young people who are at risk of developing psychosis.

Implications of all the available evidence

Whereas established risk prediction algorithms are suitable for use in older general population samples, with PsyMetRiC we are able to extend cardiometabolic risk prediction to young people with psychosis, a group who are at significantly higher cardiometabolic risk than the general population. Our findings can pave the way for a future clinical tool to encourage personalised treatment decisions with the aim of improving the long-term physical health of young people with psychosis.

were developed in samples of adults with a mean age across included studies of 50·5 years, and no studies included participants younger than 35 years. Most included studies did not include relevant predictors such as antipsychotic medication, so the authors of the review concluded that none are likely to be suitable for young people with psychosis.[8] Furthermore, an accompanying exploratory analysis found that existing algorithms significantly underpredict cardiometabolic risk in young people with or at risk of developing psychosis.[8]

Therefore, following TRIPOD reporting guidelines[9] (appendix p 19), we developed and externally validated the Psychosis Metabolic Risk Calculator (PsyMetRiC) to predict up to 6-year risk of metabolic syndrome, an age-appropriate precursor of CVD and early mortality, in young people with psychosis. We prioritised clinical usefulness and patient acceptability via input from a young person's advisory group, and by developing two PsyMetRiC versions: one with and one without biochemical results.

See Online for appendix

Methods

Data sources

We developed PsyMetRiC using pooled retrospective data from patients aged 16–35 years enrolled in the Birmingham psychosis early intervention service (EIS; sample frame n=391) or Cambridgeshire and Peterborough Assessing, Managing and Enhancing Outcomes (CAMEO) EIS (sample frame n=1113). Anonymised data from the Birmingham psychosis EIS were collected between Jan 1, 2014, and Dec 31, 2018, as part of the National Clinical Audit of Psychosis Quality Improvement programme, and were enhanced locally with medication data conforming to the Health Research Authority definition of service evaluation, which were confirmed by Birmingham Women's and Children's Hospital National Health Service (NHS) Foundation Trust. CAMEO data were identified by anonymously searching for EIS patients enrolled between Jan 1, 2013, and Nov 4, 2020, using the Clinical Records Anonymisation and Text Extraction (CRATE) tool[10] (NHS National Research Ethics Service references 12/EE/0407 and 17/EE/0442). Predictors were assessed at the closest point (within 100 days) to EIS enrolment, and outcomes were assessed up to 6 years later. We excluded patients who had less than 1 year of follow-up, had the outcome at baseline, or had missing data on all predictor or outcome variables.

To externally validate PsyMetRiC, we used the Clinical Records Interactive Search (CRIS) resource to capture anonymised data from South London and Maudsley NHS Foundation Trust (SLaM) EIS (National Institute for Health Research [NIHR] Biomedical Research Centre [BRC] CRIS Oversight Committee reference 20-005). Our sample frame included 2985 EIS patients aged 16–35 years enrolled between Jan 1, 2012, and June 3, 2020. Patients were excluded and predictors and outcomes were assessed in the same manner as for the development set.

In a sensitivity analysis, we examined the performance of PsyMetRiC in young adults who had or were at risk of

developing psychosis from the Avon Longitudinal Study of Parents and Children (ALSPAC) birth cohort (appendix p 1).[11] Our sample frame included participants identified as having experienced definite psychotic symptoms at either 18 years or 24 years, assessed via the semi-structured Psychosis-Like Symptom Interview (appendix p 1). Predictors were assessed at age 18 years, and the outcome was assessed at age 24 years. We excluded participants as described for the development set.

The ALSPAC Ethics and Law Committee and local research ethics committees provided ethical approval. Informed consent was obtained from patients following the recommendations of the ALSPAC Ethics and Law Committee at the time the data were collected.

Outcomes

We used the harmonised definition[12] of the metabolic syndrome as a binary outcome, which was at least three from the following list: ethnicity-specific waist circumference of at least 94 cm in males and at least 80 cm in females for white people, at least 90 cm in males and at least 80 cm in females for other ethnic groups, or body-mass index (BMI) greater than 29·9 kg/m²; triglyceride concentrations at least 1·70 mmol/L; HDL concentration less than 1·03 mmol/L in males or less than 1·29 mmol/L in females; systolic blood pressure greater than 130 mm Hg; or fasting plasma glucose greater than 5·60 mmol/L.

Predictor variables

Predictors were included based on a balance of clinical knowledge, past research, likely clinical usefulness, and patient acceptability after discussion of the work with the McPin Foundation Young Persons Advisory Group (YPAG), a group of volunteers aged younger than 24 years with personal experience of mental health difficulties (appendix p 10). The full model comprised age (continuous; years), ethnicity (categorical; white European or not recorded [reference], Black or African-Caribbean, Asian, or other), sex (female or male), BMI (continuous; kg/ m²), current smoking status (binary; at least one cigarette on average daily), prescription of a metabolically active antipsychotic drug (binary; based on relative cardiometabolic risk; appendix p 11), HDL concentration (continuous; mmol/L), and triglyceride concentration (continuous; mmol/L). A partial model, without HDL and triglyceride concentrations, was developed to cover eventualities where biochemical results are not available (appendix pp 5–8).

Statistical analysis

We developed PsyMetRiC using the forced entry method, after ruling out predictor multi-collinearity, to minimise risk of overfitting and as recommended for smaller datasets.[13] We did a formal sample size calculation.[14] Briefly, the sample size required was estimated from the estimated outcome prevalence, the a priori estimated R^2 of the model, and the estimated required model shrinkage. For the full model, the minimum sample required was 494, and for the partial model it was 394 (appendix p 2).[14] We did not consider non-linear terms or interactions to reduce risk of overfitting. We used multiple imputation using chained equations for missing data and we pooled estimates using Rubin's rules (appendix p 3). An initial internal validation step (500 bootstraps) was done, and coefficients were shrunk for optimism using the pooled corrected C slope as a shrinkage factor. After this step, predictive performance was assessed (see later).

The algorithms were applied to the external validation sample. The distribution of predicted outcome probabilities was inspected using histograms. Algorithm performance was primarily assessed with measures of discrimination (C statistic) and calibration (calibration plots; appendix p 4). We also recorded Nagelkerke-Cox-Snell-Maddala-Magee R^2 index, the calibration intercept (ideally close to 0), C slope (ideally close to 1), and the Brier score, which is an overall measure of algorithm performance (ideally close to 0, with scores >0·25 generally indicating a poor model).

Decision curve analysis[15] was used to assess the clinical usefulness of PsyMetRiC by estimating net benefit. Net benefit is a metric of true positives minus false positives, and is calculated as

$$\text{sensitivity} \times \text{prevalence} - (1 - \text{specificity}) \times (1 - \text{prevalence}) \times w$$

where w is the outcome odds at a given risk threshold.[16] The risk threshold is the amount of tolerable risk before an intervention is deemed necessary. Net benefit incorporates the consequences of the decisions made on the basis of an algorithm, and is therefore preferable to related measures such as sensitivity and specificity alone.[16] We also reported the standardised net benefit (net benefit/outcome prevalence) and related metrics (sensitivity and specificity). In decision curve analysis, consideration only of the range of risk thresholds that may reasonably be considered in clinical practice is customary. Our upper bound of 0·35 represents a greater than one in three chance of developing metabolic syndrome should nothing change, and it is unlikely that risk thresholds greater than this should be tolerated. We drew a decision curve plot to visualise the net benefit of both PsyMetRiC versions over varying risk thresholds compared with intervening in all patients or intervening in no-one. Net harm (ie, more false positives than true positives exposed to an intervention at a selected risk threshold) is indicated when a proposed intervention is plotted at y<0. Classical decision theory proposes that at a chosen risk threshold, the choice with the greatest net benefit should be preferred.[16]

	Development sample			SLaM EIS external validation sample (n=510)	ALSPAC risk of psychosis sensitivity analysis sample (n=505)
	Birmingham EIS (n=352)	CAMEO EIS (n=299)	Pooled development sample (n=651)		
Age, years	23·76 (4·90)	25·42 (4·77)	24·52 (4·91)	24·45 (4·75)	17·81 (0·43)
Ethnicity					
White European or not recorded	111 (32%)	250 (84%)	361 (55%)	154 (30%)	494 (98%)
Black or African-Caribbean	94 (27%)	15 (5%)	109 (17%)	250 (49%)	<5 (<1%)*
Asian or other	147 (42%)	34 (11%)	181 (28%)	106 (21%)	<5 (<1%)*
Sex					
Male	232 (66%)	208 (70%)	440 (68%)	351 (69%)	184 (36%)
Female	120 (34%)	91 (30%)	211 (32%)	159 (31%)	321 (64%)
HDL concentration, mmol/L	1·76 (0·35)	2·08 (0·49)	1·88 (0·57)	1·57 (0·37)	1·21 (0·31)
Triglycerides concentration, mmol/L	1·46 (1·18)	1·30 (0·89)	1·39 (1·06)	1·23 (0·71)	1·06 (0·77)
BMI, kg/m²	22·06 (5·13)	24·01 (5·73)	23·63 (5·43)	22·96 (6·94)	23·22 (3·55)
FPG, mmol/L	5·20 (1·02)	5·17 (1·45)	5·19 (1·28)	5·03 (1·10)	5·31 (0·49)
Systolic BP, mm Hg	121·18 (11·04)	119·88 (12·25)	120·65 (11·68)	119·96 (13·70)	115·10 (11·88)
Metabolically active antipsychotics†	239 (68%)	216 (72%)	455 (70%)	472 (93%)	58 (11%)
Current smoker	182 (52%)	133 (44%)	315 (48%)	469 (92%)‡	286 (57%)
Follow-up, years	2·44 (1·54)	1·43 (1·03)	1·86 (1·32)	2·73 (1·76)	5·18 (0·39)
Time of predictor assessment from EIS enrolment, days	23·55 (25·44)	21·93 (29·84)	16·71 (26·38)	3·05 (36·01)	§
Metabolic syndrome at baseline¶	31/383 (8%)	18/317 (6%)	49/700 (7%)	30/540 (6%)	22/527 (4%)
Metabolic syndrome at follow-up	74 (21%)	35 (12%)	109 (17%)	86 (17%)	76 (15%)

Data are mean (SD), number (%), or n/N (%). Some percentags do not add up to 100 because of rounding. ALSPAC=Avon Longitudinal Study of Parents and Children. BMI=body-mass index. BP=blood pressure. CAMEO=Cambridgeshire and Peterborough Assessing, Managing and Enhancing Outcomes. EIS=early intervention service. FPG=fasting plasma glucose. SLaM=South London and Maudsley NHS Foundation Trust. *Reported as <5 owing to ALSPAC reporting guidelines. †Listed in the appendix (p 11). ‡Smoking status was derived using the CRIS-IE-Smoking application using natural language processing software to extract ever smoking status information from open-text fields (appendix p 6). §Health record and service use data are not available in ALSPAC. ¶N numbers are the sample size before excluding cases with metabolic syndrome at baseline.

Table 1: **Demographics and clinical characteristics of patients in the algorithm development and internal and external validation sets**

	Full model	Partial model
Intercept	−6·439813	−6·973829
Age, years	0·006233226	0·00633115
Black or African-Caribbean ethnicity	0·004258861	0·07548129
Asian or other ethnicity	0·211217746	0·29285950
Male sex	0·222300765	0·31460036
Body-mass index, kg/m²	0·141186241	0·16912161
Current smoker	0·153691193	0·24751854
Prescribed a metabolically active antipsychotic	0·497552758	0·60013558
HDL, mmol/L	−0·399013329	*
Triglycerides, mmol/L	0·343528440	*

*Variable not included in model.

Table 2: **Final coefficients for the Psychosis Metabolic Risk Calculator after shrinkage for optimism**

Visual representation of PsyMetRiC

We have provided two simulated case histories applying PsyMetRiC algorithms. Additionally, we developed an online data-visualisation app using shiny for R, which allows an interactive exploration of the effect of modifiable and non-modifiable risk factors and their combinations on cardiometabolic risk in young people with psychosis according to their PsyMetRiC score.

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

Data from 651 patients were included in the pooled development sample: 352 from the Birmingham EIS and 299 from CAMEO (table 1). After 500 bootstraps, the

pooled corrected C slope was 0·90 for the full model and 0·93 for the partial model; these values were used as shrinkage factors. Final PsyMetRiC coefficients are presented in table 2. Histograms showing the distribution of predicted outcome probabilities are provided in the appendix (p 14).

At internal validation, the pooled performance statistics for the full model were C 0·80 (95% CI 0·74 to 0·86); R^2 0·25 (95% CI 0·22 to 0·28); Brier score 0·07 (95% CI 0·05 to 0·09); and intercept –0·05 (95% CI –0·08 to –0·02). For the partial model, these statistics were C 0·79 (95% CI 0·73 to 0·84); R^2 0·19 (95% CI 0·14 to 0·24); Brier score 0·10 (95% CI 0·07 to 0·13); and intercept –0·07 (95% CI –0·10 to –0·04). Calibration plots showed good agreement between observed and expected risk at most predicted probabilities, although in both PsyMetRiC versions there was evidence of slight overprediction of risk at higher predicted probabilities (appendix p 15).

Our sample frame in the SLaM EIS identified 2985 patients, 510 of whom were eligible for inclusion in the SLaM external validation set; the appendix (p 9) provides details of the missing sample analysis. After applying PsyMetRiC to the SLaM EIS patient sample, performance statistics for the full model were C 0·75 (95% CI 0·69 to 0·80); R^2 0·21 (95% CI 0·18 to 0·25); Brier score 0·07 (95% CI 0·04 to 0·10); and intercept –0·05 (95% CI –0·08 to –0·02). For the partial model, these statistics were C 0·74 (95% CI 0·67 to 0·79); R^2 0·17 (95% CI 0·14 to 0·20); Brier score 0·08 (95% CI 0·05 to 0·11); and intercept –0·07 (95% CI –0·11 to –0·03). Calibration plots showed good agreement between observed and expected risk in the full model, but in the partial model there was evidence of slight miscalibration (underprediction of risk at lower predicted probabilities, and overprediction of risk at higher predicted probabilities; figure 1). In both models, 95% CIs widened as predicted probabilities became more extreme owing to lower numbers of participants with more extreme predicted probabilities (appendix p 15).

The sample frame for the ALSPAC validation set comprised 505 patients. In the ALSPAC sample, performance statistics for the full model were C 0·73 (95% CI 0·66 to 0·79); R^2 0·20 (95% CI 0·17 to 0·23); Brier score 0·08 (95% CI 0·04 to 0·11); and intercept –0·03 (95% CI –0·07 to 0·01). For the partial model, these statistics were C 0·71 (95% CI 0·64 to 0·77); R^2 0·17 (95% CI 0·13 to 0·22); Brier score 0·09 (95% CI 0·05 to 0·13); and intercept –0·03 (95% CI –0·07 to 0·00). The appendix (p 17) shows histograms of predicted outcome probabilities. Calibration plots showed good agreement between observed and expected risk in the full model, albeit with some minor evidence of miscalibration (slight underprediction of risk at lower predicted probabilities, and overprediction of risk at higher predicted probabilities; appendix p 18). The same pattern of slight miscalibration was marginally more pronounced in the partial model.

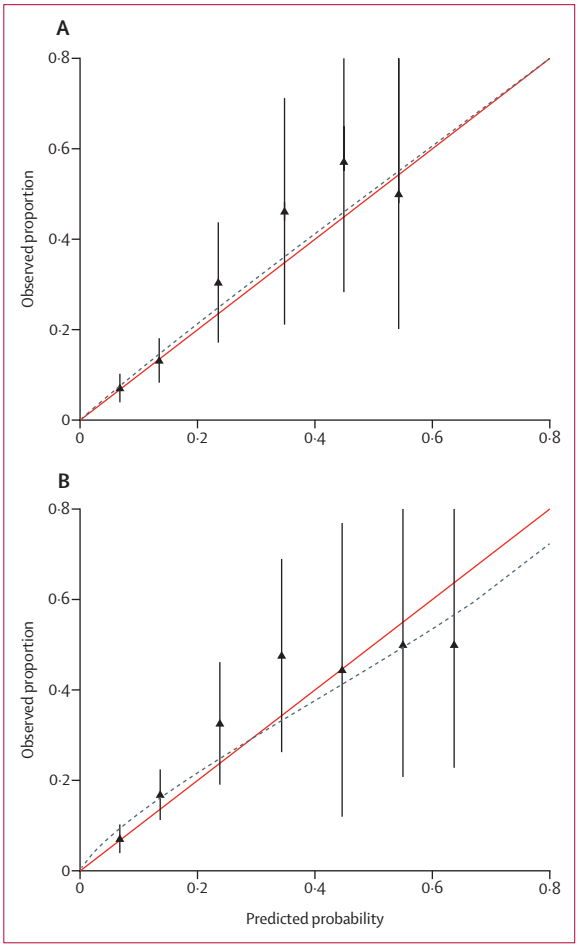


Figure 1: **Calibration plots for external validation of PsyMetRiC algorithms in an early intervention service patient sample**
Calibration plots are shown for the PsyMetRiC full model (A) and partial model (B). Calibration plots illustrate agreement between observed risk (y axis) and predicted risk (x axis). Perfect agreement would trace the red line. Algorithm calibration is shown by the dashed line. Triangles denote grouped observations for participants at deciles of predicted risk, with 95% CIs indicated by the vertical black lines. Axes range between 0 and 0·8 since very few individuals received predicted probabilities greater than 0·8. PsyMetRiC=Psychosis Metabolic Risk Calculator.

Decision curve analysis suggested that at predicted probability cutoffs greater than 0·05, both PsyMetRiC algorithms provided greater net benefit than the competing extremes of intervening in all patients or in none (figure 2). At most risk thresholds greater than 0·05, the full model provided slight improvement in net benefit compared with the partial model. The appendix (pp 12–13) provides numerical decision curve analysis results (net benefit, standardised net benefit, sensitivity, and specificity) across a range of reasonable risk thresholds. For example, if an intervention were considered necessary above a risk score of 0·18, the full model would provide a net benefit of 7·95% (95% CI 5·37–10·82), with a sensitivity of 75% (95% CI 66–82) and specificity of 74% (71–78), meaning that an additional 47% of metabolic

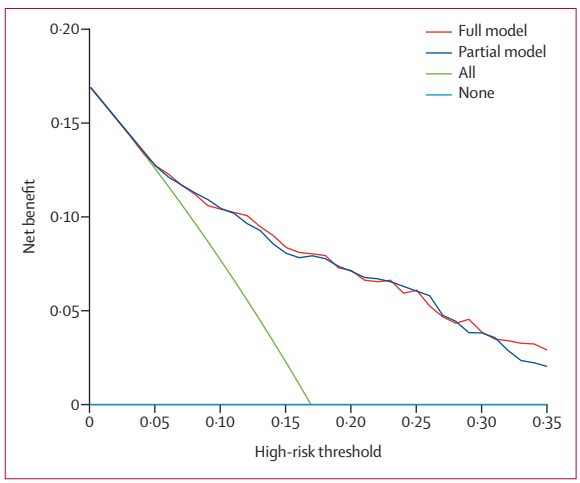


Figure 2: **Decision curve analysis plot for PsyMetRiC full and partial models**
The plot reports net benefit (y axis) of PsyMetRiC full and partial models across a range of risk thresholds (x axis) compared with intervening in all patients or intervening in no patients. PsyMetRiC=Psychosis Metabolic Risk Calculator.

syndrome cases could be prevented (standardised net benefit). At the same risk threshold, the partial model would provide a net benefit of 7·74% (95% CI 4·79–10·36), with a sensitivity of 75% (95% CI 65–81) and specificity of 74% (70–77), meaning that an additional 46% of metabolic syndrome cases could be prevented (standardised net benefit). For both models, these data equate to around an additional eight cases of metabolic syndrome that could be prevented per 100 individuals, without any increase in false positives.

Figure 3 shows decision trees outlining two simulated case scenarios to visualise the effect of modifiable and non-modifiable risk factors in young people with psychosis, as calculated from PsyMetRiC full and partial models. We have developed an online data visualisation app for both PsyMetRiC versions, which allows the user to interactively explore the effect of modifiable and non-modifiable risk factors and their combinations on cardiometabolic risk in young people with psychosis, based on PsyMetRiC scores.

For the **data visualisation app** see http://psymetric.shinyapps.io/psymetric

Discussion

We have developed and externally validated PsyMetRiC, which is to our knowledge the first cardiometabolic risk prediction algorithm tailored specifically for young people with psychosis. PsyMetRiC can predict up to 6-year risk of incident metabolic syndrome from commonly recorded clinical information, highlighting modifiable risk factors that could be addressed to reduce risk. Metabolic syndrome is a precursor to CVD and early mortality,[18] and is a suitable outcome for younger populations, since it occurs more commonly in younger adults than do more distal cardiovascular endpoints such as CVD. The external validation of both PsyMetRiC versions was good, with C statistics greater than 0·70. Calibration of the full model was good, but there was evidence of slight miscalibration of the partial model. Therefore, the partial model in particular may benefit from recalibration in larger samples. Both PsyMetRiC versions displayed greater net benefit than alternative strategies across a range of feasible risk thresholds, although at most risk thresholds our results show that the full model should be used preferentially.

Our data visualisations help to illustrate three things: first, antipsychotic medication choice imparts a substantial influence on cardiometabolic risk; second, addressing lifestyle factors can effectively reduce cardiometabolic risk even in the presence of antipsychotic medication; and third, advancing age in young adults does not influence cardiometabolic risk substantially relative to other risk factors. Although PsyMetRiC will benefit from future validation in larger samples, it has the potential to become a valuable resource to promote better management of physical health in young people with psychosis—eg, by highlighting modifiable risk factors and encouraging clinicians to make more personalised, informed decisions, such as with the choice of antipsychotic medication or lifestyle interventions, or both.

Ethnicity, smoking, and BMI are among the most commonly included predictors in existing algorithms[8] and are well known contributors to cardiometabolic risk,[19] so we included them in PsyMetRiC. Sex is also frequently considered in existing algorithms,[8] and we included it in PsyMetRiC. We found that male sex was a risk factor for incident metabolic syndrome, which aligns with meta-analytic reports that male sex is a risk factor for antipsychotic-induced metabolic dysfunction.[19] Our available sample size was too small to be able to consider separate versions of PsyMetRiC for males and females. If larger samples become available in the future, sex-stratified versions could be considered, since existing algorithms developed for the general population commonly take this step.[8]

Age is frequently included in existing algorithms,[8] and we included it in PsyMetRiC. However, existing cardiometabolic risk prediction algorithms, which were developed for older adults, weighted age to a greater extent than other predictors.[8] This is probably because most cardiometabolic risk factors contribute cumulative risk over time;[20] thus, age becomes increasingly important as one gets older. A recent exploratory analysis[8] that examined the predictive performance of the existing general population cardiometabolic risk prediction algorithms, including QRISK3[7] and PRIMROSE,[21] in young people who were at risk of developing psychosis found that each significantly underpredicted risk in the younger population, possibly owing to the way existing algorithms have modelled age. For example, in PsyMetRiC, age is weighted to a much lesser extent than other predictors, and we achieved favourable calibration in younger populations. Although QRISK3[7] and PRIMROSE[21] are good examples of well designed algorithms from large

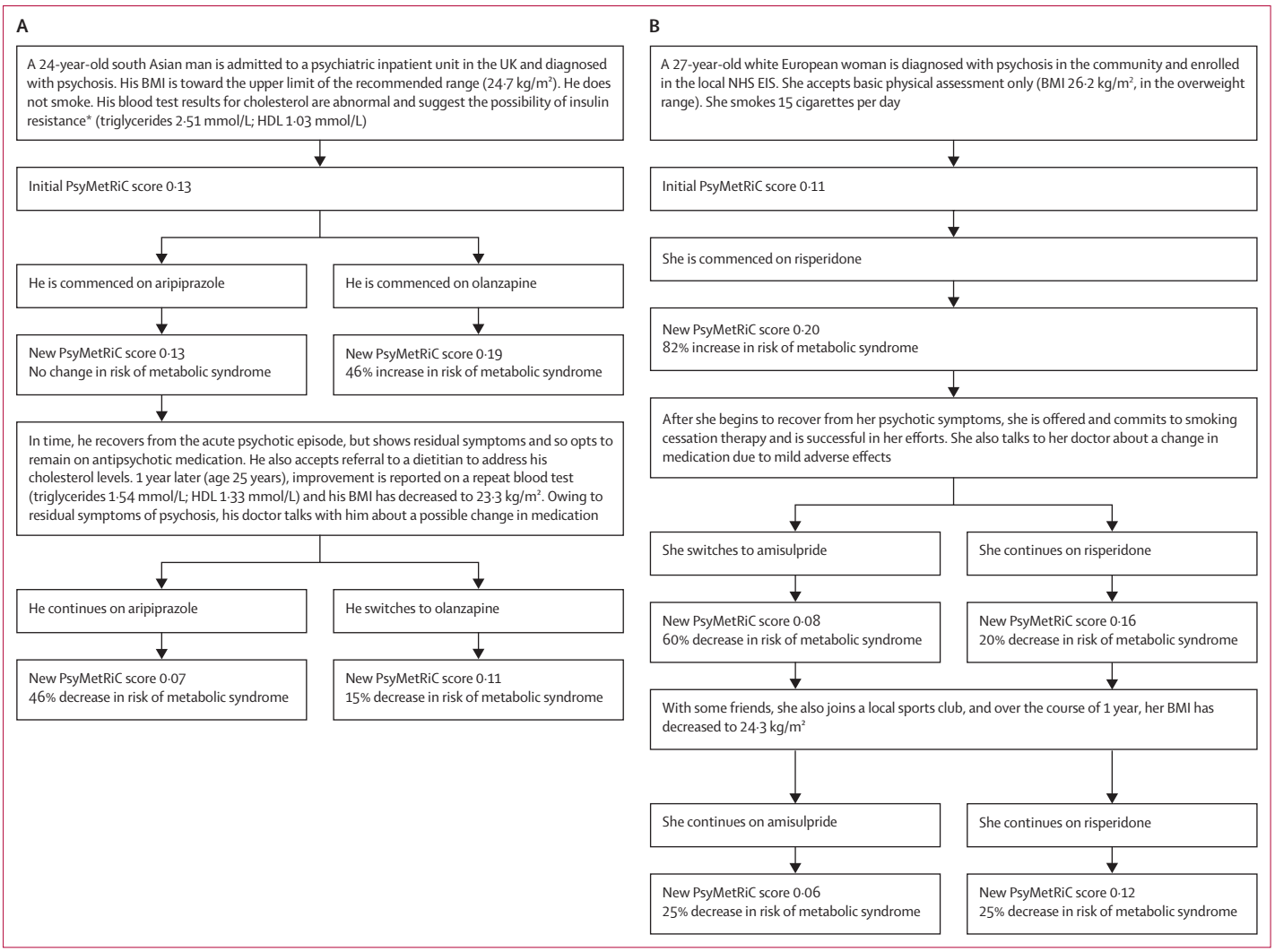


Figure 3: **Simulated case scenarios to visualise the effect of modifiable and non-modifiable risk factors on cardiometabolic risk in young people with psychosis as calculated from PsyMetRiC full and partial models**

Case scenarios are shown for the PsyMetRiC full model (A) and partial model (B). PsyMetRiC scores are presented as predicted probabilities, which can be converted to percentage chance of incident metabolic syndrome by multiplying by 100. BMI=body-mass index. EIS=early intervention service. NHS=National Health Service. PsyMetRiC=Psychosis Metabolic Risk Calculator. *A raised triglyceride:HDL ratio is indicative of insulin resistance.[17]

samples, our results suggest that PsyMetRiC is more appropriate for young people with psychosis.

Blood-based predictors, such as HDL and triglyceride concentrations, feature relatively infrequently in cardiometabolic risk prediction algorithms.[8] Meta-analytic evidence suggests abnormal triglyceride and HDL concentrations are detectable at first-episode psychosis,[22] and a raised triglyceride:HDL ratio is a hallmark of insulin resistance,[23] which is also associated with first-episode psychosis.[4] Abnormal HDL and triglyceride concentrations are associated longitudinally with cardiometabolic outcomes.[24] Guideline recommendations encourage blood-based monitoring both before and after antipsychotic exposure,[25] and so such data should be available. We found that the inclusion of blood-based predictors improved all predictive performance metrics. However, blood-based monitoring might not always be possible, and we found that the partial model still provided reliable performance estimates, although it would benefit from recalibration.

Antipsychotic medication is an important contributor to cardiometabolic risk in young people with psychosis, yet has rarely been included in existing algorithms. Some recent algorithms have included antipsychotics as predictors, grouped according to the traditional distinctions of typical and atypical or first and second generation.[8] However, the differential cardiometabolic effects of antipsychotics do not abide by these distinctions. Therefore, we instead modelled antipsychotics based on previous research (appendix p 11).

PsyMetRiC cannot yet be recommended for clinical use and requires prospective validation in larger samples, health technology assessment, and regulatory approval. However, in the future, PsyMetRiC could become a useful resource for the improved management of physical health in young people with psychosis. For example, in the presence of a very low PsyMetRiC risk score, gentle encouragement to maintain good physical health might be sufficient. This might include dietary advice or promoting daily physical activity and smoking cessation, if necessary, or both. There is little harm, yet much to gain, in offering gentle encouragement to live a healthier life, and such conversations need to become part of psychiatric consultation.

Patients and clinicians might prefer to tolerate a slightly higher threshold of risk when the proposed intervention could be deemed more burdensome or might increase the risk of other adverse effects. Regarding interventions that might be deemed more burdensome, prescribed lifestyle interventions have shown promise in lowering cardiometabolic risk in young people with psychosis,[17] but regular appointments may be difficult to maintain around work or other commitments. Regarding interventions that might increase the risk of other adverse effects, our results show that switching from metabolically active antipsychotics, or not prescribing them in the first place, is an effective means to reduce cardiometabolic risk. However, the risk of psychosis relapse or other adverse effects might reasonably be worrisome for patients and clinicians alike. Moreover, data from a meta-analysis[19] suggest that metabolically active antipsychotics could be associated with greater psychosis treatment response. Therefore, antipsychotic selection must strike an intricate balance between caring for psychiatric and physical health. Finally, trials of treatments such as metformin and statins are scarce in young people with psychosis, but evidence suggests that such medications might benefit both cardiometabolic and psychiatric outcomes.[26]

We have developed, to our knowledge, the first cardiometabolic risk prediction algorithm for young people with psychosis, harnessing data from three geographically distinct patient samples and a population-based cohort. PsyMetRiC was developed in consultation with The McPin Foundation YPAG to ensure balance between clinical practicality and patient acceptability, and we received encouraging comments from the YPAG about PsyMetRiC (appendix p 10). We developed an online interactive app permitting a visualisation of the effect of different cardiometabolic risk factors in young people with psychosis. We have published our algorithm coefficients to encourage future validation and updating. We developed two versions of PsyMetRiC to maximise clinical utility and both validated well, suggesting that PsyMetRiC is likely to be suitable for use in patients aged 16–35 years from a UK EIS population, and, from the results of our sensitivity analysis, for use in young adults at risk of developing psychosis.

Limitations of the study include missing data. We excluded participants who had the outcome at baseline, as recommended;[27] however, since predictors were assessed within a short timeframe after EIS enrolment, some metabolically sensitive individuals might have been excluded from our analysis. We also excluded participants with data missing on either all exposure or all outcome variables, which might also have introduced selection bias. The missing samples were more likely to be older and female, and less likely to be prescribed metabolically active antipsychotics. These factors might have affected some PsyMetRiC predictor coefficients. Nevertheless, we felt this exclusion step was more appropriate than imputing complete participant data. Multiple imputation can be biased when data are missing not at random, although we included auxiliary variables to reduce the fraction of missing information, limiting the effect of this bias. External validation of PsyMetRiC on larger samples is required since simulation studies have suggested a minimum of 100 outcome events for an accurate validation analysis.[28] Larger prospectively collected samples in future might also allow for updating the algorithm with interactions, non-linear terms, sex stratification, and other potentially important predictors such as other metabolically active medications, physical activity, and diet. Prospectively collected data might also predict longer-term risk. The samples in our main analysis had outcomes measured up to 6 years; however, the mean follow-up time was shorter. Although our data-driven classification of metabolically active antipsychotics is an advance over existing algorithms, the metabolically active nature of different antipsychotics lies on a continuum rather than a dichotomy. Larger samples might permit the modelling of antipsychotics individually. Prescriber bias might have downwardly biased the coefficients for antipsychotics, since metabolically active medications might have been withheld from patients considered to be at higher cardiometabolic risk.

PsyMetRiC has the potential to become a valuable resource for health-care professionals working in EISs by aiding the informed choice of antipsychotic medication, prescription of cardioprotective drugs, and non-pharmacological interventions including lifestyle adjustments to prevent the future development of cardiometabolic comorbidities and consequent years of life lost.

Contributors

BIP, RU, and GMK conceived the manuscript. BIP, EFO, RU, PKM, JY, JP, and OH acquired data. BIP, EFO, RU, PKM, JS, SZ, PBJ, and GMK analysed and interpreted the data. BIP acquired funding. BIP drafted the manuscript, with critical revision by all authors. BIP created the figures with input from EFO, PKM, and GMK. RU, PKM, PBJ, and GMK provided trial supervision. BIP and EFO had full access to, and verified the underlying data.

Declaration of interests

BIP reports a fellowship grant from the NIHR during the conduct of this study. RU reports personal fees from Sunovion, outside the submitted

work. PKM reports personal fees from Recordati and Sunovion, outside the submitted work. OH reports grants and personal fees from Angelini, Autifony, Biogen, Boehringer Ingelheim, Eli Lilly, Heptares, Global Medical Education, Invicro, Janssen, Lundbeck, Mylan, Neurocrine, Otsuka, Sunovion, Rand, Recordati, and Roche, outside the submitted work. All other authors declare no competing interests.

Data sharing

The data used in this study cannot be publicly deposited owing to patient and participant confidentiality reasons. CRIS and CRATE electronic health record data can be accessed after formal application to and ethical review by the Cambridgeshire and Peterborough NHS Foundation Trust and the South London and Maudsley NHS Foundation Trust, respectively. Access to ALSPAC data can be made following formal application to the ALSPAC executive committee. See http://www.bris.ac.uk/alspac/researchers/data-access/data-dictionary/ for a fully searchable data dictionary for ALSPAC.

Acknowledgments

BIP received funding from the NIHR (doctoral research fellowship DRF-2018-11-ST2-018). The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care. EFO is a Chain-Florey clinical PhD fellow, funded jointly by the Medical Research Council (MRC) and the NIHR Imperial BRC. GMK received funding from the Wellcome Trust (intermediate clinical fellowship grant 201486/Z/16/Z); MQ: Transforming Mental Health (data science award grant MQDS17/40); the MRC (MICA: Mental Health Data Pathfinder grant MC_PC_17213; and Therapeutic Target Validation in Mental Health grant MR/S037675/1); and the BMA Foundation (J Moulton grant 2019). PBJ received funding from the MRC and MQ (as described earlier), and programmatic funding from the NIHR (RP-PG- 0616-20003). JS, JP, and PBJ received salary support from the NIHR Applied Research Collaboration East of England. RU received funding from the NIHR (Health Technology Assessment grant 127700) and MRC (Therapeutic Target Validation in Mental Health grant MR/S037675/1). The MRC, Wellcome Trust (grant 102215/2/13/2), and the University of Bristol provide core support for ALSPAC. A comprehensive list of grants funding is available on the ALSPAC website. This research was specifically funded by the Wellcome Trust (grant 08426812/Z/07/Z), Wellcome Trust and MRC (grant 217065/Z/19/Z), and MRC (grant MR/M006727/1). SZ is supported by the NIHR BRC at University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol. Access to London EIS data were supported in part by the NIHR Cambridge BRC, and by the NIHR BRC at South London and Maudsley NHS Foundation Trust and King's College London. We thank all the families who took part in ALSPAC, the midwives for their help in recruiting them, and the whole ALSPAC team, including interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses. We also thank Rachel Temple, Anja Hollowell, and Thomas Kabir from the McPin Foundation for their kind support in facilitating the authors' attendance at YPAG meetings, and all participants of the YPAG for their enthusiasm for the project and insightful comments. Finally, we thank Megan Pritchard (SLaM BRC Nucleus) and Jonathan Lewis (Cambridgeshire and Peterborough NHS Foundation Trust Database Manager) for their invaluable help and support in setting up the local database search queries.

References

1 Plana-Ripoll O, Pedersen CB, Agerbo E, et al. A comprehensive analysis of mortality-related health metrics associated with mental disorders: a nationwide, register-based cohort study. *Lancet* 2019; **394:** 1827–35.

2 Firth J, Siddiqi N, Koyanagi A, et al. The *Lancet Psychiatry* Commission: a blueprint for protecting physical health in people with mental illness. *Lancet Psychiatry* 2019; **6:** 675–712.

3 Naylor C, Parsonage M, McDaid D, Knapp M, Fossey M, Galea A. Long-term conditions and mental health: the cost of co-morbidities. London: The King's Fund and Centre for Mental Health, 2012.

4 Perry BI, McIntosh G, Weich S, Singh S, Rees K. The association between first-episode psychosis and abnormal glycaemic control: systematic review and meta-analysis. *Lancet Psychiatry* 2016; **3:** 1049–58.

5 Pillinger T, Beck K, Stubbs B, Howes OD. Cholesterol and triglyceride levels in first-episode psychosis: systematic review and meta-analysis. *Br J Psychiatry* 2017; **211:** 339–49.

6 De Hert M, Detraux J, Vancampfort D. The intriguing relationship between coronary heart disease and mental disorders. *Dialogues Clin Neurosci* 2018; **20:** 31–40.

7 Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *BMJ* 2017; **357:** j2099.

8 Perry BI, Upthegrove R, Crawford O, et al. Cardiometabolic risk prediction algorithms for young people with psychosis: a systematic review and exploratory analysis. *Acta Psychiatr Scand* 2020; **142:** 215–32.

9 Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med* 2015; **162:** 55–63.

10 Cardinal RN. Clinical records anonymisation and text extraction (CRATE): an open-source software system. *BMC Med Inform Decis Mak* 2017; **17:** 50.

11 Boyd A, Golding J, Macleod J, et al. Cohort profile: the 'children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* 2013; **42:** 111–27.

12 Alberti KG, Eckel RH, Grundy SM, et al. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation* 2009; **120:** 1640–45.

13 Steyerberg EW, Eijkemans MJ, Harrell FE Jr, Habbema JD. Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. *Stat Med* 2000; **19:** 1059–79.

14 Riley RD, Snell KI, Ensor J, et al. Minimum sample size for developing a multivariable prediction model: part II—binary and time-to-event outcomes. *Stat Med* 2019; **38:** 1276–96.

15 Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006; **26:** 565–74.

16 Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. *Diagn Progn Res* 2019; **3:** 18.

17 Fernandez-Abascal B, Suarez-Pinilla P, Cobo-Corrales C, Crespo-Facorro B, Suarez-Pinilla M. In- and outpatient lifestyle interventions on diet and exercise and their effect on physical and psychological health: a systematic review and meta-analysis of randomised controlled trials in patients with schizophrenia spectrum disorders and first episode of psychosis. *Neurosci Biobehav Rev* 2021; **125:** 535–68.

18 Isomaa B, Almgren P, Tuomi T, et al. Cardiovascular morbidity and mortality associated with the metabolic syndrome. *Diabetes Care* 2001; **24:** 683–89.

19 Pillinger T, McCutcheon RA, Vano L, et al. Comparative effects of 18 antipsychotics on metabolic function in patients with schizophrenia, predictors of metabolic dysregulation, and association with psychopathology: a systematic review and network meta-analysis. *Lancet Psychiatry* 2020; **7:** 64–77.

20 Reinikainen J, Laatikainen T, Karvanen J, Tolonen H. Lifetime cumulative risk factors predict cardiovascular disease mortality in a 50-year follow-up study in Finland. *Int J Epidemiol* 2015; **44:** 108–16.

21 Osborn DP, Hardoon S, Omar RZ, et al. Cardiovascular risk prediction models for people with severe mental illness: results from the prediction and management of cardiovascular risk in people with severe mental illnesses (PRIMROSE) research program. *JAMA Psychiatry* 2015; **72:** 143–51.

22 Misiak B, Stanczykiewicz B, Laczmanski L, Frydecka D. Lipid profile disturbances in antipsychotic-naive patients with first-episode non-affective psychosis: a systematic review and meta-analysis. *Schizophr Res* 2017; **190:** 18–27.

23 Murguia-Romero M, Jimenez-Flores JR, Sigrist-Flores SC, et al. Plasma triglyceride/HDL-cholesterol ratio, insulin resistance, and cardiometabolic risk in young adults. *J Lipid Res* 2013; **54:** 2795–99.

For **ALSPAC grant details** see http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf

24 Nordestgaard BG, Varbo A. Triglycerides and cardiovascular disease. *Lancet* 2014; **384:** 626–35.

25 Barnes TR, Drake R, Paton C, et al. Evidence-based guidelines for the pharmacological treatment of schizophrenia: updated recommendations from the British Association for Psychopharmacology. *J Psychopharmacol* 2020; **34:** 3–78.

26 Hayes JF, Lundin A, Wicks S, et al. Association of hydroxylmethyl glutaryl coenzyme A reductase inhibitors, L-type calcium channel antagonists, and biguanides with rates of psychiatric hospitalization and self-harm in individuals with serious mental illness. *JAMA Psychiatry* 2019; **76:** 382–90.

27 Wolff RF, Moons KGM, Riley RD, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med* 2019; **170:** 51–58.

28 Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. *Stat Med* 2016; **35:** 214–26.