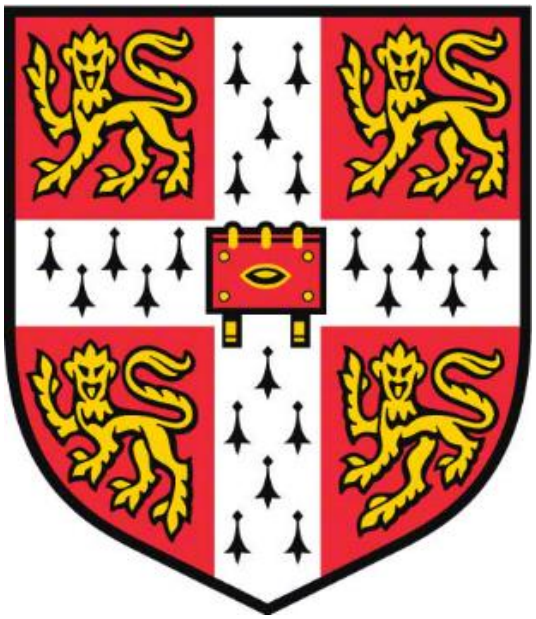
The role of the sociocultural context in explaining variance in incidence of psychosis and higher rates of disorder in minorities.

Hannah Elisabeth Jongsma

Darwin College



This thesis is submitted to the Degree Committee for the Faculties of Clinical Medicine and Veterinary Medicine for the degree of Doctor of Philosophy.

March 2018

Summary

The role of the sociocultural context in explaining variance in incidence of psychosis and higher rates of disorder in minorities.

Hannah E. Jongsma

Over the past few decades, epidemiological evidence has accrued to establish variance in psychosis risk across both geographical locations and demographic characteristics such as the excess risk in migrants and their descendants. Yet, the causes of this variation in rates between places and ethnic groups are still unclear, and I aimed to address this in this thesis.

I conducted a systematic review and meta-analyses to synthesise existing literature on psychosis incidence in the six countries included in the EUropean network of national schizophrenia networks studying Gene-Environment Interactions (EU-GEI) study: England, The Netherlands, Spain, France, Italy and Brazil. I subsequently analysed data from two parts of the EU-GEI study: a 17-centre service-based incidence study of psychosis, and a case-control arm utilising community volunteers. In the latter, I aimed to explain excess risk in ethnic and religious minorities using a theoretical sociocultural distance model I developed using literature from the social sciences. Here, I proposed that culturally distant minorities were particularly at risk of social exclusion, and this outsider experience led to increased psychosocial disempowerment (a lack of control over one's life), which increased psychosis risk. I also explored if this model could explain any excess risk in those with increased genetic African ancestry in England.

Incidence varied substantially between the studies included in the systematic review, although methodological differences could not be excluded as an explanation. The EU-GEI incidence study confirmed substantial variation by place, and demonstrated a higher incidence in ethnic minorities and for young men, as well as in areas characterised by a low percentage of owner-occupied housing. The sociocultural distance model could explain most of the excess psychosis risk in ethnic minorities, although some excess risk remained, particularly in the Black ethnic group. Social and cultural distance appeared to be more important predictors than psychosocial disempowerment, suggesting that chronic social injustices rather than acute stress play an important role. This model did not explain excess risk in religious minorities: those following any religion retained an excess risk. It could explain the excess risk in those with increased genetic African ancestry, although this was a small, exploratory sample and this will need replicating in larger studies.

This thesis demonstrated, for the first time, that excess risk in ethnic minorities could be explained by the sociocultural distance model. Overall, the findings from this thesis confirm substantial variation in psychosis risk by person and place, and suggest that the social reality of the environment plays a crucial role in explaining this.

Is Obama Anything but Black?

So lots of folk – mostly non-Black – say Obama's not Black, he's biracial, multiracial, Black-and-White, anything but just Black. Because his mother was White. But race is not biology; race is sociology. Race is not genotype; race is phenotype. Race matters because of racism. And racism is absurd because it's about how you look. Not about the blood you have. It's about the shade of your skin and the shape of your nose and the kink of your hair. Booker T. Washington and Frederick Douglass had White fathers. Imagine them saying they were not Black.

Imagine Obama, skin the color of a toasted almond, hair kinky, saying to a census worker- I'm kind of white. Sure you are, she'll say. Many American Blacks have a White person in their ancestry, because White slave owners liked to go a-raping in the slave quarters at night. But if you come out looking dark, that's it. (So if you are that blond, blue-eyed woman who says "My grandfather was Native American and I get discrimination too" when black folk are talking about shit, please stop it already.) In America, you don't get to decide what race you are it is decided for you. Barack Obama, looking as he does, would have had to sit in the back of the bus fifty years ago. If a random Black guy commits a crime today, Barack Obama could be stopped and questioned for fitting the profile. And what would that profile be?

"Black Man".

Chimamanda Ngozi Adichie (2013)

Contents

List of Tables

List of Figures

List of Abbreviations

AESOP	Aetiology and Ethnicity in Schizophrenia and Other Psychoses study
AIC	Akaike Information Criterion
ALSPAC	Avon Longitudinal Study of Parents and Children
AMSTAR	Assessing Methodological Quality of Systematic Reviews
APA	American Psychiatric Association
CASH	Comprehensive Assessment of Symptoms
CI	Confidence Interval
DIGS	Diagnostic Interview for Genetic Studies
DSM	Diagnostic and Statistical Manual
DUP	Duration of Untreated Psychosis
ELFEP	East London First Episode Psychosis study
EU-GEI	EUropean network of national schizophrenia networks studying Gene-Environment Interactions study
FEP	First Episode of Psychosis
GWAS	Genome-Wide Association Study
HR	Hazard Ratio
ICD	International Classification of Disease
IQR	Interquartile Range
IR	Incidence Rate
IRR	Incidence Rate Ratio
LAMI	Low and Middle Income
LRT	Likelihood Ratio Test
MAR	Missing at Random
MCAR	Missing Completely at Random
MNAR	Missing Not at Random
MRC	Medical Research Council
NICE	National Institute for Health and Care Excellence
NOS	Not Otherwise Specified (psychosis)
NUTS	Nomenclature of Territorial Units for Statistics
OPCRIT	Operational Criteria Checklist
OR	Odds Ratio
PACE	Personal Assessment and Crisis Evaluation (study)
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SCAN	Schedules for Clinical Assessment in Neuropsychiatry
SCID	Structured Interview of DSM-IV
SIR	Standardised Incidence Ratio
WHO	World Health Organization
WP2	Work Package 2: Functional Enviromics (of the EU-GEI study).
WVS	World Values Survey

Author abbreviations

HEJ	Hannah Jongsma
JBK	James Kirkbride
PBJ	Peter Jones

Declaration of authorship

I hereby declare that this thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration, except when specifically indicated in the text and acknowledgement.

This thesis makes extensive use of data from the EU-GEI study, a large international collaboration. I have had no role in study design, data collection or data entry. I have cleaned all demographic data (age, sex, minority status) for the analyses presented in Chapter 4 in this thesis, but not the diagnostic data. For the analyses presented in Chapter 6 and 7, I contributed to data cleaning for the following variables: age, sex, ethnicity, discrimination (all in cooperation with Charlotte Gayer-Anderson), paternal age, first language, fluency in the majority language and religious affiliation (all entirely my responsibility). All other variables were cleaned and prepared for analysis by others. The genetic data used in Chapter 7 was prepared by Alexander Richards and colleagues. All analyses presented in this thesis are entirely my own. Any errors in this thesis are entirely my own.

This thesis is not substantially the same as any work I have submitted or is being concurrently submitted for a degree or other qualification at the University of Cambridge or any other University or similar institution. 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.

This thesis has resulted in one peer-reviewed publication (Chapter 4):

Jongsma, H.E., Gayer-Anderson, C., Lasalvia, A., et al (2017). Treated Incidence of Psychotic Disorders in the Multinational EU-EI Study. *JAMA Psychiatry.* DOI:10.1001/jamapsychiatry.2017.3554

This thesis does not exceed the word limit set by the Degree Committee for the Faculties of Clinical Medicine and Veterinary Medicine.

Acknowledgements

I am deeply indebted to both my supervisors; James Kirkbride and Peter Jones. Your mentoring, support, guidance and patience are much appreciated, and I continue to be grateful for the trust you put in me, the belief you have in my and my work, and the freedom you give me to pursue my ideas. I am also grateful for the community offered by both your research groups, here in Cambridge and the PsyLife group in London. I am particularly grateful to Jan Stochl (Cambridge) for statistical advice and Francesca Solmi (UCL) for advice on multiple imputation.

This thesis would not have been possible without all the participants and researchers involved in the EU-GEI study. I would particularly like to thank Gonzalo Lopez, Rosana Shuhama, Elsje van der Ven, Ilaria Tarricone, Chiara Bonetto, Fabio Seminerio and Andrei Szöke for their help in retrieving missing data. I am particularly grateful to Diego Quattrone, who uncovered the vast majority of OPCRIT diagnoses for the incidence study, to Craig Morgan who continues to lead WP2, and to Charlotte Gayer-Anderson for her time and effort spent on cleaning and managing the data.

I am grateful to Robin Murray and Jean-Paul Selten for their feedback on my research ideas for Chapter 6 and 7 when they were circulated to the EU-GEI group. I owe a debt of gratitude to Alexander Richards at the University of Cardiff for his patient explanations on genetic ancestry and his feedback on the work presented in Chapter 7. For Chapter 5, Section 5.2 in particular, I am grateful to Katherine Furman for allowing me to read her PhD thesis. It provided a brilliant refresher of some of Nancy Cartwright's and Katie Steele's lectures on causality we both attended during our MSc in philosophy, and a very useful insight in how to think about my research as a causal selection problem. For the systematic review presented in Chapter 2, I am grateful to the Medical School Librarian, Isla Kuhn, for her assistance in adapting the search strategy to the various databases. I also like to acknowledge the work of Caitlin Turner (a former MPhil in Public Health candidate) on the work she's done to expand this systematic review beyond the EU-GEI context (not included in this thesis).

I am grateful to the European Seventh Framework Programme and NIHR CLAHRC East of England for funding my PhD, and to the Fearnside Fund, Guarantors of Brain and my SIRS Travel Fellowship for providing me with the opportunity to present my work at scientific conferences.

Last, but not least, I'd like to acknowledge the support of my friends and family (particularly those who sent me chocolate). My parents have fostered my curiosity and independence from a young age, and I wouldn't have been there if it wasn't for this. Thanks to my group group (Abi, Becca, Falk, Laura, Pauline & Sabine) for having my back when I most needed it. And Adam, your unwavering support and belief in me and this thesis were infectious, and I'm really grateful you came into my life.

Chapter 1 – General Introduction

In this thesis, I investigate heterogeneity in incidence rates of psychotic disorders and examine the role of the social environment in explaining both this variation and excess risk in (ethnic) minority groups. Before doing so, this introduction introduces and reviews the literature on psychotic disorders and their various diagnostic classifications, and provides a broad overview of their general epidemiology. This Chapter also gives a brief description of migration during the 20th century in England, the Netherlands, France, Spain, Italy and Brazil, as these are the countries include in the study I will analyse data from. I also give an overview of the association between ethnicity, migration and health, and provide a detailed discussion of the epidemiology of psychotic disorders in migrants and their descendants. Finally, I clarify the gap in the literature I seek to address and the particular intellectual viewpoint I take through this thesis.

This thesis will make use of data from the EUropean network of national schizophrenia networks studying Gene-Environment Interactions (EU-GEI) study, an international incidence and case-sibling-control study of the determinants and outcomes of psychotic disorders. Full details of this study are given in Chapter 3.

1.1 Introduction to psychotic disorders

In this Section, I introduce the psychotic disorders and describe the various nosological systems that are used to differentiate between disorders.

1.1.1 Description and diagnostic classification

Psychotic disorders constitute an umbrella term used for a group of disorders characterised by hallucinations, delusions and disordered thought structure (formal thought disorder). The latter is usually betrayed through disrupted speech construction. Hallucinations are perceptions in clear consciousness that have no sensory input, and can occur in any modality. Some are particularly common in the subtype of psychosis called schizophrenia such as hearing voices in the third person, or hearing one's thoughts being spoken out loud. Delusions are persistent, false beliefs not in keeping with a person's cultural or social background (Fletcher & Frith, 2009). Most are self-referential and persecutory but some are more bizarre; for example, a person may believe that their spouse has been replaced by an imposter.

Psychotic disorders are generally divided into so-called affective and non-affective psychotic disorders, depending on the relative prominence of mood symptoms (the effect on affect). Delusions, hallucinations and thought disorder are known as positive symptoms (characterised by their presence), whereas lack of motivation and inability to experience pleasure (avolition and anhedonia), are known as negative symptoms and characterised by absence of normal function. Mood (affect) disturbance often determines the diagnostic category into which a psychotic state is classified. Depressed mood and elation are aspects of the definition of depressive psychosis and bipolar 1 disorder, whereas an absence of affective reactivity (blunted affect) is often seen in schizophrenia. Although not part of the definition of psychotic disorders, cognitive impairments are common, disabling and may precede the onset of positive features (Barnett et

al., 2005; Bora, Yucel, & Pantelis, 2009; Fioravanti, Carlone, Vitale, Cinti, & Clare, 2005; Khandaker, Barnett, White, & Jones, 2011).

The two nosological systems currently in widespread use are the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-5) and the International Statistical Classification of Disease and Health Related Problems, 10th revision (ICD-10). The former is produced by the American Psychiatric Association (APA) and as such is mainly used in the USA (American Psychiatric Association, 2013), whereas the latter is produced by the World Health Organization (WHO) and is used globally (World Health Organization, 1992a) and throughout this PhD thesis. In ICD-10, the psychotic disorders are found in Chapter V 'Mental and Behavioural Disorders'; a full list can be found in Table 1.1 below.

Table 1.1: ICD-10 diagnoses of psychotic disorders

Diagnosis	ICD-10 code	Description
Substance-induced psychoses	F1X.5	Includes psychosis induced by a range of substances including alcohol, opioids, cannabinoids and hallucinogens.
Schizophrenia	F20	Characterised by fundamental and characteristic distortions of thinking and perception, and inappropriate or blunted affect. Includes paranoid, hebephrenic, catatonic, undifferentiated, residual, simple and other schizophrenia as well as post-schizophrenic depression
Schizotypal disorders	F21	Characterised by eccentric behaviour and anomalies of thinking and affect resembling schizophrenia, but without definite schizophrenic anomalies
Persistent delusional disorder	F22	Where long-standing delusions as the only, or most conspicuous, clinical characteristic which can't be classified as organic, schizophrenia or affective.
Acute and transient psychotic disorders	F23	Acute onset of psychotic symptoms (less than two weeks), usually with complete recovery within a few months
Induced delusional disorder	F24	Delusional disorder shared by two or more people with close emotional links where one suffers from genuine psychotic disorder and delusions are induced in the other
Schizoaffective disorder	F25	Episodic disorders where both affective and schizophrenic symptoms are prominent but do not justify other diagnosis
Other nonorganic psychotic disorders	F28	Delusional or hallucinatory disorders that do not justify other diagnosis
Unspecified nonorganic psychosis	F29	Psychosis NOS
Manic episode	F30	Elevated mood out of keeping with circumstances, either with or without delusions or hallucinations or excessive excitement, motor activity and flight of ideas.
Bipolar affective disorder	F31	Current manic episode (as above) and has had at least one other affective episode (hypomanic, manic, depressive or mixed) in the past.
Depressive episode	F32	Lowering of mood, reduction of energy and decrease in activity (mild, moderate or severe, depending on number of symptoms) with or without the presence of hallucinations, delusions, psychomotor retardation or severe stupor.
Recurrent depressive disorder	F33	Disorder characterised by repeated episodes of depression without any history of independent manic episodes.

Whilst psychosis has been described as far back as 1550 BC (Gold & Gold, 2014), the origins of contemporary classification systems can be traced back to Emil Kraepelin in the late nineteenth century (Kraepelin, 1899). Kraepelin, through empirical study of patients with largely long-term illness, grouped psychopathology into two categories, reflecting what he believed to be different underlying biological disease entities. His *dementia praecox* is the basis for the current non-affective disorders, and his *manic*

depressive psychoses are now known as bipolar disorder (Kraepelin, 1899). The term *group of schizophrenias* was first introduced by Eugen Bleuler in 1911 (Bleuler, 1911), and Kurt Schneider introduced what he called first-rank symptoms of schizophrenia in 1957 (Schneider, 1957). These included auditory hallucinations, passivity experiences, thought withdrawal, insertion and broadcasting as well as delusional perceptions. Schneider believed these to be particularly characteristic of schizophrenia (Schneider, 1957). Together, these continue to be the basis for present-day diagnostic classifications.

1.2 Epidemiology of psychotic disorders

In this Section, I give a broad overview of the epidemiology of psychotic disorders. I mention the incidence of psychosis, and discuss the prevalence of psychosis as well as its associated morbidity and mortality. I will discuss the incidence in more detail in Chapters 2 and 4. I also discuss both the genetic and environmental determinants of psychosis in the general population, and will turn to the link between migration and mental health more specifically in the next Section.

1.2.1 Incidence, prevalence, mortality and morbidity

Psychotic disorders have a low incidence: the most recent (prior to this thesis) systematic review of the literature showed a median incidence of schizophrenia of 15.2 per 100,000 person-years (10^{th}-90^{th} percentile: 7.7-43.0) (McGrath et al., 2004). Nonetheless, psychotic disorders contribute substantially to the global burden of disease (Whiteford, Ferrari, Degenhardt, Feigin, & Vos, 2015). This is largely for two reasons: psychotic disorders have an early age of onset compared with many other non-communicable diseases and a substantial proportion of those with a first episode of psychosis (FEP) face a chronic or intermittent course of illness (Owen, Sawa, & Mortensen, 2016; Saha, Chant, Welham, McGrath, & Lapsley, 2005).

Prevalence is the epidemiological measure used to indicate the proportion of the population that have a disorder at a specified time (point prevalence) or during a specified period (yearly prevalence or lifetime prevalence, for instance). The lifetime prevalence of schizophrenia is often quoted to be around 1%, and this is likely traced back to the DSM-IV, where it was indicated to be between 0.5 and 1% (American Psychiatric Association, 1994). The most recent international systematic review of the existing literature showed a median point prevalence of schizophrenia of 0.5% (10^{th} percentile: 0.2%, 90^{th} percentile: 1.0%), a median lifetime prevalence of 0.4% (10^{th} percentile: 0.2%, 90^{th} percentile: 1.2%) and a median lifetime morbid risk of 0.7% (10^{th} percentile: 0.3%, 90^{th} percentile: 2.7%) (Saha et al., 2005). This concept of lifetime morbid risk is conceptually similar to that of a cumulative incidence in a birth cohort. These estimates are similar to those found by an earlier narrative review of the evidence (Torrey, 1987), and similar to estimates of the prevalence of rheumatoid arthritis in North America and Northern Europe (Alamanos & Drossos, 2005). The prevalence of bipolar I disorder (a DSM diagnosis including at least one severe manic episode (American Psychiatric Association, 2013)) is similar. A Swedish study showed an annual prevalence of 1% (Carlborg, Ferntoft, Thuresson, & Bodegard, 2015), and a systematic review

established a pooled annual prevalence of 0.72%, and a lifetime prevalence of 0.8% (Waraich, Goldner, Somers, & Hsu, 2004). A Finnish study, using a combination of clinical interviews and registry based diagnoses (the latter for non-responders), demonstrated a lifetime prevalence of schizophrenia of 0.87%, of bipolar I disorder of 0.24%, and of all psychotic disorders of 3.5% (Perälä et al., 2007), illustrating the importance of the remaining non-affective and affective disorders in contributing to the overall lifetime prevalence of all psychotic disorders. Whilst this study is likely to be highly accurate, it might not be generalisable: Northern European countries have a higher prevalence of psychotic disorders, see for instance Saha, Chant, Welham, & McGrath, 2006.

The morbidity and mortality associated with psychotic disorders are immense. Premature mortality is substantially increased compared with the general population (Chang et al., 2011; Hoang, Stewart, & Goldacre, 2011; Saha, Chant, & McGrath, 2007; Tiihonen et al., 2009), with a recent systematic review estimating an average reduced life-expectancy of 14.5 years (Hjorthøj, Stürup, McGrath, & Nordentoft, 2017) and a further recent UK-based cohort study concluding that the mortality gap between the general population and individuals with bipolar disorder and schizophrenia was widening as general population life expectancy increases (Hayes, Marston, Walters, King, & Osborn, 2017). Cardiovascular disease appears to be the leading cause of death in individuals with severe mental illness (Hayes et al., 2017; Hennekens, Hennekens, Hollar, & Casey, 2005; Laursen et al., 2013; Weiner, Warren, & Fiedorowicz, 2011), and increased levels of suicidality appeared to be the cause of death most elevated relative to the general population (although the absolute rates of suicide are low compared with other common causes of death) (Brown, 1997; Hayes, Miles, Walters, King, & Osborn, 2015).

Individuals with a psychotic disorders face not only a reduced life expectancy, but also increased levels of co-morbidity and poorer quality of life. Although many recover from an initial FEP, many individuals face a chronic or intermittent disease trajectory (Owen et al., 2016). This is associated with poorer physical (Daumit et al., 2006; Leucht et al., 2013; Singleton, Bumpstead, O'Brien, Lee, & Meltzer, 2003) and lifestyle (Vancampfort et al., 2012) outcomes as well as a markedly poorer social outcomes. Many individuals with psychosis remained unemployed, even after making a symptomatic recovery (Marwaha & Johnson, 2004; Revier et al., 2015).

1.2.2 Determinants of psychotic disorders

Non-organic psychotic disorders are multi-causal: no single underlying cause has been identified. Determinants of psychosis are often described as part of a stress-vulnerability model: a genetic and early-life vulnerability combined with environmental risk factors (Zubin & Spring, 1977). In this Section I address these genetic, early life, and environmental risk factors, and I will address the issue of causality in psychosis from a theoretical perspective and in more detail in Chapter 5.

1.2.2.1 Genetic risk and psychotic disorders

The genetic risk for the psychoses is complex and, with the exception of rare syndromes such as 22q11 deletion syndrome (Baker & Skuse, 2005; Bassett et al., 2003, 2005), involves a large number of genes each carrying a small effect size (Craddock & Sklar, 2013; Harrison, 2015; Schizophrenia Working Group of the Psychiatric Genomic Consortium, 2015). Schizophrenia for instance, is highly heritable with 50% monozygotic twin concordance (Dean, 2012). Recent advances in genetic research, particularly the increased availability and decreased cost of carrying out genome-wide association studies (GWASs), mean that understanding of genetics of psychotic disorders has improved substantially (Vassos et al., 2017). Yet, heritability is considered complex and non-Mendelian (Giegling et al., 2017). There is also considerable overlap in genetic risk between schizophrenia and bipolar I disorder (Craddock & Sklar, 2013; The International Schizophrenia Consortium, 2009) and other psychiatric disorders (Cross-Disorder Group of the Psychiatric Genetics Consortium, 2013). The most recent meta-analysis of GWASs of schizophrenia has identified 108 independently associated loci (places on the genome) (Schizophrenia Working Group of the Psychiatric Genomic Consortium, 2015), which each explained a very small part of the genetic variance. The largest meta-analysis of bipolar-I GWASs has established 56 independently associated loci across five regions (Mühleisen et al., 2014), and further study estimates there might be as many as 105 loci (Chen et al., 2013), but these have not yet been discovered (largely, due to the large sample sizes needed for such discoveries).

1.2.2.2 A life course approach: prenatal and childhood risk factors.

Psychotic disorders are thought to have neurodevelopmental origins (Howes & Murray, 2014), and risk factors for development of psychosis start in utero. Severe maternal malnutrition (Susser & Lin, 1992) and prenatal exposure to infection increase risk of later disorder (Khandaker, Zimbron, Lewis, & Jones, 2013). Obstetric complications and other perinatal hazards such as complications in pregnancy, abnormal fetal growth and complications during labour (Brown & Derkits, 2010; Cannon, Jones, & Murray, 2002), markers of disordered motor development and neurodevelopment in childhood (Bramon et al., 2005; Jones, Rodgers, Murray, & Marmot, 1994) and cognitive problems predating the onset of disorder (Reichenberg et al., 2010) have also been associated with increased risk of disorder. High educational attainment in childhood is considered protective against the development of psychosis (Khandaker, Stochl, Zammit, Lewis, & Jones, 2014; Khandaker et al., 2011), and study authors have suggested this might link to neurodevelopmental origins of disorder and the existence of a 'cognitive reserve' which increases available coping strategies (Khandaker, Stochl, et al., 2014).

Childhood adversities are also strongly associated with later psychosis risk (Bendall, Jackson, Hulbert, & McGorry, 2008; Matheson, Shepherd, Pinchbeck, Laurens, & Carr, 2012; Varese et al., 2012), including emotional, psychological abuse, physical and sexual abuse, neglect, peer victimisation and bullying (Varese et al., 2012). A longitudinal study (n<1,000 adolescents) suggested a bidirectional relation between

childhood trauma and psychotic experiences in adolescence: trauma both predicted these experiences and later trauma was a consequence of them. After accounting for this bidirectionality, trauma was still strongly predictive of psychotic experiences in a dose-response relationship, suggesting a causal role (Kelleher et al., 2013).

An emerging body of evidence is accumulating regarding the association between increased childhood inflammatory markers (interleukin 6 and C-reactive protein) and psychotic experiences and psychosis in adolescence and adulthood (Khandaker, Pearson, Zammit, Lewis, & Jones, 2014; Khandaker, Zammit, Lewis, & Jones, 2014), indicating a potential important role for inflammation in the pathogenesis of psychotic disorders. This may not only provide new targets for treatment, but could also be the mechanism behind high comorbidity between psychotic disorders and cardiovascular disease (Khandaker, Pearson, et al., 2014).

1.2.2.3 Individual socio-environmental risk factors

These early life vulnerabilities are not the only determinants of psychosis risk; known adolescent or adult stressors are also important, predominantly relating to the social environment. Being of an ethnic minority background is a long-standing (Ødegaard, 1932) and widely replicated risk factor for developing a psychotic disorder (Bourque, van der Ven, & Malla, 2011) (see Section 1.4)). The association between urban birth, upbringing and residence and excess risk of psychosis is also frequently reported (Vassos, Pedersen, Murray, Collier, & Lewis, 2012). Further adult risk factors in the social domain include unemployment (Reininghaus et al., 2008) and social isolation (or lack of social support) (Gayer-Anderson & Morgan, 2013; Kohn & Clausen, 1955; Reininghaus et al., 2008). Causality of these risk factors is difficult to determine, particularly when using case-control data: the prodromal phase of illness (where symptoms might already manifest, but before a full-blown psychotic episode) can be lengthy, and might involve increased isolation or downward social mobility. Cannabis use is implicated too (Arseneault et al., 2002; Henquet, Murray, Linszen, & Van Os, 2005; Manrique-Garcia et al., 2012; Moore et al., 2007) - particularly frequent use (Moore et al., 2007), and use from an early age (Arseneault et al., 2002).

Recent advances in genetic methodology have however suggested that some risk factors might be subject to a degree of gene-environment selection. Individuals with a higher genetic risk of schizophrenia might be more likely to move to more densely populated areas (Colodro-Conde et al., 2017; Sariaslan et al., 2015), and might be more likely to develop a substance use disorder (Hartz et al., 2017). The latter was tested using nicotine, alcohol and cocaine dependence, so did not include cannabis use.

1.2.2.4 Neighbourhood-level risk factors

Complimenting this research on individual-level risk factors is a growing body of evidence on neighbourhood-level risk factors. This research tenet started in the early 20th century, with Faris and Dunham's seminal work in Chicago (Faris & Dunham, 1939). They established that neighbourhoods with

increased residential mobility and percentage of African-Americans and those of no fixed abode faced an increased incidence of psychosis (Faris & Dunham, 1939). This work was later replicated in Bristol (Hare, 1956). More recent studies have examined ethnic density: the size of a particular ethnic group in proportion to the total population in a specified area (Bhugra et al., 2011). A low ethnic density is hypothesised to be associated with increased incidence of psychotic disorders in ethnic minorities (Bécares, Nazroo, & Stafford, 2009; Boydell et al., 2003). It is thought that a high ethnic density is associated with strong support networks, alleviating the effects of discrimination (Bécares et al., 2009). Results of recent epidemiological investigations are mixed: some studies established a relationship (Boydell et al., 2003; Schofield, Ashworth, & Jones, 2011; Veling, Susser, van Os, et al., 2008), but others showing a more nuanced picture of differential effects by ethnic group (Kirkbride, Jones, Ullrich, & Coid, 2014).

Other neighbourhood-level risk factors also concern the social (as opposed to the physical) environment. For instance, increased socio-economic deprivation is associated with increased incidence (Kirkbride et al., 2014; Lasalvia et al., 2014). High crime and low education (Bhavsar, Boydell, Murray, & Power, 2014) appear to be particularly important domains of deprivation. Income inequality within neighbourhoods (Burns & Esterhuizen, 2008; Kirkbride et al., 2014) and social fragmentation (Allardyce et al., 2005) and disorganisation (Veling, Susser, Selten, & Hoek, 2015) appear to be associated with increased incidence too. These studies are cross-sectional, so any observed associated could at least partially be due to downward social drift whereby individuals in the prodromal stages of disorder move to neighbourhoods with higher levels of the exposure of interest.

1.3 Ethnicity, migration and (mental) health

A large part of this thesis is concerned with the excess psychosis risk in migrants and their descendants. In this section, I introduce how ethnicity can be measured, and I give a brief migration history of the six countries included in the EU-GEI study (England, the Netherlands, Spain, France, Italy and Brazil) that forms the basis of my work. I also introduce the effects migration has on the physical and mental health of those who undergo it, and in the next Section discuss the epidemiological evidence surrounding higher rates of psychotic disorders in migrants and their descendants.

1.3.1 Race, ethnicity, migration and skin colour

There are various ways of differentiating migrants and their descendants from the majority population for the purposes of studying differences in morbidity and mortality across population groups, the two most common being ethnicity and race. The concept of race is used to divide humankind into subspecies on the basis of biological characteristics of physical attributes, supposedly reflecting differences in genetic make-up (Afshari & Bhopal, 2002; Kelly & Nazroo, 2008). There no scientific basis for this, as all humans belong to a single species. The genetic variations between racial groups are small, and the physical features used to distinguish between races are not associated with disease or behaviours (Afshari & Bhopal, 2002). Furthermore, this term also has a long history of abuse and racist motives (Bhopal, 1997; Kelly & Nazroo,

2008; Senior & Bhopal, 1994), making uncritical use of this term problematic (Cooper & David, 1986). Demonstrating the superiority of the white race was the leading motive in both 19th century phrenology and Nazi eugenics, for instance.

For these reasons the concept of race has become unfavourable, and the concept of ethnicity has gained popularity. Ethnicity reflects a social identity: a social group an individual belongs to and either self-identifies with, or is identified with by others. This is a fluid, imprecise and dynamic classification, to some extent depending on context (Kelly & Nazroo, 2008; Bhopal, 1997). My understanding of ethnicity for the purposes of this thesis is a group that sees itself, or is seen by others, as a distinct group, and has some extent of common origins, ancestry or social background, a shared culture, language, traditions or history that are maintained between generations and provide a sense of identity (Ahmad & Bradby, 2007; Hutchinson & Smith, 1996; Senior & Bhopal, 1994). Ethnicity can be measured in a number of ways: skin colour, birthplace, name, geographical origins, self-ascribed ethnicity or a combination of these measurements (Bhopal, 1997). Skin colour is based on race, subjective, and an imprecise measurement. It fails to differentiate between similar looking but culturally different minorities, such as Moroccan and Turkish minorities in the Netherlands. Using birthplace fails to identify second and later-generation migrants and conversely, if one is born to, for example, English parents in India, but migrated to England, one is not commonly considered part of an ethnic minority. Currently, in the UK the preferred way of measurement is self-ascribed ethnicity as this accurately reflects the view that ethnicity is largely a matter of self-perception (Bhopal, 2004), although different EU-GEI countries follow different customs (see Sections 4.3.1, 6.3.1 and 7.3.4 for the measurements of ethnicity in this thesis).

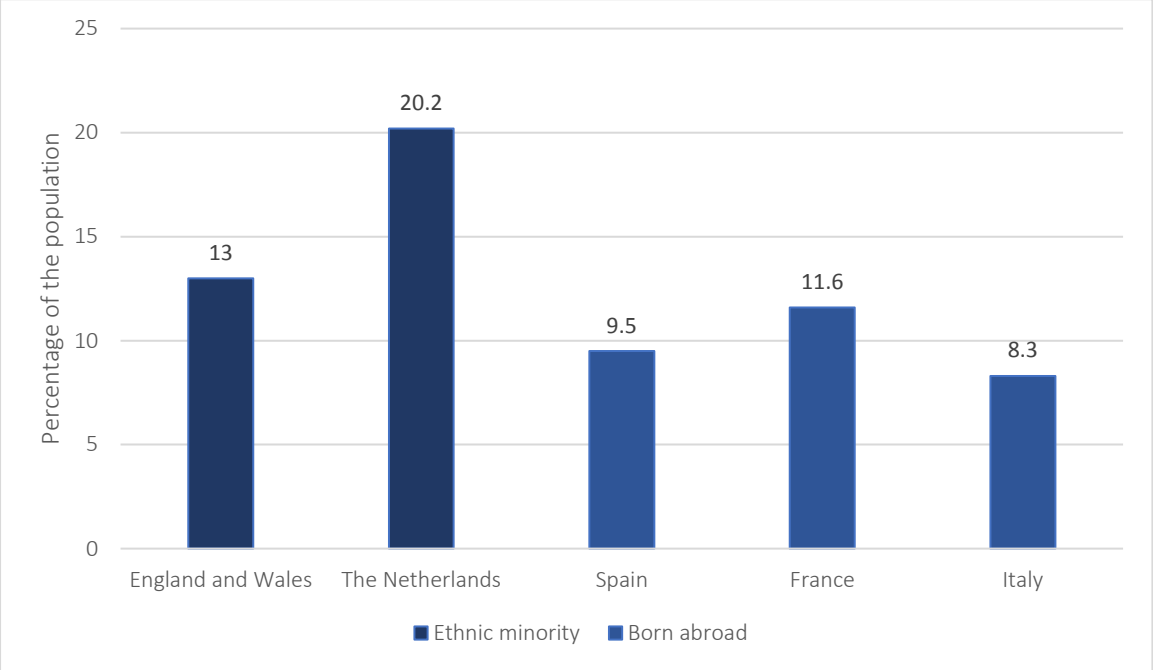
Regardless of the definition of ethnicity used, within what are considered ethnic groups in epidemiology there is often a lot of heterogeneity (Senior & Bhopal, 1994; Williams, Costa, & Leavell, 2017), and this applies to the EU-GEI study too. Examples of this are the 'Asian' group, which in England tends to include everyone from India, Pakistan or Bangladesh (although in other countries this might refer to individuals from China or Japan), or a 'Black' group, which is derived from skin colour but is now used to indicate an ethnicity (Senior & Bhopal, 1994), and includes ethnic groups from across the globe. Furthermore, any research examining health differences between ethnic groups needs to take into account the fluid and dynamic nature of (self-ascribed) ethnic identities, and the socio-economic context in which ethnicity is shaped (Senior & Bhopal, 1994). Failure to do so can unintentionally contribute to racism (Karlsen & Nazroo, 2002; Scambler, 2010). I will return to this elaborately in Chapter 5.

1.3.2 A brief history of 20th century migration in the countries involved in the EU-GEI study

The definition of minority-status differs per EU-GEI country (see below), but Figure 1.1 gives an overview of the percentage of the population that belongs to an ethnic minority (England and Wales, the Netherlands) or that is born abroad (Spain, France, Italy), depending on how minority status is officially recorded. Brazil is not included in this graph, because they solely make distinctions between ethnic groups

on skin colour, and it is hard to compare this directly to the remaining countries. As can be seen in Figure 1.1 below, in the Netherlands the highest percentage of the population belongs to an ethnic minority background, and the percentage is lowest in Italy.

Figure 1.1: Percentage of the population belonging to an ethnic minority, or born abroad



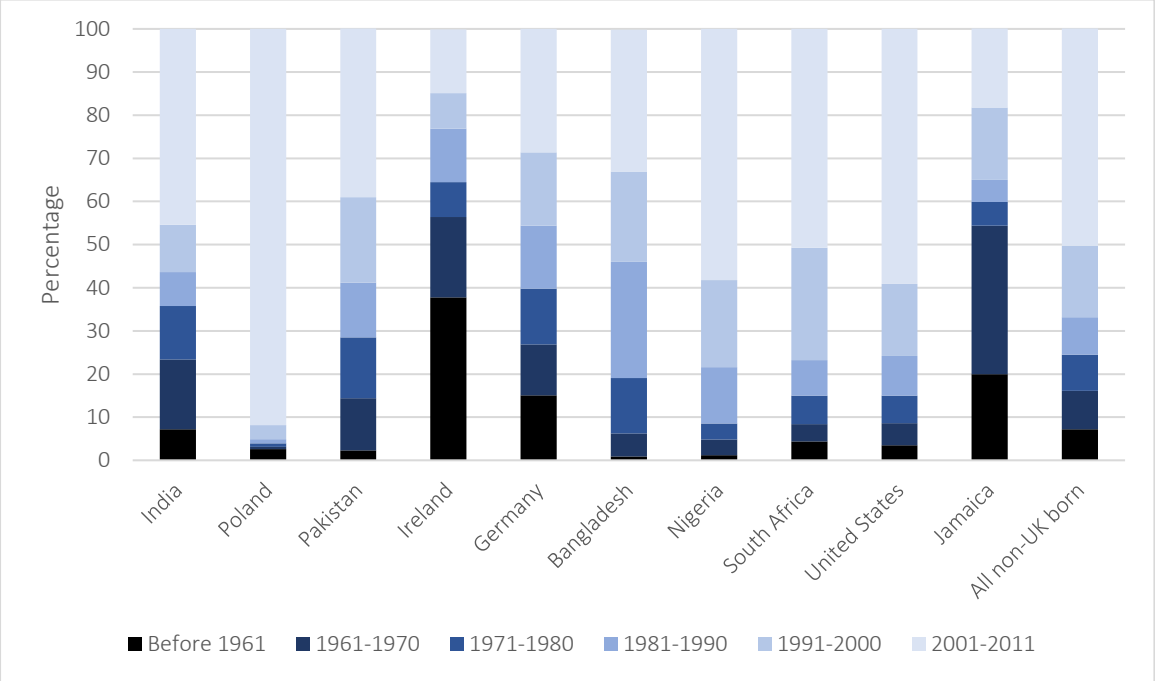
1.3.2.1 England

The United Kingdom has a long-standing history of immigration, with numbers of migrants increasing after the Second World War, and particularly since the 1990s.The 1948 British Nationality Act granted subjects of the British Empire the right to live and work in the UK (although they were subject to tighter immigration rules again from 1961 onwards)(Migration Watch UK, 2014) leading to significant immigration from former British colonies and Commonwealth countries throughout the second half of the twentieth century (Migration Watch UK, 2014). Following the Treaty of Maastricht in 1992, freedom of movement of persons was introduced in the European Union, allowing citizens to work and live in any member state (European Parliament, 2017). Whilst net migration has been positive since the 1951 Census, it has been rising consistently sharper since the mid-1990s, largely driven by non-EU migrants (ONS Digital, 2015).

The percentage of the resident population of England and Wales born abroad is recorded as 13 percent (7.5 million individuals) in the 2011 Census, up from 4% (1.9 million individuals) in the 1951 Census (Office for National Statistics, 2013a). In the 2011 Census, 4.8 million people (9%) are recorded as having a non-UK passport, of which 2.3 million (4%) hold a passport from a European Union country (Office for National Statistics, 2012). As can be seen in Figure 1.2 below, there is a distinct difference between the time of arrival of people born in the ten countries with the highest proportion of nationals in the UK. Almost 40% of Irish immigrants arrived before 1961, and whilst large proportions of the Commonwealth citizens (India, Pakistan, Bangladesh, Jamaica) arrived before the 1990s, recent years have seen a surge in Nigerian

immigrants, as well as Polish immigrants following the Eastern expansion of the European Union. Around half of non-UK born population has lived in the UK for more than ten years, with a quarter each arriving between five and ten years ago and within the last five year (Office for National Statistics, 2013a).

Figure 1.2: Top ten non-UK countries of birth of usual residents in England and Wales (2011) by year of arrival. Source: Office for National Statistics (2013).



1.3.2.2 The Netherlands

The Netherlands, too, has a long-standing history of migration, which also significantly increased after the end of the Second World War. The post-war independence of Suriname and Indonesia caused significant immigration from these two former colonies, and significant migration has also accrued from the Dutch Antilles and Aruba, which are still part of the Kingdom of the Netherlands. From the 1960s onwards, substantial numbers of 'guest workers' were recruited to carry out low-paying jobs in the booming economy, initially from Spain and Italy, but after the economy recovered from the 1973 oil crisis, mainly from Morocco and Turkey (Jennissen, 2011; Nicolaas & Sprangers, 2007). It was thought that their presence would be temporary, but this has, in fact, resulted in a large and permanent Moroccan and Turkish minority in the Netherlands (Focus Migration, 2007d; Jennissen, 2011; Nicolaas & Sprangers, 2007).

Presently, over 20% of the roughly 17 million inhabitants of the Netherlands has a migration background, meaning that they or at least one of their parents were born abroad (Centraal Bureau voor de Statistiek, 2016). Twelve percent of inhabitants of the Netherlands has a non-Western migration background (around 2.1 million individuals) and almost ten percent a Western migration background (around 1.7 million). Almost half of these people have been born in the Netherlands and thus are second-generation migrants (in the official Dutch definition, though many might have Dutch passports). The largest population groups

are of Turkish (398,000; 2.3%), Moroccan (385,000; 2.3%) Indonesian (367,000; 2.2%), German (360,000; 2.1%) and Surinamese (349,000; 2.1%) origin. The Eastern expansion of the European Union has led to an increase in Polish, Romanian and Bulgarian first generation immigrants, and since 2014 the number of asylum seekers, particularly from Syria and Eritrea, has increased substantially (Centraal Bureau voor de Statistiek, 2016).

1.3.2.3 France

France, too, has a colonial migration history, with immigrants arriving as early as the 18th and 19th century, and migration substantially increasing since the end of the Second World War. In 1946, immigrants made up 5% of the population, rising to 8% in 2005, mainly from former colonies in the Maghreb region (Morocco, Algeria and Tunisia) (Focus Migration, 2007b).

France classes everyone born on mainland France as well as in their overseas territories (Guadeloupe, French Guiana, Martinique, Réunion, Mayotte, French Polynesia, Saint Pierre and Miquelon, Wallis and Futuna, Saint Martin, Saint Bartélemy, New Caledonia, French Southern Antarctica) as 'born in France', and makes no distinction between these categories. As of the first of January 2014, 7.6 million inhabitants (11.6% of the population) were born outside of France, although only 5.8 million of these were classed as immigrants (8.9% of the population) (Institut national de la statistique et des études économiques, 2015; Institut National de la statistique et des études économiques, 2014). Of those born in France, 600,000 (1.0%) had a foreign nationality, these were mainly young children born to immigrant parents. Of the total number of immigrants, 1.73 million (22.7%) originated from the Maghreb region, 811,000 (10.7%) from other African countries and 1.85 million (24.3%) from countries within the European Union. The EU country with the largest immigrant population in France was Portugal (Institut national de la statistuque et des études économiques, 2016).

1.3.2.4 Spain

Historically, Spain's migration history has mainly been one of emigration. In the first half of the twentieth century, there was substantial emigration to South America, and later to other western European countries. The latter came to an abrupt halt with the 1973 oil crisis (Bover & Velilla, 1997; Focus Migration, 2007c). Until the late 1980s the number of foreign immigrants in Spain was less than 0.02% of the overall population, and even in 1995 it was only 0.05%. Most immigrants have traditionally come from Europe and South America, although from the end of the 1980s onwards immigration from Africa has seen the sharpest increase, from around 15% of total immigrants in 1983 to 25% in 1995 (Bover & Velilla, 1997).

The most recent year for which data is available (2016), is the first year since the onset of the financial crisis where net immigration occurred (Instituto Nacional de Estadística, 2017), and previous net emigration is largely explained by Spanish nationals emigrating due to the financial crisis and associated high levels of unemployment (Instituto Nacional de Estadística, 2015). Currently, Spain has a total foreign-

born population of around 4.4 million (9.5% of the total population) reflecting significant recent migration. The largest foreign groups are Romanian (678,000; 1.5%), Moroccan (667,000; 1.4%), British (294,000; 0.6%), Italian (203,000; 0.4%) and Chinese nationals (178,000; 0.4%), with further groups mainly coming from other EU-countries and Latin- and South America (Instituto Nacional de Estadística, 2017).

1.3.2.5 Italy

Italy, too, was historically an emigration country. Until around 1885 this was predominantly to other European countries, but from then on mainly to the United States, Argentina and Brazil (Bonifazi, Heins, Strozza, & Vitiello, 2009). It was not until the second half of the 1970s that Italy started receiving immigrants from low- and middle-income countries and later from other European countries (mainly from Central and Eastern Europe). In the 1951 Census, 47,177 foreign residents were recorded, rising to 356,000 in 1991 and 1.3 million in 2001 (Bonifazi et al., 2009).

Currently, the total number of inhabitants with foreign citizenship is almost 5 million or 8.3% of the total population (Istat, 2017). The number of non-EU foreign nationals holding a residence permit on 1 January 2016 was almost 4 million (6.5%), with the largest number coming from Morocco (510,000; 0.8%), Albania (483,000; 0.8%) and China (334,000; 0.6%) (Istat, 2016b). There was a significant change in reason for issuing permit, shifting from being issued for those who come to Italy for work to those who come to Italy for asylum and humanitarian reasons following recent ongoing international conflicts (Istat, 2016b). Immigration appears to have slowed down recently; whilst between 2011 and 2014 it grew substantially, little change is reported between 2014 and 2015 when around 280,000 individuals immigrated to Italy (Istat, 2016a).

1.3.2.6 Brazil

Brazil, a former Portuguese colony, has a very different migration history to the other countries in the EU-GEI study. Colonialisation not only meant the immigration of a substantial number of European immigrants, but also large numbers of African slaves (Amaral & Fusco, 2005; Focus Migration, 2007a). Following independence of Portugal in 1822 (Central Intelligence Agency, 2017), there were three waves of immigration in the late 19th and early 20th century bringing a large number of European (mainly Portuguese) and, to a lesser extent, Japanese immigrants. In the latter half of the 20th century, immigration became more local (from other South American countries) and many Brazilians started to emigrate in search of economic opportunities (Amaral & Fusco, 2005; Focus Migration, 2007a).

The total Brazilian population in the 2010 Census is around 191 million inhabitants, and, reflecting its centuries long immigration history, this is a very ethnically mixed population. Brazil does not classify people by country of birth or ethnicity, but by self-identification of skin colour according to the following categories: White, Black, Brown, Yellow and Indigenous. Forty-seven percent of the population is White (91 million), and the next biggest population group is the 'Brown' population group (82 million; 43%),

followed by the Black group (15 million; 8%). Other population groups ('Yellow' and Indigenous) are much smaller (Instituto Brasileiro de Geographica e Estatistica, 2011). No Census data on immigration was available, although UNICEF indicated that in 2013, 306,000 individuals (0.2%) migrated to Brazil, still predominantly from Portugal. Emigration in 2013 stood at 1.1 million individuals (0.6%), predominantly moving to the United States and Japan (UNICEF, n.d.).

1.3.3 Migration and health

Migration has three broad stages during which exposure to personal or relational stressors can adversely influence migrants' (mental) health: a pre-migration state where the decision to migrate is made and the migration is prepared; the act of migrating itself; and a post-migration state where the individual settles into a new cultural and social framework (Bhugra et al., 2011; Jones & Bhugra, 2010). Whereas first generation migrants (those who move from one country, region or place of residence to settle in another (Bhugra et al., 2011)) experience all three stages, their direct and subsequent descendants only share a similar post-migration experience with their parents (Jones & Bhugra, 2010). The effects of exposure to migration-related stressors might not always be immediate. Ødegaard's seminal study, for instance, found that schizophrenia peaked ten to twelve years after migration (Ødegaard, 1932). There is some evidence that, at least initially, migrants have better health than the non-migrant population in their host country (Rechel, Mladovsky, Ingleby, Mackenbach, & Mckee, 2013). This 'healthy migrant effect' is considered partially due to the act of migrating requiring a degree of good health (Rechel et al., 2013). Nonetheless, data collection on the health of migrants is patchy at best and consequently knowledge about migrant health, at least in Europe, is limited (Rechel et al., 2013).

It appears that any healthy migrant effect for first generation migrants is of relatively short duration: a systematic review found that migrants experienced poorer self-perceived health than the general population, after adjusting for age and sex (Smith Nielsen & Krasnik, 2010). Some studies included in this review found that this was partially accounted for by post-migratory socio-economic status and discrimination (Cooper, 2002; Lindström, Sundquist, & Östergren, 2001; Wiking, Johansson, Sundquist, & Wiking, 2004), although this might in itself be a result of migrant status or ethnic origin (Davies, Basten, & Frattini, 2006; Ingleby, 2006) and migration itself might be a social determinant of health (Marmot, Allen, Bell, Bloomer, & Goldblatt, 2012). Migrants also appear to be at increased risk of diabetes (Vandenheede et al., 2012), some communicable diseases (EuroHIV, 2007; European Centre for Disease Surveillance and Control, 2010), maternal and child health problems (Gissler et al., 2009), and occupational health hazards and injuries compared with the native population in the same occupations (Schenker, 2008; 2010).

This pattern of poorer health appears to extend to poorer mental health too, although evidence here is more mixed (Bhugra et al., 2011; Rechel et al., 2013). A recent Canadian review of the literature found that prevalence of common mental disorders in migrants was initially lower than the general population, but increased to similar levels over time (Kirmayer et al., 2011). In Europe, a proportion of immigrants and

ethnic minorities experience more depressive symptoms (Missinne & Bracke, 2012), with refugees at double the risk of other migrants, possibly due to their excessive exposure to pre-migratory risk factors for poor mental health and a more stressful migration experience (Lindert, Von Ehrenstein, Priebe, Mielck, & Brähler, 2009). Certain migrant groups (South Asian women in the UK, Northern European migrants to Australia) also appear to be at excess risk of suicide (McKenzie, Bhui, Nanchahal, & Blizard, 2008; Morrell, Taylor, Slaytor, & Ford, 1999).

1.4 Migration, ethnicity and psychotic disorders

The most consistent epidemiological finding pertaining to mental ill health and migration is the excess risk of psychotic disorders in migrants and their descendants. This finding is long-established (Ødegaard, 1932), and well-replicated (Bourque et al., 2011; Cantor-Graae & Selten, 2005). I will discuss the evidence regarding this finding and its determinants in this Section.

There are no reasons to assume this increased incidence is an artefact of demographics (there might simply be more young men in minority groups, who have the highest risk) as incidence is still higher after adjusting for age and sex (see, for instance, Fearon et al., 2006). Furthermore, whilst it appears to be true that there are, for instance, ethnic differences in pathways to care (Morgan, Mallett, Hutchinson, & Leff, 2004) and in the probability of being physically restrained whilst accessing mental health services (NHS Digital, 2017) there is no evidence that there are systemic racist diagnostic biases (at least, not concerning Black Caribbeans in the UK (Hickling, McKenzie, Mullen, & Murray, 1999; Lewis, Croft-Jeffreys, & David, 1990)). I therefore assume that the differential incidence rates across different ethnic groups are a true reflection of differential risk.

Theoretically, there are several reasons why rates of disorder might be increased in ethnic minority groups, including (but not limited to) ethnicity itself, pre-migration circumstances, migration itself and the post-migratory social circumstances of minorities. I discuss these potential causes in the remaining Section.

1.4.1 Ethnicity

If ethnicity, *per se,* is a reliable risk indicator of increased psychosis, incidence of psychosis in the general population in countries of origin such as Caribbean, North African, or Sub-Saharan African countries would be expected to be approximately as high as the incidence in the population groups from those countries in their host countries in Western Europe. The falsity of this assertion is already demonstrated in Ødegaard's seminal study of Norwegian migrants in Minnesota: rates were higher in Norwegian emigrants than they were in the general population in Norway (Ødegaard, 1932). Rates in several relevant countries of origin have been investigated, predominantly in Caribbean countries. In Trinidad (Bhugra et al., 1996), Surinam (Hanoeman, Selten, & Kahn, 2002; Selten et al., 2005) and Jamaica (Hickling, 1995) rates were between 11 and 22 per 100,000 person years, much lower than rates of psychosis in Black Caribbean populations in

Western Europe, and, in fact, much more in line with general population rates in host countries (Kirkbride et al., 2012; Veling et al., 2006). Two of the three studies were methodologically strong, either using the same case-finding methodology as the WHO ten-country study (Bhugra et al., 1996; Jablensky et al., 1992) or relying on well-established patient registers (Hanoeman et al., 2002), while the third relied on reporting through existing service infrastructure (Hickling, 1995). There is little evidence from other countries of origin unfortunately. The search strategy employed to expand the systematic review (Chapter 2) globally identified only one further study in South-Africa (Burns & Esterhuizen, 2008), and an ongoing multinational research project (Morgan et al., 2017).

An initial explanation offered for this finding was the so-called 'unhealthy migrant' effect whereby those already more vulnerable to developing psychosis were more likely to migrate (Ødegaard, 1932). This is however an implausible mechanism: migrants' health at least initially appeared to be at least as good as the general population health in host countries (see above). There was also no empirical evidence to underpin this putative cause (Selten, Cantor-Graae, Slaets, & Kahn, 2002; van der Ven et al., 2014). Selten and colleagues established that even if the entire population of Surinam would migrate to the Netherlands (trebling the denominator) and would not contribute any extra cases, the Surinamese population in the Netherlands still faced an increased incidence of schizophrenia (Selten et al., 2002). Using Swedish conscript and population registry data, Van der Ven and colleagues assessed whether known risk factors for psychosis were more common in conscripts who later migrated than in those who remained in Sweden. They showed that the only risk factor more prevalent in later emigrants was urban upbringing, and concluded their findings provided evidence against the selective migration hypothesis (van der Ven et al., 2014).

A second reason why ethnicity shouldn't be considered a reliable risk indicator is that it appears that various broad ethnic groups are differentially affected across various host countries. An example is people of Black Caribbean descent. In England, their risk is higher than any other ethnic group, with a pooled risk of 5.6 times higher than the general population (Kirkbride et al., 2012). In the Netherlands, people of Black Caribbean origin (from Surinam and the Dutch Antilles) are at increased risk too, but their estimated risk is limited to 2-3 times that of the general population (Veling et al., 2006). In the Netherlands, Moroccan immigrants face the highest increase in risk with incidence rate ratios (IRRs) between 4 and 6 (Veling et al., 2006), whereas in France North African migrants face no significantly increased risk (IRR: 1.4; 95% confidence interval [CI]: 0.4-5.6) (Tortelli et al., 2014), although the wide confidence interval suggests this study might be underpowered. Sub-Saharan migrants appear most at risk in France (IRR: 7.1; 95%CI: 2.3-21.8) (Tortelli et al., 2014).

1.4.2 Exposure to pre-migratory risk factors and the act of migrating

Various studies have sought to investigate the relationship between pre-migratory circumstances, migration itself and excess psychosis risk in migrants. These studies have examined various aspects of this

relationship, such as age at migration (Kirkbride, Hameed, Ioannidis, et al., 2017; Pedersen & Cantor-Graae, 2012; Veling, Hoek, Selten, & Susser, 2011) and refugee-status (Anderson, Cheng, Susser, McKenzie, & Kurdyak, 2015; Hollander et al., 2016). Refugees are at increased risk of psychosis compared with the general migrant population (Anderson et al., 2015; Hollander et al., 2016), and this is thought to be due to a combination of increased exposure to pre-migratory stressors and a more traumatic migration experience (and post-migratory factors likely also play a role) (Rechel et al., 2013).

It is not clear which age at migration conveys the greatest risk. A Dutch study showed that an earlier age of migration (0-4 years) is associated with the highest risk (Veling et al., 2011), whereas a Danish study demonstrated that in immigrants from the developed world, those who migrated aged 10-14 appeared to be at a similarly high risk (Pedersen & Cantor-Graae, 2012), and findings from East Anglia indicate that those who migrated during childhood (5-12 years) carried the highest risk (Kirkbride, Hameed, Ioannidis, et al., 2017).

While there is some evidence that pre-migratory exposure to stressors and migration contribute to excess risk in psychotic disorders in migrants, it is also clear that, despite being born in the host country and thus never having experienced either of these, second- and even third-generation migrants are still at increased risk of developing psychotic disorders or having psychotic experiences (Bourque et al., 2011; Oh, Abe, Negi, & Devylder, 2015). Moreover, the association between generational status and increased psychosis risk appears to be complex: some migrant groups in some host countries experience a higher risk in the second compared to the first generation, whereas in other groups and in other host countries the risk appears to be attenuated (Bourque et al., 2011). A further study suggests that given the same age structure, the risk across generations is approximately equal within ethnic groups (Coid et al., 2008). Pre-migratory factors and migration are unable to explain these findings.

1.4.3 Post-migratory social circumstances

The authors of the most recent meta-analysis into the higher risk of psychosis in ethnic minorities argue that it is therefore likely that post-migratory factors play an important role in this excess risk (Bourque et al., 2011) and there is some evidence to support this. For instance, in the Netherlands, there are some indications that, at a population level, there is a link between increased incidence of psychosis and high self-perceived discrimination (Veling et al., 2007), although this wasn't replicated at an individual-level (Veling, Hoek, & Mackenbach, 2008). Most of the epidemiological investigations into the role of social circumstances as a risk factor for psychosis tend to look at their role in explaining variation in incidence in the general population at a neighbourhood level (Drukker, Krabbendam, Driessen, & van Os, 2006; Kirkbride et al., 2014; Veling et al., 2015), and not specifically at how they explain excess risk in (ethnic) minorities. An exception to this neighbourhood-level focus is a case-control study (Morgan et al., 2008), where the authors showed that markers of social disadvantage were more prevalent in the Black Caribbean population, and suggested this contributed to their excess risk.

1.5 Where this thesis fits in

The latest international systematic review and meta-analysis of the incidence of schizophrenia dates back to 2004 (McGrath et al., 2004), and I will synthesise literature pertaining to all psychotic disorders for the six countries included in the EU-GEI study published since then (Chapter 2). The latest international incidence study dates back even further: it was published in 1992 and with data collected in the 1980s (Jablensky et al., 1992). Our understanding of psychosis has changed substantially since then, with an increasing evidence base supporting a role of the (social) environment in explaining heterogeneity between settings. I will therefore use the EU-GEI study to update the incidence epidemiology of psychosis (Chapter 4).

The higher rates of psychosis in ethnic minorities are poorly understood, and despite the alleged importance of the post-migratory social environment in explaining excess psychosis risk, there is a dearth of epidemiological studies investigating this, and a limited understanding of why and how the social environment contributes to this. In this thesis, I aim to contribute to filling this important gap in the literature by taking a step back from epidemiology and examining literature from the social sciences to inform my epidemiological hypotheses (Chapter 5), which I will subsequently test using the case-control data of the EU-GEI study (Chapters 6 and 7).

1.6 Intellectual position taken in this thesis

I approach this thesis from the viewpoint of social psychiatry, and as such it sits within a wider literature considering psychotic disorders to be predominantly disorders of social functioning, and where mental illness more broadly is considered in its social context (Ventriglio, Gupta, & Bhugra, 2016). This has been conceptualised in many ways. In neuroscience for example, biological constructivism examines how the brain and culture mutually form each other (Malinowska, 2016), and predictive processing is invoked as a model to explain how the brain interacts with the (social) world (Clark, 2016). Essentially, the brain can be seen as a Bayesian prediction engine which always tries to be one step ahead of sensual perception, and uses any discrepancy between this anticipation and perception to update its belief systems (Clark, 2016; Fletcher & Frith, 2009). Within this framework, psychotic disorders (in particular the occurrence of hallucinations and delusions) are conceptualised as errors in this processing of the social world (Clark, 2016; Fletcher & Frith, 2009; Gold & Gold, 2014), and psychotic disorders are disorders of social functioning (Abed & Abbas, 2011; Gold & Gold, 2014; van Os, Kenis, & Rutten, 2010).

Some of the epidemiological risk factors described in this Chapter, could also be interpreted in this light. The experience of childhood trauma could bias subsequent 'cognitive schemas': the lens through which individuals interpret the world; this could be further biased by later social disadvantage (Howes & Murray, 2014). Cannabis use and social isolation as risk factors could be seen as a deficit in appropriate interaction with the social environment. The excess risk in ethnic minorities has also been explained in this way: Abed

and Abbas argue that psychosis is an exaggerated threat response to a lack of contact with in-group members, excessive contact with out-group members or a combination of both (Abed & Abbas, 2011). This is also their explanation of why visible minorities are at excess risk: inter-group differences are more salient (Abed & Abbas, 2011). Gold and Gold draw a similar conclusion as they argue that hallucinations and delusions result from a functionally disconnected stress-response system, which responds particularly strongly to the social threat of malicious intentions and this is more salient to visible minorities, and simply more numerous in more densely populated urban areas (Gold & Gold, 2014).

This idea that the brain and mind, and their associated disorders, should be understood in reference to the social world also lends itself well to a more sociological framework which, as will be discussed in Chapter 5, seeks to identify structural social drivers that influence mental health more broadly and psychotic disorders, specifically (Horwitz, 2017). Most structural drivers of disorder (such as lack of fulfilling job opportunities, living in sub-par housing and lack of access to good-quality education) are not randomly distributed in the population, but tend to occur in clusters (Thoits, 2017). Minorities tend to be hit harder by these structural drivers as they occur with greater frequency in minority groups (Morgan et al., 2008; Wilson, 2010). On top of this, minorities also face unique challenges in the form of institutional and overt discrimination, which act not only as additional stressors but might also amplify the effects of other structural drivers such as those just mentioned (even if these are not explicitly targeting minority groups)(Wilson, 2010). Perceived discrimination and hostility also shape the way in which individuals from a minority background interpret themselves within their social environment: the world is more threatening if you constantly see yourself as different (Abed & Abbas, 2011; Gold & Gold, 2014).

It is from this intellectual position that findings from the empirical Chapters will be interpreted: variation in incidence by putative environmental risk factors reflect differences in the social reality brought about by these risk factors, and insofar as exposures tested in the case-control study explain variance in psychosis risk, this is because they represent differences in social reality of participants.

1.7 Structure of this thesis

The second Chapter of this thesis seeks to systematically review and meta-analyse the literature on the incidence of non-organic adult psychotic disorders in England, the Netherlands, Spain, France, Italy and Brazil in order to understand the epidemiological landscape, and serves as a background to the EU-GEI incidence study (Chapter 4). In Chapter 3, I describe the methodology of the EU-GEI study, which includes settings from the aforementioned six countries and serves as the data source for the remainder of this thesis. Chapter 4 contains details on the incidence arm of the EU-GEI study, where I estimate the crude and directly standardised incidence of all psychotic disorders, non-affective and affective disorders in the 17 catchment areas included in this study. I will also examine variation by putative environmental risk factors.

Chapter 5 contains a philosophical and sociological exploration of the putative causes of higher rates of psychotic disorders in ethnic minorities, proposing that if we are to fully understand this well-replicated epidemiological finding we ought to examine the social context of ethnic minority status. I specifically propose that minorities' sociocultural distance from the majority will increase their psychosocial disempowerment and subsequent risk of disorder. Chapter 6 seeks to test empirically the ideas put forward in Chapter 5, using case-control data from the EU-GEI study. Here, I follow the methodology originally proposed for the AESOP study, of estimating crude odds ratios (ORs) and progressively adding exposures in an attempt to 'explain away' crude excess risk. Chapter 7 explores whether ancestry, *per se*, as measured by genetic distance from the White British population, is a predictor of psychosis risk in the EU-GEI subsample recruited in England when social factors are also taken into account. Chapter 8 discusses and contextualises the main findings of this thesis and discusses what this thesis means for the aetiology of psychotic disorders and for public health and policy. I also discuss recommendations for future research.

Chapter 2 - A systematic review and meta-analysis of the incidence of psychotic disorders in England, the Netherlands, Spain, France, Italy and Brazil.

2.1 Background

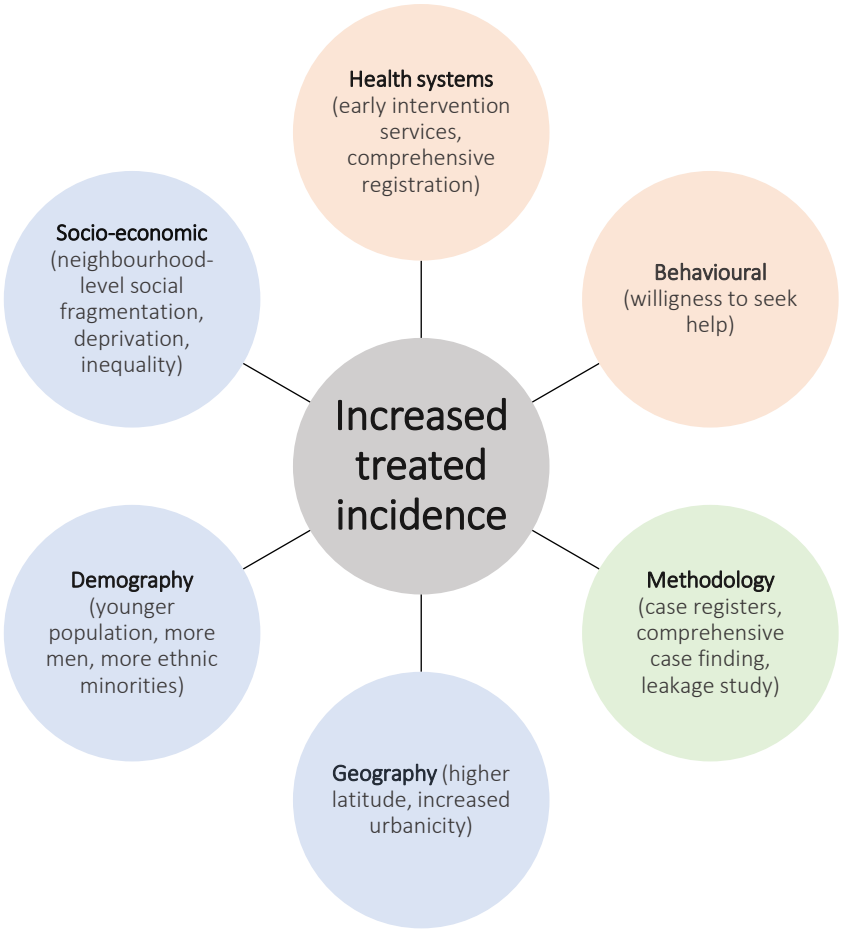
In recent years interest in examining the incidence of psychotic disorders has increased substantially, and research efforts have broadened beyond North America (Anderson et al., 2015), the UK (Kirkbride et al., 2012), the Netherlands (Veling et al., 2006) and Scandinavia (Leão et al., 2006; Sundquist, Frank, & Sundquist, 2004), where epidemiological research was traditionally concentrated (Alonso et al., 2007; Gigantesco, Lega, & Picardi, 2012; Lasalvia et al., 2014; Menezes & Scazufca, 2007; Morgan et al., 2017; Tarricone et al., 2012). This means that an increasingly large and varied body of evidence is available to assess and synthesise incidence rates and to investigate any potential heterogeneity.

The most recent international meta-analysis on this topic dates from 2004, and only includes studies published up until 2002 (McGrath et al., 2004). More recent meta-analyses have been carried out, but these were limited to one country (Kirkbride et al., 2012), a specific population group (Bourque et al., 2011), a particular risk factor (Esterberg, Trotman, Holtzman, Compton, & Walker, 2010; Matheson et al., 2012; Moore et al., 2007; Varese et al., 2012; Vassos et al., 2012) or involved a re-analysis of individual incident cases (van der Werf et al., 2014).

These meta-analyses suggested that incidence varies between settings (Kirkbride et al., 2012; McGrath et al., 2004), and previous studies suggested an aetiological role for environmental risk factors. Meta-analyses have reported a higher incidence in settings at a higher latitude (Kinney et al., 2009; Saha et al., 2006), despite lack of agreement on the underlying mechanism. Risk of schizophrenia is also known to be higher in more urban or densely populated areas (Vassos et al., 2012), and in more socially fragmented (Faris & Dunham, 1939; Giggs & Cooper, 1987; Hare, 1956; Hollingshead & Redlich, 1958; Maylaih, Weyerer, & Hafner, 1989) or deprived areas (Kirkbride et al., 2014; Lasalvia et al., 2014), as well as areas with higher levels of economic inequality (Burns & Esterhuizen, 2008; Kirkbride et al., 2014). No evidence was found for variation in incidence rates over time (Kirkbride et al., 2012), by study quality (Kirkbride et al., 2012; McGrath et al., 2004) or other methodological features (McGrath et al., 2004). Nonetheless, a study directly comparing a first contact study with a case register found a three-fold increase in incidence in the latter (Hogerzeil, van Hemert, Rosendaal, Susser, & Hoek, 2014), and studies carrying out a leakage study have demonstrated this methodological feature increases the number of cases identified by approximately ten percent (Cooper et al., 1987; Kirkbride et al., 2006). These hypothesised influences on the incidence of psychosis are summarised in Figure 2.1 below. This Figure contains two further categories of influences on treated incidence: socio-economic, geographical and demographic factors linked to true

incidence, and health services and behavioural factors influencing the interaction between true and treated incidence. The latter reflects the epidemiological notion of a 'disease pyramid', where not all individuals suffering from a disorder will access the formal health system (Goldberg & Huxley, 1980). These health systems factors are not explicitly tested but will be discussed briefly in Section 2.5.

Figure 2.1: Factors influencing treated incidence of psychosis.



This Figure captures influences on true incidence (in blue), influences on the interaction between true and treated incidence (in orange) and research design biases that could influence reported incidence (in green).

To synthesise the accumulating international research on the incidence of psychotic disorders and its putative socio-economic correlates since the most recent international review of psychosis incidence, I systematically identified and synthesised studies relating to the incidence of psychotic disorders conducted across the six countries included in the EU-GEI study: England, France, Italy, The Netherlands, Spain and Brazil. I also assessed if any observed variance could be explained by study quality, time, urbanicity, latitude, economic inequality and self-perceived freedom and trust as indicators of social fragmentation. This systematic review, meta-analysis and meta-regression also serves as a literature review to aid in the interpretation and contextualisation of the EU-GEI incidence study, which is presented in Chapter 4.

2.2 Hypotheses

In this Chapter, I tested the following hypotheses:

1. Incidence of psychotic disorders would vary substantially between studies;
2. Incidence would be higher in more urban areas, settings at a higher latitude, more economically unequal countries, and in countries were self-reported levels of freedom and trust are lower.
3. There would be no association between study quality and incidence of psychotic disorders.

2.3 Methods

The research question, search strategy, data extraction procedure, inclusion criteria and statistical methodology were specified *a priori* in line with reporting guidelines for systematic reviews and meta-analyses (National Collaborating Centre for Methods and Tools, 2011). I wrote a protocol and submitted this to PROSPERO, an international prospective register of systematic reviews hosted by the Centre for Reviews and Dissemination at the University of York. It was submitted on 13 April 2015 under registration number CRD42015019276. The original protocol is included as Appendix 2A.

There were three post-submission changes to the original protocol:

- The inclusion of national socio-economic correlates as predictor variables for meta-regression;
- The inclusion of Embase and Web of Science as databases to be searched instead of ScIELO;
- The removal of the language of publication inclusion criterion in favour of including studies that were published in languages other than English.

In this Chapter, a study refers to a body of research from which data has originated. A citation refers to a unique report, published or otherwise, of this data. As such, a single study could be covered by more than one citation (Kirkbride et al., 2012).

A recent systematic review already identified all relevant studies in England until 2009 (Kirkbride et al., 2012), and I updated this search strategy to include studies up until 31 December 2014. The data from the English review were made available for synthesis within the present review. The identification of studies in the present systematic review therefore only included English studies identified published between 1 January 2010 and 31 December 2014. In order to aid comparability and reliability, I followed the English methodology as closely as possible (Kirkbride et al., 2012), and amalgamated results from the two searches at the analysis stage. The chronological starting point of the systematic review and meta-analysis (1 January 2002) followed the end-point of study identification in the previous international systematic review and meta-analysis (McGrath et al., 2004).

2.3.1 Inclusion criteria

Studies were eligible for inclusion if they met the following criteria:

a. Published between 1 January 2002 and 31 December 2014;
b. Wholly or partially conducted in Brazil, England, France, Italy, the Netherlands or Spain;
c. Published in scientific or grey literature;
d. Containing original data on the incidence of non-organic adult psychosis (16-64 years);
e. Published in the English, Dutch, French, Italian, Spanish or Portuguese language.

2.3.2 Information sources

I (HEJ) conducted a systematic literature search in the PubMed, PsycINFO, Embase and Web of Science databases. I manually checked reference lists of all citations included in the systematic review for further studies. Where possible, I contacted the lead author of each citation with the list of studies included in the review and asked to identify any further studies, including those appearing in grey or unpublished literature. I contacted a colleague in Chile (Antonia Errazuriz) who conducted a systematic review in Latin and South American countries as part of her PhD thesis to ensure no Brazilian studies were overlooked (Errazuriz, 2013).

2.3.3 Search strategy

The search strategy was based on a search strategy previously used in England (Kirkbride et al., 2012), and adhered to Cochrane Systematic Reviewing guidelines (The Cochrane Collaboration, 2008). This strategy was known to be effective, and has been chosen to increase reliability in the systematic identification of studies between countries. The search strategy can be found below (Table 2.1).

2.3.4 Selection process

The PubMed search was carried out first, and HEJ, PBJ and JBK each screened an equal share of the titles. For the other databases, HEJ assessed all titles. The title screening had the objective of excluding true negatives; I excluded studies which were evidently irrelevant to the purpose of the review whereas I kept studies of potential or definite relevance for abstract review.

I reviewed all abstracts and scored them against inclusion criteria. When studies met inclusion criteria, or I couldn't reach a conclusion, I assessed the full text article. I resolved uncertainties in agreement with JBK, based on the inclusion criteria.

2.3.5 Data collection process and data management

I managed citations in Mendeley (version 1.17.10), and extracted data from eligible citations onto a standardised Excel spreadsheet.

Where I identified missing data items, I made three attempts to contact the corresponding author: an initial email to the address listed on the paper, a reminder after two weeks and a final reminder after one month (where needed).

Table 2.1 Search strategy

#1	schizo* .ti,ab
#2	(psychotic or psychosis or psychoses).ti,ab
#3	(bipolar disorder*).ti,ab
#4	(delusion* disorder).ti,ab
#5	(((severe or serious or chronic) and mental and (illness* or disorder*)).ti,ab.
#6	SMI.ti,ab
#7	(mani* depressi*).ti,ab
#8	chronic psychosis
#9	exp psychosis
#10	schizoaffective disorder
#11	(#1 or #1 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10)
#12	(inciden* or epidemiolog*).ti,ab
#13	((first* or 1st* or hospital*) and (episode* or contact* or admission* or admit*)).ti,ab
#14	(case and register*).ti,ab
#15	case control*.ti,ab
#16	prospectiv* or population* or communit* or survey*).ti,ab
#17	(#12 or #13 or #14 or #15 or #16 or #17)
#18	(Brasil or Brazil)
#19	France
#20	(Italy or Sicily)
#21	(Spain)
#22	(Holland or the Netherlands)
#23	(#19 or #20 or #21 or #22)
#24	(Brasilian or Brazilian)
#25	(French)
#26	(Italian or Sicilian)
#27	(Dutch)
#28	(Spanish)
#29	(#24 or #25 or #26 or #27 or #28)
#30	(Sao Paulo or Ribeirão Preto)
#31	(Creteil or Clermont-Ferrand)
#32	(Bologno or Palermo)
#33	(Amsterdam or Gouda)
#34	(Barcelona or Oviedo or Valencia)
#35	(#30 or #31 or #32 or #33 or #34)
#36	(#23 or #29 or #35)
#37	(#11 and #17 and #36)

2.3.6 Outcome measures and study details

The main outcome variable collected was the incidence rate, expressed per 100,000 person-years, of one or more of the following (categories of) psychotic disorders:

- all psychotic disorders;
- non-affective psychosis;
- schizophrenia;
- affective psychosis;
- bipolar disorder;
- psychotic depression;
- substance-induced disorders;

- symptoms[1], and
- other.

I recorded the exact diagnostic outcome, including classification system and relevant codes. Studies investigating the same diagnostic outcome, albeit using different diagnostic classification systems, were grouped together using the same pragmatic approach as Kirkbride et al (2012). Based on the latest international meta-analysis, I anticipated that diagnostic classification used made no difference to the recorded incidence (McGrath et al., 2004). Uncertainties on how to categorise citations were resolved in cooperation with PBJ.

I either directly ascertained incidence rates from the citation, or derived them manually by dividing the number of cases by the denominator person years. The former was used when both were available. If, in this Chapter only, I give a point estimate without an estimate of uncertainty (an incidence rate without a 95% confidence interval), the estimate of uncertainty could not be obtained based on the original publication.

I collected three types of data items: study-level variables, rate-level variables and meta-variables. The former provides information about the design of the study, rate-level variables include information on the estimate of incidence in each citation, and meta-level variables are included to explore heterogeneity in rates by various covariates including latitude, urbanicity and study quality. Meta-level variables might not be explicitly measured in the original study, but can be derived from it. A full list of variables included can be found in Table 2.2 below. I retrieved incidence rates for all available sociodemographic strata such as by age, sex, ethnicity, country of birth or levels of deprivation.

Table 2.2: Data items collected

Study-level	Rate-level	Meta-level
Authors	Incidence rates	Study quality
Study title	Size of the numerator	Urbanicity
Publication source	Size of denominator population	GINI-coefficient (inequality)
Year of publication	Method of standardisation	Self-reported levels of freedom (World Values Survey)
Study type (incidence / special population)		Self-reported levels of trust (World Values Survey
Geographical setting		
Study length (recruitment dates and duration)		
Age range		
Diagnostic outcomes studied		
Usage of OPCRIT diagnosis		
Method of confirming diagnosis		
Diagnostic classification system used		
Source of denominator data		

[1] Citations reporting symptoms were only included if other diagnostic outcomes were also reported, as this diagnostic outcome alone did not fulfil inclusion criteria.

2.3.7 Covariates

2.3.7.1 Study-level variables

Type of incidence study

I recorded five types of incidence study: first contact, first admission, first diagnosis, first GP record and a case register. *A priori*, I expected that type of incidence study would influence reported incidence rates. First admission studies were anticipated to record a lower incidence as they would only capture in-patient cases and exclude those who remained in the community unless, of course, such services were the only available treatment option. Based on a study showing a threefold increase in incidence rates in a case register compared with a first contact study (Hogerzeil et al., 2014), I also expected that case registers reported higher rates of disorder. I therefore systematically assessed the impact of study type on incidence rates.

Mid-year point of case ascertainment

This variable refers to the mid-year point of case ascertainment of each study. So, if a study recruited participants from 2004-2006, the mid-year point would be 2005. I included this to assess whether incidence of psychotic disorders varied over time. If a study reported data from more than one distinct time period, I used the mid-point for each time period and derived the arithmetic mean for regression purposes.

2.3.7.2 Meta-level variables

Study quality

I systematically assessed study quality and risk of bias, using seven criteria (Table 2.3), and include this in meta-regression. These quality criteria were originally developed for the English meta-analysis (Kirkbride et al., 2012), and were based on epidemiological theory of best practice.

Table 2.3: Quality criteria

Criterion	Reason for inclusion
Clearly defined catchment area	Assessing accuracy of incidence rates (numerator)
Accurate denominator data	Assessing accuracy of incidence rates (denominator)
Population-based case-finding	Minimising selection bias
Usage of standardised research diagnosis	Minimising diagnostic bias / assessing diagnostic reliability
Blinding of clinician to demographic variables	Minimising diagnostic bias / assessing diagnostic reliability
Inclusion criteria clearly listed	Measure of reporting quality
Leakage study conducted	Minimising ascertainment bias

Gini Index of economic inequality

The Gini Index is a measure of the income distribution of a country's residents. It is frequently used as a measure of inequality, and official data is kept by the World Bank (The World Bank, 2017). The index ranges from 0 (perfect equality) to 100 (perfect inequality). Looking at data collected since 2001, the

highest recorded score was 69.2 (Jamaica, 2001, and the lowest recorded score was 23.7 (Slovenia, 2008). I obtained data for countries in the meta-analysis for the most recent year available (2005 for France, 2010 for Italy, the Netherlands, Spain, and England, and 2012 for Brazil, see Section 2.4.2).

Self-reported levels of trust

The World Values Survey was started in 1981 and aims to investigate changing human values and their impact on social and political life. Surveys are carried out in four-year waves, the most recent completed wave lasted from 2010 – 2014 (World Values Survey Association, 2015). For each country, I used the most up-to-date data (2005-2007 for England, France and Italy, and 2010-2012 for Brazil, the Netherlands and Spain). The minimum sampling size per country was 1,200 participants, who must be representative of the general adult population. Exact sampling strategies varied per country, but were quality-checked centrally (World Values Survey Association, 2017). I could only obtain results for the UK as a whole, not for England specifically.

For the meta-regression I was interested in both self-reported levels of freedom and trust. Levels of trust were ascertained using the following question:

Generally speaking, would you say that most people can be trusted or that you need to be very careful in dealing with people?

With the following answering options:

1. *Most people can be trusted*
2. *Need to be very careful*

I downloaded raw data for each country from the World Values Survey website, and derived the percentage in each country answering that most people can be trusted (option 1. above).

Self-reported levels of freedom

I derived self-reported levels of freedom from the same World Values Survey questionnaires using the following question:

Some people feel they have completely free choice and control over their lives, while other people feel what they do has no real effect on what happens to them. Please use this scale where 1 means 'no choice at all' and 10 means 'a great deal of choice' to indicate how much freedom of choice and control you feel you have over the way your life turns out.

I used the arithmetic mean per country for analyses.

Urbanicity

I ranked settings from 1 to 21, with 1 being the most urban setting and 21 being the most rural setting. Cities were accounted for in terms of number of inhabitants, and the urbanicity of rural areas was approximated by looking at population density (number of inhabitants per square kilometre). Because exact details of catchment areas were often not available, I could not rely on official governmental statistics and had to rely on data available in the public domain (eg. Google). The index was created jointly by HEJ and JBK.

Latitude

The latitude of each setting was recorded as degrees away from the equator. When studies took part across different geographical settings, and incidence data was unavailable per setting, I averaged latitude for regression purposes. I used data available from the public domain (as above).

2.3.8 Prioritisation for analysis

Where more than one citation per study was identified, I applied the following criteria to prioritise citations for inclusion:

a) Data relevant to the specific outcome and/or exposure under investigation;
b) Data presented with a corresponding estimate of the standard error;
c) Data most closely relating to the entry criteria for the study (for example, when examining rates in ethnic minorities, the citation with the most details in reference to this specific analysis was retained);
d) Published data;
e) Citations published in the highest ranking journals.

2.3.9 Assessment of small study effects or publication bias

I assessed small study effects, the most common of which is publication bias, by visual interpretation of funnel plots or, where possible ($n \geqq 10$), formal Egger's test (Harbord, Harris, & Sterne, 2009; Steichen, 2009). In the absence of bias, dots on the funnel plot are approximately randomly distributed along the null hypothesis on the vertical axis (Sterne & Harbord, 2009).

2.3.10 Data synthesis and statistical methodology

After a brief description of the study yield, a narrative synthesis of results was the first stage of data analysis. This allowed for the identification of broad themes, and was done by visual interpretation of forest plots as well as careful examining of studies included in the systematic review. I report results of the narrative synthesis by country and, where possible, diagnostic categories within countries. Diagnostic categories for which only a small number of citations were found (less than four), were not included in any subsequent meta-analyses, but citations were retained for the narrative synthesis.

I then carried out meta-analyses for each diagnostic outcome. In order to do so, I assessed if all studies included a measure of uncertainty in the form of a standard error. This could either be reported, or derived using $1/\sqrt{d}$, where *d* is the number of cases. For the data to be synthesised and quantitatively analysed, I had to transform incidence rates to the natural logarithmic scale and derive corresponding standard errors. Studies for which no standard error could be obtained could not be included in the meta-analysis, but were retained in the systematic review. For synthesis of results across sub-groups of the population (sex and broad ethnic group), incidence rate ratios (IRRs) were computed and synthesised where possible. I divided the incidence rate in the exposed group by the rate in the unexposed group, and transformed this to the natural logarithmic scale. I derived standard errors using $\sqrt{(\frac{d1}{N1}+\frac{d0}{N0})}$ where *d1* was the number of cases in the exposed group, *N1* was the denominator in the exposed group and *d0* and *N0* were the number of cases and denominator in the unexposed group respectively (Kirkwood & Sterne, 2003).

Heterogeneity between studies included in the meta-analysis was assessed using the Q-test, and quantified using the I^2 statistic. I anticipated high statistical heterogeneity (Kirkbride et al., 2012), making it difficult to draw any firm conclusions from the review (Higgins & Thompson, 2002). This means that visual interpretation of forest plots is paramount in interpreting results, and the pooled estimate will play a secondary role. Nonetheless, I chose to compute and display pooled estimates, to facilitate interpretation of results and to minimise the possibility for *ad hoc* analysis (Borgenstein, Hedges, Higgins, & Rothstein, 2009).

Considering these high anticipated levels of heterogeneity, I *a priori* considered it more appropriate to use a random-effect meta-analysis. Both fixed effect and random effect methodology weight studies by the inverse of variance, but the computation of this weighting differs. Whereas a fixed effect meta-analysis only incorporates within-study variability, a random effect methodology enables the incorporation of both within-study variability and an estimate of the between-study variance (Borgenstein et al., 2009). Consequently, random-effect meta-analysis yields more conservative (wider) and realistic confidence intervals compared with a fixed-effect meta-analysis. The precise method used was originally proposed by DerSimonian and Laird (DerSimonian & Laird, 1986).

To investigate the source of this expected high heterogeneity, I used meta-regression. This technique is used to examine possible associations between covariates (latitude, urbanicity, study quality, mid-year of recruitment, inequality, freedom and trust) and incidence of psychotic disorders (Borgenstein et al., 2009). Initially, I tested univariable associations between predictor variables and psychosis incidence, and subsequently build a multivariable model using a backward stepwise regression model, where I initially entered all variables and dropped them in descending order of association (assessed by Wald p-value), until all remaining variables are significantly associated. I only report multivariable results when significant ($p<0.05$). All analyses were carried out in StataMP13 (StataCorp, 2013).

2.3.11 Cumulative evidence

To assess the quality of methodology and reporting of this systematic review and meta-analysis, and increase confidence in findings, I applied the Assessing the Methodological Quality of Systematic Reviews (AMSTAR) criteria (National Collaborating Centre for Methods and Tools, 2011). These criteria are known to have high face and construct validity, good inter-rater agreement and test-retest reliability (National Collaborating Centre for Methods and Tools, 2011), and are considerably more straightforward to report than the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) criteria on transparent reporting of systematic reviews and meta-analyses (Moher, Liberati, Tetzlaff, Altman, & Group, 2009). I followed the latter in reporting, but I didn't explicitly assesse them in this Chapter. The AMSTAR criteria are listed in Table 2.4 below, and this Table also details where in this Chapter they will be discussed. A full report on the criteria can be found in Section 2.5.2.

Table 2.4: AMSTAR criteria

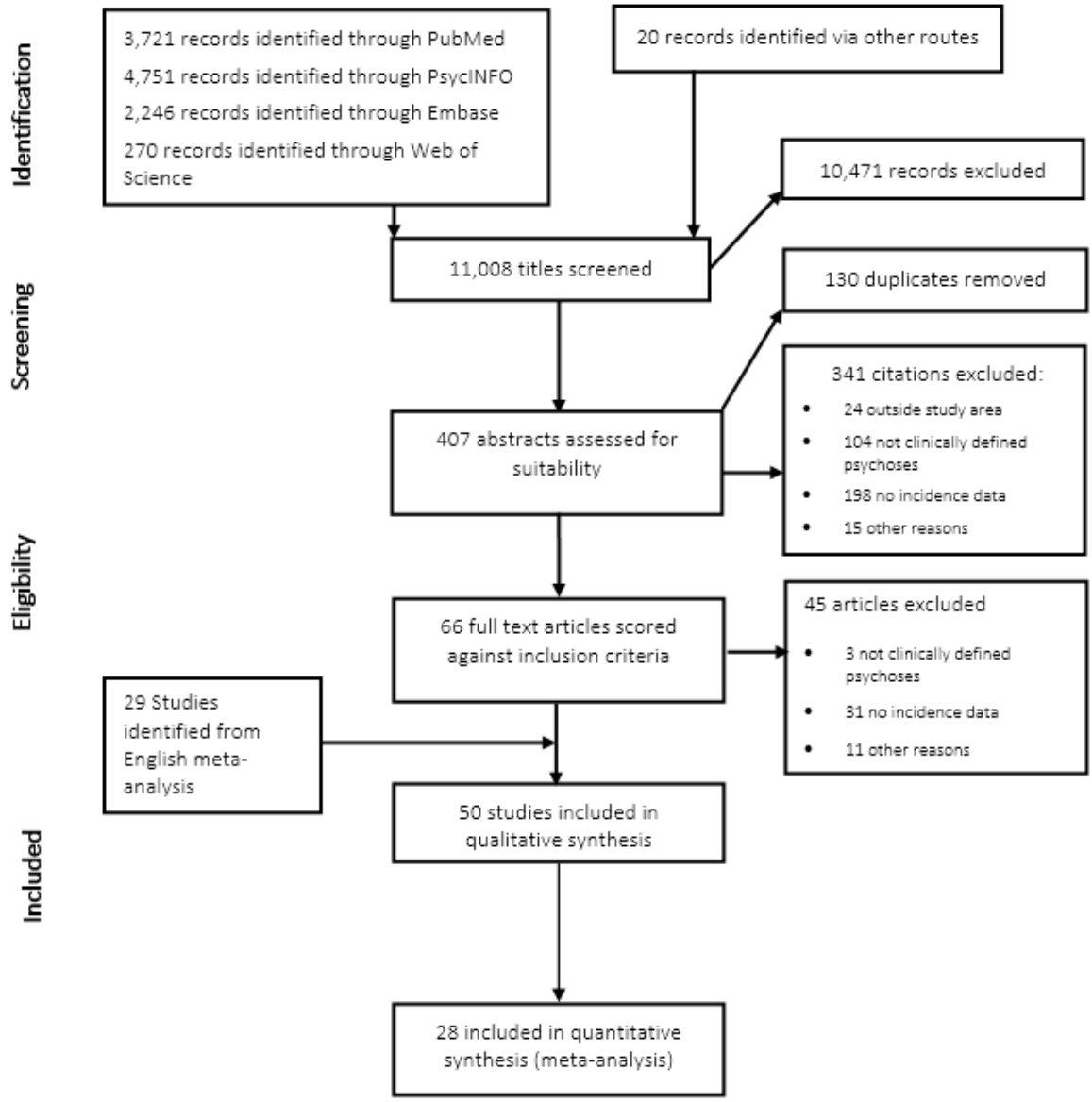
AMSTAR Item	Section
A priori design	2.3
Duplicate study selection and data extraction	2.3.4/2.3.5
Comprehensive literature search	2.3.3
Status of publication as inclusion criteria	2.3.3
Lists of studies included and excluded	2.4.1/Appendix 2B
Characteristics included studies	2.4.1
Scientific quality assessment	2.3.8 /2.4.1
Method of combining data items appropriate	2.3.9/2.4.6
Publication bias assessed	2.3.8
Conflict of interest included	Table 2.5

2.4 Results

2.4.1 Study selection

A summary of the search process can be found in the PRISMA flow chart (Figure 2.2). The search yielded 11,008 across all databases, including 20 citations identified via manual reference checking and contact with authors.

Figure 2.2: PRISMA Flowchart



The titles of these 11,008 records were screened to eliminate true negatives. When there was uncertainty regarding inclusion, the record was maintained for abstract review. During this stage, 10,471 records were excluded. Furthermore, 130 duplicates identified in more than 1 database were removed, leaving 407 abstracts for review.

I reviewed abstracts against the inclusion criteria of containing original data on the incidence of non-organic adult psychosis in Brazil, France, Italy, the Netherlands or Spain (studies from England were identified as part of an earlier review using the same methodology). This led to the exclusion of 341 citations (Figure 2.2). Abstracts not containing sufficient information to reach a conclusion were retained for full text review. I read the remaining 66 citations in full, and checked them against inclusion criteria.

This led to the exclusion of 45 citations. A basic description of these citations and the reason for their exclusion can be found in Appendix 2B.

The 21 citations meeting the inclusion criteria were combined with 29 studies from England previously identified via the original search strategy, leading to a total of 50 citations included for the systematic review, and 28 could be included in the meta-analysis. These 50 citations are summarised in Tables 5 and 6 below.

2.4.2 Yield

Of the 50 citations, most were from England (n=29; 58%), followed by The Netherlands (n=12;, 24%). Fewer studies were published in Italy (n=5; 10%), France (n=2; 4%), Spain (n=1; 2%) and Brazil (n=1; 2%). One of the citations identified was published in Dutch (Boonstra, Wunderink, De Wit, Noorthoorn, & Wiersma, 2008), whilst the remainder of citations was published in English.

Forty-three citations included a measure of uncertainty (either reported or derived), and could therefore theoretically be included in a meta-analysis. However, four citations (Bhavsar et al., 2014; Cheng et al., 2011; Kirkbride, Stubbins, & Jones, 2012; Mahmmood & Fisher, 2006) only looked at adults in the age range of early intervention services (aged up to 35). As this is the age group in which incidence peaks (Kirkbride et al., 2012), these studies were not representative of the general adult population. Including these studies would have artificially increased any pooled estimate of incidence rates and they were summarised separately. Two further citations were excluded from meta-analysis but retained for narrative review: one citation covering the armed forces only as it was not considered representative of the general adult population (Turner et al., 2006), and a further citation presenting data from both a large study already included in the meta-analysis, and from a case register, including the same patient group (Hogerzeil et al., 2014).

A number of large studies yielded several relevant citations (for example, the Aetiology and Ethnicity in Schizophrenia and Other Psychoses (AESOP) study, the East London First Episode Psychosis (ELFEP) study and a large study from The Hague). For each meta-analysis, a decision on inclusion of citations was made on the basis of the criteria presented in Section 2.3.7. This meant that some citations (Kirkbride et al., 2008; Veling et al., 2006) were not included in any meta-analyses, as data from these citations was covered in more detail in other citations. A further four citations could not be included in meta-analysis (but were retained for the systematic review) as they only presented data pertaining to a sub-group analysis covering fewer than four citations (Fearon et al., 2006;Kennedy et al., 2005; Lloyd et al., 2005; Morgan et al., 2006).

Table 2.5: Basic details of studies included in systematic review

ID	Authors	Title	Country	Year	Journal	Conflicts of interest declared
1	Boydell, J.; Van Os, J.; Lambri, M.; Castle, D.; Allardyce, J.; McCreadie, R. G.; Murray, R. M.	Incidence of schizophrenia in south-east London between 1965 and 1997	England	2003	British Journal of Psychiatry	Yes
2*	Singh, S; Wright, C; Joyce, E; Barnes, T; Burns, TR.	Developing early intervention services in the NHS: A survey to guide workforce and training needs	England	2003	Psychiatric Bulletin	Yes
3*	Veen N.D., Selten J.P., Schols, D., Laan W., Hoek, H.W.; Van der Tweel I., Kahn R.S.	Diagnostic stability in a Dutch psychosis incidence cohort	The Netherlands	2004	British Journal of Psychiatry	Yes
4	Boydell, J.; van Os, J.; McKenzie, K.; Murray, R. M.	The association of inequality with the incidence of schizophrenia-an ecological study	England	2004	Psychiatry & Psychiatric Epidemiology	Yes
5*	Proctor, S. E.; Mitford, E.; Paxton, R.	First episode psychosis: a novel methodology reveals higher than expected incidence; a reality-based population profile in Northumberland, UK	England	2004	Journal of Evaluation in Clinical Practice	Yes
6*	Singh, S. P.; Burns, T.; Amin, S.; Jones, P. B.; Harrison, G.	Acute and transient psychotic disorders: Precursors, epidemiology, course and outcome	England	2004	British Journal of Psychiatry	Yes
7	Kennedy, N.; Boydell, J.; Kalidindi, S.; Fearon, P.; Jones, P. B.; van Os, J.; Murray, R. M.	Gender differences in incidence and age at onset of mania and bipolar disorder over a 35-year period in Camberwell, England.	England	2005	American Journal of Psychiatry	Yes
8	Lloyd, T; Kennedy, N; Fearon, P; Kirkbride, JB; Mallett, R; Leff, J; Holloway, J; Harrison, G; Dazzan, P; Morgan, K; Murray, RM.; Jones, PB.	Incidence of bipolar affective disorder in three UK cities: Results from the AESOP study	England	2005	British Journal of Psychiatry	Yes
9	Nixon, N. L.; Doody, G. A.	Official psychiatric morbidity and the incidence of schizophrenia 1881-1994	England	2005	Psychological Medicine	Yes
10	Kennedy, N.; Everitt, B.; Boydell, J.; van Os, J.; Jones, P.	Incidence and distribution of first-episode mania by age: Results from a 35-year study	England	2005	Psychological Medicine	Yes
11	Drukker M., Krabbendam L., Driessen G., Van Os J.	Social disadvantage and schizophrenia. A combined neighbourhood and individual-level analysis	The Netherlands	2006	Social Psychiatry and Psychiatric Epidemiology	Yes
12	Veling W., Selten J.P., Veen N.D, Laan W., Blom J.D., Hoek H.W.	Incidence of schizophrenia among ethnic minorities in the Netherlands: a four-year first-contact study	The Netherlands	2006	Schizophrenia Research	Yes
13	Fearon, P.; Kirkbride, J. B.; Morgan, C.; Dazzan, P.; Morgan, K.; Lloyd, T.; Tarrant, J.; Lun Alan Fung, W.; Holloway, J.; Mallett, R.; Harrison, G.; Leff, J.; Jones, P. B.; Murray, R. M.; Muga, F.; Mietunen, J.; Ashby, M.; Hayhurst, H.; Craig, T.; McCabe, J.; Samele, C.; Gwenzi, E.; Sharpley, M.; Vearnals, S.; Hutchinson, G.; Burnett, R.; Kelly, J.; Orr, K.; Salvo, J.; Greenwood, K.; Raune, D.; Lambri, M.; Auer, S.;	Incidence of schizophrenia and other psychoses in ethnic minority groups: Results from the MRC AESOP Study	England	2006	Psychological Medicine	Yes

	Rohebak, P.; McIntosh, L.; Doody, G.; Window, S.; Williams, P.; Bagalkote, H.; Dow, B.; Boot, D.; Farrant, A.; Jones, S.; Simpson, J.; Moanette, R.; Sirip Suranim, P. Z.; Ruddell, M.; Brewin, J.; Medley, I.					
14*	Kirkbride, J. B.; Fearon, P.; Morgan, C.; Dazzan, P.; Morgan, K.; Tarrant, J.; Lloyd, T.; Holloway, J.; Hutchinson, G.; Leff, J. P.; Mallett, R. M.; Harrison, G. L.; Murray, R. M.; Jones, P. B.	Heterogeneity in incidence rates of schizophrenia and other psychotic syndromes: findings from the 3-center AeSOP study	England	2006	Archives of General Psychiatry	No (author affiliations only)
15	Morgan, C.; Dazzan, P.; Morgan, K.; Jones, P.; Harrison, G.; Leff, J.; Murray, R.; Fearon, P.	First episode psychosis and ethnicity: initial findings from the AESOP study	England	2006	World Psychiatry	Yes
16*	Gould, M.; Theodore, K.; Pilling, S.; Bebbington, P.; Hinton, M.; Johnson, S.	Initial treatment phase in early psychosis: Can intensive home treatment prevent admission?	England	2006	Psychiatric Bulletin	Yes
17	Turner, MA.; Finch, PJ.	Psychosis in the British Army: A 2-Year Follow-Up Study	England	2006	Military Medicine	Unknown
18	Mahmmood, M. A.; Fisher, H.; Power, P.	The incidence of first episode psychosis in inner London: findings from the Lambeth Early Onset (LEO) service	England	2006	Schizophrenia Research	No
19*	Veling W., Susser E., van Os J., Mackenback J.P., Selten J.P., Hoek H.W.	Ethnic density of neighbourhoods and incidence of psychotic disorders among immigrants	The Netherlands	2008	American Journal of Psychiatry	Yes
20*	Menezes P.R., Scazufca M., Busatto G., Coutinho L.M., McGuire P.K., Murray R.M.	Incidence of first-contact psychosis in Sao Paulo, Brazil	Brazil	2007	British Journal of Psychiatry	Yes
21	Kirkbride, J. B.; Fearon, P.; Morgan, C.; Dazzan, P.; Morgan, K.; Murray, R. M.; Jones, P. B.	Neighbourhood variation in the incidence of psychotic disorders in Southeast London	England	2007	Social Psychiatry & Psychiatric Epidemiology	No
22*	Kirkbride, J. B.; Morgan, C.; Fearon, P.; Dazzan, P.; Murray, R. M.; Jones, P. B.	Neighbourhood-level effects on psychoses: re-examining the role of context	England	2007	Psychological Medicine	Yes
23*	Pelayo-Terán J.M., Pérez-Iglesisas R., Ramírez-Bonilla M., Gonzáles-Blanch C., Martínez-García O., Pardo-García G., Rodr/guez-Sánchez J.M., Rois-Santiáñez, Tordesillas-Gutiérrez D., Mata I, Vázques-Barquero J.L., Crespo-Facorro B.	Epidemiological factors associated with treated incidence of first-episode non-affective psychosis in Cantabria: insights from the clinical programme on Early Phases of Psychosis	Spain	2008	Early Intervention in Psychiatry	Yes
24*	Coid, J. W.; Kirkbride, J. B.; Barker, D.; Cowden, F.; Stamps, R.; Yang, M.; Jones, P. B.	Raised incidence rates of all psychoses among migrant groups: findings from the East London first episode psychosis study	England	2008	Archives of General Psychiatry	Yes
25	Crebbin, K.; Mitford, E.; Paxton, R.; Turkington, D.	First-episode psychosis: an epidemiological survey comparing psychotic depression with schizophrenia	England	2008	Journal of Affective Disorders	Yes
26*	Kirkbride, J. B.; Barker, D.; Cowden, F.; Stamps, R.; Yang, M.; Jones, P. B.; Coid, J. W.	Psychoses, ethnicity and socio-economic status	England	2008	British Journal of Psychiatry	Yes
27	Kirkbride, J. B.; Boydell, J.; Ploubidis, G. B.; Morgan, C.; Dazzan, P.; McKenzie, K.; Murray, R. M.; Jones, P. B.	Testing the association between the incidence of schizophrenia and social capital in an urban area	England	2008	Psychological Medicine	Yes

28	Crebbin, K.; Mitford, E.; Paxton, R.; Turkington, D.	First-episode drug-induced psychosis: a medium term follow up study reveals a high-risk group	England	2009	Social Psychiatry & Psychiatric Epidemiology	Yes
29	Kirkbride, J. B.; Croudace, T.; Brewin, J.; Donoghue, K.; Mason, P.; Glazebrook, C.; Medley, I.; Harrison, G.; Cooper, J. E.; Doody, G. A.; Jones, P. B.	Is the incidence of psychotic disorder in decline? Epidemiological evidence from two decades of research	England	2009	International Journal of Epidemiology	Yes
30*	Reay, R.; Mitford, E.; McCabe, K.; Paxton, R.; Turkington, D.	Incidence and diagnostic diversity in first-episode psychosis	England	2009	Acta Psychiatrica Scandinavica	Yes
31*	Zandi, T; Havenaar, J.M.; Smits, M; Limburg-Okken, A.G.; Van Es, H; Chan, W; Algra, A; Kahn, R.S.; Van den Brink, W	First contact incidence of psychotic disorders among native Dutch and Moroccan immigrants in the Netherlands: influence of diagnostic bias	The Netherlands	2010	Schizophrenia Research	Yes
32*	Veling, W; Hoek, H.W.; Selten, J.P.; Susser, E	Age at migration and future risk of psychotic disorders among immigrants in the Netherlands: a 7-year incidence study	The Netherlands	2011	American Journal of Psychiatry	Yes
33	Cheng, F.; Kirkbride, J.B.; Lennox, B.R.; Perez, J.; Masson, K.; Lawrence, K.; Hill, K.; Feeley, L.; Painter, M.; Murray, G.K.; Gallagher, O.; Bullmore, E.T.; Jones, P.B.	Administrative incidence of psychosis assessed in an early intervention service in England: first epidemiological evidence from a diverse, rural and urban setting	England	2011	Psychological Medicine	Yes
34*	Tarricone, I; Mimmi, S; Paparelli, A; Rossi, E; Mori, E; Panigada, S; Carchia, G; Bandieri, V; Michetti, R; Minnena, G; Boydell, J; Morgan, C; Berardi, D	First-episode psychosis at the West Bologna Community Mental Health Centre: results of an 8-year prospective study	Italy	2012	Psychological Medicine	Yes
35*	Gigantesco, A; Lega, I; Picardi, A; SEME Collaborative Group	The Italian SEME Surveillance System of Severe mental Disorders Presenting to Community Mental Health Services	Italy	2012	Clinical Practice and Epidemiology in Mental Health	Yes
36*	Turola, M.C.; Comellini, G; Galuppi, A; Nanni, M.G; Carantoni, E; Scapoli, C	Schizophrenia in real life: courses, symptoms and functioning in an Italian population	Italy	2012	International Journal of Mental Health Systems	Yes
37*	Sutterland, A.L.; Dieleman, J; Storosum, J.G.; Voordouw, B.A.; Kroon, J; Veldhuis, J; Denys, D.A.; De Haan, L; Sturkenboom, M.C.	Annual incidence rate of schizophrenia and schizophrenia spectrum disorders in a longitudinal population-based cohort study	The Netherlands	2013	Social Psychiatry and Psychiatric Epidemiology	Yes
38*	Kroon, J.S; Wohlfart, T.D.; Dieleman, J; Sutterland, A.L.; Storosum, J.G/; Denys, D; De Haan, L, Sturkenboom, M.C.	Incidence rates and risk factors of bipolar disorder in the general population: a population-based cohort study	The Netherlands	2013	Bipolar Disorders	Yes
39*	Hardoon, S.; Hayes, J.F.; Blackburn, R.; Petersen, I.; Walters, K.; Nazareth, I.; Osborn, D.P.J	Recording of Severe Mental Illness in United Kingdom Primary Care	England	2013	PLoS ONE	Yes
40	Kirkbride, J.B.; Stubbins, C.; Jones, P.B.	Psychosis incidence through the prism of early intervention services	England	2013	British Journal of Psychiatry	Yes
41*	Tortelli, A; Morgan, C.; Szoke, A; Nascimento, A; Skurnik, N; Monduit de Caussade, E; Fain-Donabedian, E; Fridja, F; Henry, M; Ezembe, F; Murray, R.M.	Different rates of first admissions for psychosis in migrant groups in Paris	France	2014	Social Psychiatry and Psychiatric Epidemiology	Yes

42*	Szöke, A; Charpeaud, T; Galliot, A-M; Vilain, J; Richard, J-R; Leboyer, M; Llorca, P-M; Schürhoff, F.	Rural-urban variation in incidence of psychosis in France: a prospective epidemiologic study in two contrasted catchment areas	France	2014	BMC Psychiatry	Yes
43*	Lasalvia, A; Bonetto, C; Tosato, S; Zanetta, G; Cristofalo, D; Salazzari, D; Lazzarotto, L; Bertani, M; Bisolli, S; De Santi, K; Cremonese, C; De Rossi, M; Gardellini, F; Ramon, L; Zucchetto, M; Amaddeo, F; Tansella, M; Ruggeri, M; PICOS-Veneto Group	First-contact incidence of psychosis in north-eastern Italy: influence of age, gender, immigration and socioeconomic deprivation	Italy	2014	British Journal of Psychiatry	Yes
44	Hogerzeil, S.J.; van Hemert, A.M; Rosendaal, F.R; Susser, E; Hoek, H.W;	Direct comparison of first-contact versus longitudinal register-based case finding in the same population: early evidence that the incidence of schizophrenia may be three times higher than commonly reported	The Netherlands	2014	Psychological Medicine	Yes
45	Bhavsar, V; Boydell, J; Murray, R; Power, P.	Identifying aspects of neighbourhood deprivation associated with increased incidence of schizophrenia	England	2014	Schizophrenia Research	Yes
46*	Veling, W; Susser, E; Selten, J.P.; Hoek, H.W.	Social disorganisation of neighbourhoods and incidence of psychotic disorders: a 7-year first contact incidence study	The Netherlands	2014	Psychol ogical Medicine	Yes
47*	Selten, JP; van Os, J; Nolen, WA	First admissions for mood disorders in immigrants to the Netherlands	The Netherlands	2003	Social Psychiatry & Psychiatric Epidemiology	Yes
48*	Schofield, P; Ashworth, M; Jones, R.	Ethnic isolation and psychosis: re-examining the ethnic density effect	England	2011	Psychological Medicine	Yes
49*	Cocchi, A; Balbi, A; Corlito, G; Ditta, G; Di Munzio, W; Nicotera, M; Meneghelli, A; Pisano, A; Preti, A	Early intervention in psychosis: a feasibility study financed by the Italian Center on Control of Maladies	Italy	2014	Early Intervention in Psychiatry	Yes
50*	Boonstra, N; Wunderink L; de With, PHM; Noorthoorn, E; Wiersma, D	De administratieve incidentie van niet-affectieve psychosen in Friesland en Twente	The Netherlands	2008	Tijdschrift voor psychiatrie.	Yes

* Indicates studies also included in a meta-analysis.

Table 2.6: Further characteristics of studies included in the systematic review

ID	First Author	Type of incidence study	Subgroups	Diagnostic outcomes	Quality	Mid-year
1	Boydell	Case register & first contact	Time groups	Schizophrenia	7	1966/1970/1975/1980/1985/1990/1995
2*	Singh	First contact	--	Psychosis (Schizophrenia, Affective psychosis or drug-related psychosis), schizophrenia (including schizoaffective, delusional and acute and transient psychotic disorder), affective psychosis, drug-related psychosis	2	2000
3*	Veen	First contact	Incidence, and diagnosis at follow-up	Schizophrenic disorder (schizophrenia, schizophreniform disorder and schizoaffective disorder), psychotic mood disorder (major depressive disorder (with psychotic features) and bipolar disorder (with psychotic features)), other non-organic psychotic disorders (delusional disorder, brief psychotic disorder and psychotic disorder not otherwise specified), organic psychotic disorder (psychotic disorder due to a general medical condition and substance-induced psychotic disorder)	6	1998
4	Boydell	Case register +	Sex, deprivation	Schizophrenia	6	1993
5*	Proctor	Case register	Age group, sex,	All psychosis, schizophrenia-like psychosis, affective psychosis	1	1999
6*	Singh	First contact	Incidence and diagnosis at 3-year follow-up	First episode psychosis, acute onset transient psychotic disorder, schizophrenia, affective psychosis, substance-induced psychosis, non-affective psychosis, true acute onset transient psychotic disorder	6	1993
7	Kennedy	Case register	Age, sex	Bipolar I disorder, first manic episode	6	1983
8	Lloyd	First contact	Three centres, sex	Mania with or without psychotic symptoms or bipolar affective disorder	7	1998
9	Nixon	Case register +	Three time points	RDC schizophrenia, ICD-9 schizophrenia, ICD-10 schizophrenia	4	1900/1979/1993
10	Kennedy	Case register +	Age, sex, ethnicity	DSM-IV Bipolar I, bipolar affective disorder/mania/hypomania	6	1983
11	Drukker	Case register	Neighbourhood social deprivation, residential instability	Schizophrenia	3	1998
12	Veling	First contact	Generation of migrants ($1^{st}/2^{nd}$), gender	Schizophrenic disorder (schizophrenia, schizophreniform disorder and schizoaffective disorder)	5.5	1998/2001
13	Fearon	First contact	Ethnicity, sex	All psychosis, narrow schizophrenia, manic psychosis, depressive psychosis, other psychosis	6	1998
14*	Kirkbride	First contact	Three centres	Affective psychosis I, affective psychosis II, Certain' or 'very likely' Schizophrenia or paranoid psychosis (295.0-295.9 except 295.7 with or without an alternative diagnosis) Schizophrenia or paranoia	7	1998
15	Morgan	First contact	Three centres, ethnic groups (Black African and Black Caribbean only)	All psychoses, schizophrenia	6	1998

16*	Gould	First presentation	Age, sex, ethnicity (number of cases)	All psychotic illnesses (schizophrenia, schizoaffective disorder, delusional disorder, depression with psychotic symptoms and mania with psychotic symptoms)	6	2002
17	Turner	First admission	None	Non-affective psychosis, schizophrenia, delusional disorder, acute psychotic disorder, schizoaffective disorder, non-organic psychotic disorder	3	200
18	Mahmmood	First contact	Age	Psychosis	3	2003
19*	Veling	First contact	Ethnicity (incidence rates) Ethnicity x ethnic density (IRR)	Schizophrenia Spectrum Disorder (schizophrenia, schizoaffective disorder, schizophreniform disorder), major depressive disorder with psychotic features, bipolar disorder with psychotic features, delusional disorder, brief psychotic disorders, psychotic disorder NOS	5.5	1998/2003
20	Menezes	First contact	Five year age groups	Psychotic disorders, non-affective psychoses, affective psychoses	6	2003
21	Kirkbride	First contact	Standardised rates	Broad psychoses, non-affective psychoses, affective disorders	6	1998
22*	Kirkbride	First contact	Sex, age, ethnicity (incidence rates) Ethnic density, voter turnout, population density, deprivation (Adjusted IRR)	Non affective psychosis, schizophrenia, other non-affective psychosis	6	1998
23*	Pelayo-Terán	First contact	Age, Sex	Schizophrenia Spectrum disorder (schizophrenia, schizophreniform disorder, schizoaffective disorder, brief psychotic disorder, psychosis NOS)	4	2003
24*	Coid	First contact	Sex, age, ethnicity, country of birth, generation.	Schizophrenia , affective psychosis, non-affective psychosis (including schizophrenia, schizophreniform disorder, and schizoaffective disorder), other non-affective psychosis, all psychosis	7	1998
25	Crebbin	Case register	Age, sex, type of contact.	Psychotic depression, schizophrenia, all first episode psychosis	3	2002
26*	Kirkbride	First contact	Ethnicity	All psychoses, schizophrenia, other non-affective psychoses, affective psychoses	6	1998
27	Kirkbride	First contact	Social capital, trust, ethnic fragmentation, deprivation.	Schizophrenia	6	1998
28	Crebbin	Case register	Sex, type of contact	Drug induced psychosis, first-episode schizophrenia with drug misuse, first-episode schizophrenia	2	2002
29	Kirkbride	Case register +	Three time periods	All psychoses, non-affective psychoses, schizophrenia, other non-affective psychoses, substance-induced psychoses, affective psychoses, manic psychoses, depressive psychoses	6	1979/1993/1998
30*	Reay	First contact	Sex, age group	All psychosis, schizophrenia, schizophrenia spectrum disorder, bipolar, psychotic depression, affective psychosis	4	2002
31*	Zandi	First contact	CASH/CASH-CS, Ethnicity (Dutch and Moroccan only)	Suspected psychotic disorder (unconfirmed), all psychotic disorders (schizophrenia, schizophreniform disorder, schizoaffective disorder, delusional disorder brief psychotic disorder, psychotic disorder NOS, major depressive disorder with psychotic symptoms, bipolar disorder with psychotic symptoms), schizophrenic disorders (schizophrenia,	4	2003

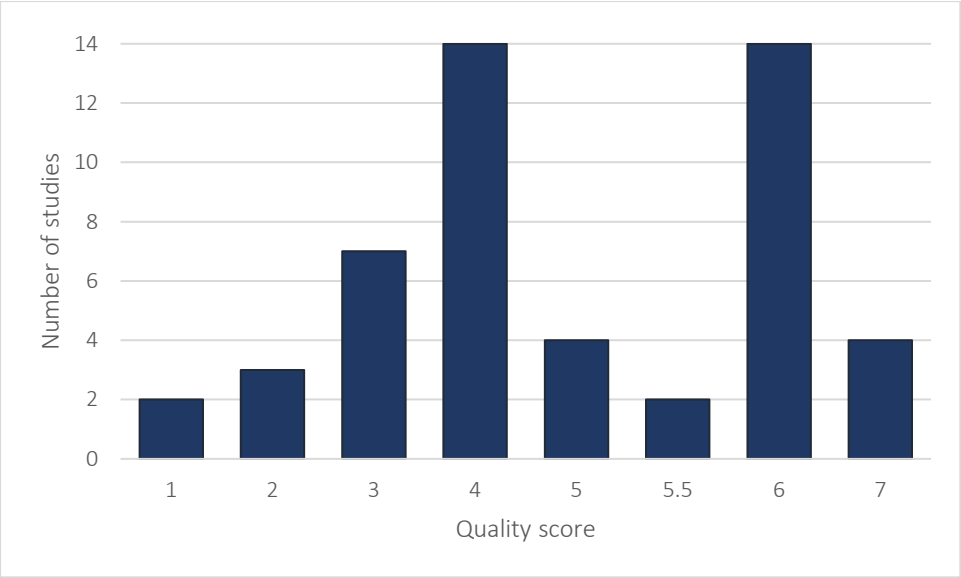
				schizoaffective disorder, schizophreniform disorder), non-psychotic disorders (mood disorders without psychosis, factitious disorder, disassociative disorder)		
32*	Veling	First contact	Age at migration, ethnicity x generation	All psychotic disorders	4	1998/2003
33	Cheng	First contact	Age, sex, ethnicity	All clinically relevant psychotic disorder (F10-39)	4	2005
34*	Tarricone	First contact	Age at onset, age at first contact, sex, ethnicity (migrant/native)	All psychoses, affective psychoses, non-affective psychoses, schizophrenia, substance-induced psychoses	5	2005
35*	Gigantesco	First contact	Geographical region	Psychotic disorders (schizophrenia, schizophreniform disorder, schizoaffective disorder, delusional disorder), bipolar I, major depressive disorder with psychotic symptoms or suicide attempts, anorexia nervosa	4	2002
36*	Turola	First diagnosis	Sex	Schizophrenia	3	1993
37*	Sutterland	First GP record	Sex, deprivation x urbanicity	Schizophrenia Spectrum Disorder (schizophrenia, schizophreniform disorder, schizoaffective disorder, delusional disorder, brief psychotic disorder, psychotic disorder NOS), schizophrenia	4	2001
38*	Kroon	First GP record	Sex, age, deprivation	Bipolar I, Bipolar II	4	2001
39*	Hardoon	First GP record	Age, sex, deprivation, urbanicity	Schizophrenia, bipolar disorder, other non-organic psychosis	4	2005
40	Kirkbride	First presentation	Early intervention service, sex, age, ethnicity	All psychotic disorders	5	2010
41*	Tortelli	First admission	Country of birth, sex, sex x country of birth	All psychoses	3	2007
42*	Szöke	First contact	Urban/rural	All psychoses, non-affective psychoses, affective psychoses	4	2011
43*	Lasalvia	First contact	Sex, age, urbanicity, socioeconomic deprivation	All psychoses, non-affective psychoses, affective psychoses, schizophrenia, bipolar disorder/psychotic mania, psychotic depression	5	2006
44	Hogerzeil	Case register	Sex, age	Schizophrenia	4	2003
45	Bhavsar	First contact	Age, sex, deprivation (subdomains too), proportion of BME, population density, ethnic density (Black African/Black Caribbean)	Schizophrenia	5	2004
46*	Veling	First contact	Social disorganisation: socio-economic level, residential mobility, ethnic diversity, proportion of single person households, population density, crime rate, voter turnout + cumulative.	Schizophrenia Spectrum Disorder, major depressive disorder with psychotic features, bipolar disorder with psychotic features, delusional disorder, brief psychotic disorder, psychotic disorder NOS	4	1998/2003
47*	Selten	Case register	Country of birth	Manic-depressive psychosis (depressed type (major depressive disorder)), manic depressive psychosis (circular type (Bipolar-I))	3	1993
48*	Schofield	First GP record	Ethnicity, ethnicity x ethnic density (IRR only)	All psychotic disorders	4	2001
49*	Cocchi	First contact	None	Schizophrenia and related syndromes (F20-29)	1	2008
50*	Boonstra	First contact	None	Non-affective psychoses	2	2002

2.4.3 Descriptive statistics: covariates

2.4.3.1 Study quality

The median study quality across the 50 citations included in this review was 4 (interquartile range (IQR): 4-6). The scores followed an approximately bimodal distribution, with 14 citations scoring a 4 or 6 (Figure 2.3 below). Citations scoring a 6 most frequently did not have any blinding of clinicians to demographic variables (n=10; 71.4%), and citations scoring a 4 most commonly failed to use a standardised research diagnosis (n=10; 71.4%), to carry out a leakage study (n=9, 64.3%), as well as to blind clinicians to demographic variables (n=13; 92.9%). There was a weak negative correlation between study quality and year of publication (-0.27), but this did not reach statistical significance (p=0.06).

Figure 2.3: Total quality score, by study



2.4.3.2 Type of incidence study

I recorded type of incidence study for all studies included in the systematic review. As can be seen in Table 2.7, 60% of citations (n=30) were first contact studies, and a further 12 (24%) were case register studies[2].

Table 2.7: Distribution of study type

Type of study	N	%
First contact	30	60
Case register	12	24
First GP record or diagnosis	5	10
First presentation	2	4
First admission	2	4

[2] One citation (Boydell et al., 2003) reported data from both a case register and a first contact study.

2.4.3.3 Mid-year of case ascertainment

For citations included in the meta-analyses (n=28) mid-year case ascertainment ranged from 1993 to 2011 (Table 2.6), although the wider systematic review included some citations including data from as far back as 1900 (Nixon & Doody, 2005).

2.4.3.4 World Values Survey levels of self-reported trust and freedom by nation

A total of 7,631 responses were recorded for the six countries included in this analysis. Overall, 31.6% of participants agreed that most people can be trusted, although this percentage differed significantly (χ^2: $1.9*10^3$, $p<0.01$) across countries, ranging from only 6.5% in Brazil to 66.1% in the Netherlands (Table 2.8). When asked about the perceived degree of free choice and control over their lives, ranging from no choice at all (1) to a great deal of choice (10) the mean value across the sample was 7.0 (standard deviation: 2.0). The range was much narrower across countries: from 6.3 (2.1) in Italy to 7.7 (2.4) in Brazil, although still statistically significantly different (F-statistic: 61.0, $p<0.01$) (Table 2.8).

Table 2.8: World Values Survey data and Gini index per country

Country	Wold Values Survey				Gini index	
	N (respondents)	Year	Trust[1]	Freedom[2]	Year	Gini index (0-100)[3]
The Netherlands	1,902	2010-2012	1,258 (66.1)	6.9 (.14)	2010	28.9
United Kingdom	1,041	2005-2007	311 (29.9)	7.3 (1.9)	2010	38.0
Italy	1,012	2005-2007	278 (27.5)	6.3 (2.1)	2010	35.5
Spain	1,189	2010-2012	225 (18.9)	7.0 (1.9)	2010	35.8
France	1,001	2005-2007	186 (18.6)	6.7 (2.1)	2005	31.7
Brazil	1,486	2010-2012	97 (6.5)	7.7 (2.4)	2005	52.7
Overall	7,631		2,385 (31.6)	7.0 (2.0)		
χ^2/F-statistic			$\chi^{2:}$ $1.9*10^3$	F: 61.0		
p-value			<0.01	<0.01		

[1] Number of people (percentage) stating that 'most people can be trusted'
[2] Mean (standard deviation). Range: 1 (no choice at all) – 10 (a great deal of choice)
[3] Where 0 is perfect equality and 100 is perfect inequality.

2.4.3.5 Gini index of income inequality, by country

Levels of inequality varied substantially across countries. Brazil was the country where income was least equally distributed, with a Gini index of 52.7. There was less variation across the European settings: the United Kingdom was the most unequal country in the European subset of the sample (Gini index: 38.0), and The Netherlands was the most equal country in the sample with a Gini index of 28.9 (Table 2.8).

2.4.3.6 Urbanicity

The ranking of urbanicity based on population numbers or density can be found in Table 2.9 below. Inner London was the most urban setting, and Northumberland the most rural.

2.4.3.7 Latitude

Latitude of the various settings can be found in Table 2.9 below. For cities, the difference in latitude between settings (different London boroughs, inner and outer Paris) wasn't approximated as such

differences are minimal. For Friesland (53.1°N) and Twente (52.3°N) in the Netherlands, I averaged the latitude between the two settings as incidence data was not presented for the two settings separately.

Table 2.9: Urbanicity ranking and latitude

Study ID	Setting (study)	Setting	Country	Urbanicity	Latitude (°N)
16	Camden & Islington	Inner London	England	1	50.5
14, 21, 22, 48	South-East London	Inner London	England	1	50.5
22,26	East London	Inner London	England	1	50.5
41	20th District	Inner Paris	France	2	48.9
49		Rome	Italy	3	41.9
2	West and South-West London	Outer London	England	4	50.5
42	Val-de-Marne	Outer Paris	France	5	48.9
49		Milan	Italy	6	45.5
3, 12, 19, 31, 46		The Hague	Netherlands	7	52.1
14, 21, 22		Bristol	England	8	51.5
34		Bologna	Italy	9	44.5
31		Utrecht	Netherlands	10	52.1
6, 14, 22		Nottingham	England	11	53.0
43		Veneto	Italy	12	45.7
49		Salerno	Italy	13	40.7
11		Maastricht	Netherlands	14	50.9
36		Ferrara	Italy	15	44.8
49		Cantanzaro	Italy	16	38.9
49		Grosetto	Italy	17	42.8
23		Cantabria	Spain	18	43.3
50		Friesland & Twente	Netherlands	19	52.7
42		Puy de Dome	France	20	45.7
5,30		Northumberland	England	21	55.2

2.4.4 Narrative synthesis, by country

As can be seen in Table 2.10 below, the diagnostic emphasis differed slightly per country: English citations predominantly reported on all psychotic disorders and schizophrenia, whereas in Dutch and Italian citations non-affective disorders were the diagnostic outcome most frequently studied.

Table 2.10: Diagnostic outcomes per country

Country	All psychotic disorders *n(%)*	Non-affective disorders *n(%)*	Schizophrenia *n(%)*	Affective disorders *n(%)*	Bipolar disorder *n(%)*	Psychotic depression *n(%)*	Substance induced disorders *n(%)*	Symptoms *n(%)*	Other *n(%)*	Total *n(%)*
England	17 (22.6)	10 (13.3)	18 (24)	8 (10.7)	7 (9.3)	4 (5.3)	4 (5.3)	-	7 (9.3)	75 (100)
Netherlands	3 (10.7)	6 (21.4)	4 (14.3)	1 (3.6)	4 (14.3)	3 (10.7)	1 (3.6)	-	6 (21.4)	28 (100)
Italy	2 (11.8)	4 (23.5)	3 (17.6)	2 (11.8)	2 (11.8)	2 (11.8)	1 (5.9)	-	1 (5.9)	17 (100)
France	3 (42.9)	2 (28.6)	-	2 (28.6)	-	-	-	-	-	7 (100)
Spain	1 (33.3)	1 (33.3)	-	1 (33.3)	-	-	-	-	-	3 (100)
Brazil	1 (33.3)	1 (33.3)	-	1 (33.3)	-	-	-	-	-	1 (100)
All countries	**27 (20.3)**	**24 (18.0)**	**25 (18.8)**	**15 (11.2)**	**13 (9.8)**	**9 (6.7)**	**6 (4.5)**	**-**	**14 (10.5)**	**133 (100)**

2.4.4.1 England

The majority of citations in the systematic review (n=29; 58%) were set in England. These citations came from a wide variety of settings, varying from the most urban to the most rural setting in the review, and covering a wide range of diagnostic outcomes and subgroups. Eight citations covered the AESOP study (Fearon et al., 2006; Kirkbride et al., 2009; Kirkbride et al., 2008; Kirkbride et al., 2007; Kirkbride et al., 2006, 2007; Lloyd et al., 2005; Morgan et al., 2006), and two citations (Coid et al., 2008; Kirkbride et al., 2008) reported data from the ELFEP study. A further four citation each covered the Camberwell study (Boydell et al., 2003; Boydell, van Os, McKenzie, & Murray, 2004; Kennedy et al., 2005; Kennedy et al., 2005) and the Personal Assessment and Crisis Evaluation (PACE) study (Crebbin, Mitford, Paxton, & Turkington, 2008, 2009; Proctor, Mitford, & Paxton, 2004; Reay, Mitford, McCabe, Paxton, & Turkington, 2010).

The crude incidence of all psychotic disorders varied from 21.0 (95%CI: 18.6 -23.5) per 100,000 person-years reported in Northumberland (Nixon & Doody, 2005; Singh, Burns, Amin, Jones, & Harrison, 2004; Singh, Wright, Joyce, Barnes, & Burns, 2003) to around 50 or higher in Southeast (Morgan et al., 2006, Schofield et al, 2011), and East London (Coid et al., 2008; Kirkbride et al., 2008).

The incidence of non-affective disorders was lowest in a 2-year follow-up study conducted in the British army at 11.0 (Turner et al., 2006), and highest in East (Coid et al., 2008) and Southeast London (Kirkbride et al., 2007) at around 37 per 100,000 person-years. Rates were between 14 and 23 per 100,000 person-years elsewhere (Kirkbride et al., 2006, 2007; Proctor et al., 2004; Reay et al., 2010; Singh et al., 2004).

The most frequently reported diagnostic outcome in England was schizophrenia (n=18; 24%, Table 2.10). Crude incidence varied more than ten-fold, from 4.4 (95%CI: 3.4-5.7) per 100,000 person years (Reay et al., 2010) to 54.6 per 100,000 person-years (Bhavsar et al., 2014). The latter study was conducted in Lambeth, South-East London and solely included patients aged 16-35 (Bhavsar et al., 2014), whereas the former study included all patients aged 16 or over and was set in Northumberland (Reay et al., 2010). These two settings represented the most extreme settings in the entire dataset in terms of urbanicity (see Table 2.10), and incidence of psychotic disorders, including schizophrenia, was known to peak in those under the age of 35 (Kirkbride et al., 2012). Excluding these two extremes, reported crude incidence still varied almost fivefold from 7.1 (95%CI: 5.5-9.1) (Singh et al., 2003) to 32.4 (95%CI: 28.9-36.5) (Coid et al., 2008) per 100,000 person-years.

The four citations reporting only rates the under-35s were all in England. Rates of all psychotic disorders ranged from 42.6 (95%CI: 38.4-47.2) (Kirkbride et al., 2012) to 100.0 (95%CI: 89.4-111.9) (Mahmmood & Fisher, 2006) per 100,000 person-years, with a third study reporting a rate of 50.0 (95%CI: 44.5-56.2) (Cheng et al., 2011). The fourth citation reported an incidence rate of schizophrenia of 54.6 (95%CI: 49.5-60.2) (Bhavsar et al., 2014) per 100,000 person-years.

Incidence of affective disorders in England was lower, varying from 5.2 per 100,000 person-years in Nottingham (Singh et al., 2004), to 13.4 In Southeast London (Kirkbride et al., 2007) and 13.5 in East London (Coid et al., 2008). Three further citations reported rates of between 8 and 10 per 100,000 person-years (Kirkbride et al., 2006; Proctor et al., 2004; Reay et al., 2010). Within the affective disorders, a GP registry study reported a crude incidence rate of bipolar disorder of 15.0 (95%CI: 14.4-15.5) per 100,000 person-years (Hardoon et al., 2013), whereas another citation reported a rate of 3.7 (95%CI: 2.7-4.8). This citation also reported an incidence of psychotic depression of 6.0 (95%CI: 4.9-7.2) per 100,000 person-years (Reay et al., 2010)

2.4.4.2 The Netherlands

Compared with England, studies from the Netherlands presented a less varied epidemiological landscape. Whilst most diagnostic outcomes were covered (Table 2.10), this was often not by more than three or four studies. Of the twelve Dutch citations included in the analysis, five were from the same study based in the city of The Hague (Veen et al., 2004; Veling et al., 2015; Veling et al., 2006, 2011; Veling, Hoek, et al., 2008), and a further citation compared data from this first-contact study to the Longitudinal Psychiatric Register of The Hague (a case register) (Hogerzeil et al., 2014). A first contact study from Utrecht reported a crude rate of all psychotic disorders of 29.5 (95%CI: 23.9-36.8) per 100,000 person-years (Zandi et al., 2010), which was slightly lower than the only other study investigating this outcome, where a rate of 33.1 per 100,000 (95%CI: 30.6-35.8) was derived (Veling et al., 2011). The rate of non-affective psychosis was similar across settings: the only two studies investigating this outcome (the study from The Hague and a study conducted in two rural areas) both reported an incidence of approximately 22 per 100,000 person years (Boonstra, Wunderink, de Wit, Noorthoorn, & Wiersma, 2008; Veling et al., 2006; Veling, Hoek, et al., 2008). The case-register reported the highest incidence of schizophrenia at 69 (95%CI: 64-74) per 100,000 person-years. This was more than twice as high as the final rate from the study from The Hague reporting rates on the basis of first contact with mental health services, which stood at 21 (95%CI: 18-24) per 100,000 person-years (Hogerzeil et al., 2014; Veling et al., 2015; Veling et al., 2011). A national medical records study identified an incidence of schizophrenia spectrum disorders of 22 (95%CI: 19-24) per 100,00 person-years and of schizophrenia of 12 (95%CI: 10-14) per 100,000 person-years (Sutterland et al., 2013).

In terms of the affective disorders, the first two years of the The Hague study report a crude incidence of 3.3 per 100,000 person-years (Veen et al., 2004), and the overall study reported a crude incidence of bipolar disorder of 1.3 (95%CI: 0.9-1.9) and of psychotic depression of 1.4 (95%CI: 0.9-1.8) per 100,000 person-years (Veling, Susser, van Os, et al., 2008). A national cohort study reported an incidence rate of bipolar disorder of 9.8 (95%CI: 8.7-11.0) per 100,00 person-years (Selten, van Os, & Nolen, 2003), and further GP record study reported an incidence rate of 7.0 (95%CI: 5.7-8.3) (Kroon et al., 2013).

2.4.4.3 Italy

Interest in incidence studies in Italy appeared to have increased significantly in recent years, with all five citations from this country published in 2012 or later (Cocchi et al., 2015; Gigantesco et al., 2012; Lasalvia et al., 2014; Tarricone et al., 2012; Turola et al., 2012). Overall rates of all psychotic disorders appeared to be lower in Italy than in England and the Netherlands, with the highest reported rate of 23.6 (95%CI: 20.2-27.6) per 100,000 person years (Lasalvia et al., 2014) at the lower end of what was reported in England and The Netherlands. The incidence of non-affective psychoses (the diagnostic outcome most frequently studied, n=4, Table 2.10) varied almost fivefold, from 3.0 (95%CI: 1.0-9.4) in Salerno (Cocchi et al., 2015) to 14.3 (95%CI: 13.0-15.8) per 100,000 person years in the Veneto region (Lasalvia et al., 2014). Affective disorders also varied from a median of 1.7 (IQR: 0.9-3.0) per 100,000 person-years (Tarricone et al., 2012) to 3.8 (95%CI: 3.1-4.6) (Lasalvia et al., 2014) across the two citations that reported them.

2.4.4.4 France

Two French citations reported details from three different settings: an inner-city area in Paris, a Parisian suburb and a rural area in central France. Rates of first admission of all psychotic disorders in inner Paris (31.5 per 100,000 person-years; 95%CI: 28.0-35.5 (Tortelli et al., 2014)) and first contact in suburban Paris (36.2; 95%CI: 29.8-44.1) (Szöke et al., 2014) were approximately twice as high as crude incidence rate in rural Central France (17.1; 95%CI: 12.5-23.4 (Szöke et al., 2014)). Rates of non-affective disorder were only reported for suburban Paris (36.4 per 100,000 person-years) and rural Central France (11.5) as were rates of affective disorders (14.2 and 5.7 respectively) (Szöke et al., 2014).

2.4.4.5 Spain

The only Spanish citation identified was conducted in the rural region of Cantabria in Northern Spain, and reported an overall crude incidence of schizophrenia spectrum disorders of 13.7 (95%CI: 11.7-16.1) per 100,000 person years (Pelayo-Terán et al., 2008).

2.4.4.6 Brazil

A single citation on incidence of psychotic disorders was identified in Brazil. This study was carried out in São Paulo, and reported a crude rate of all psychoses of 15.8 per 100,000 person years (95%CI: 14.3-17.4), which is lower than the authors expected considering the size of the city (Menezes & Scazufca, 2007).

2.4.5 Meta-analysis and meta-regression per diagnostic outcome

Table 2.11 below is an overview of the meta-analyses conducted. For three different analyses the sample size of citations was limited to four, and for no analysis the sample size exceeded twenty. Heterogeneity was extremely high across analyses (at least 78.4%). No country-specific meta-analyses were carried out, as sample sizes were insufficiently large for all countries apart from England. However, in the overall analyses of all psychotic disorders and non-affective disorders citations were displayed and summarised by country.

Table 2.11: Overview of meta-analyses

Country	Diagnostic Outcome	Subgroup	Studies included (ID)	N	I^2 (%)
All	All psychoses	All	2, 5, 6,16, 21, 24, 30, 31, 32, 34, 35, 41, 42, 43, 48	20	97.4
All	All psychoses	Sex	24, 32, 34, 43	4	93.8
All	All psychoses	Age	14, 24, 43, 46	4	n/a
All	All psychoses	Ethnicity	26, 31, 32, 41, 43, 48	6	78.4
All	Non-affective psychosis	All	3, 5, 6, 22, 23, 30, 35, 37, 41, 42, 43, 49, 50	20	96.0
All	Schizophrenia	All	6, 11, 22, 24, 30, 36, 37, 39, 43	10	98.3
All	Schizophrenia	Sex	22, 36, 39, 43	4	93.9
All	Affective psychosis	All	3, 5, 6, 30, 41, 42, 43	11	93.9
All	Bipolar disorder	All	19, 30, 35, 38, 39, 43, 47	9	99.0

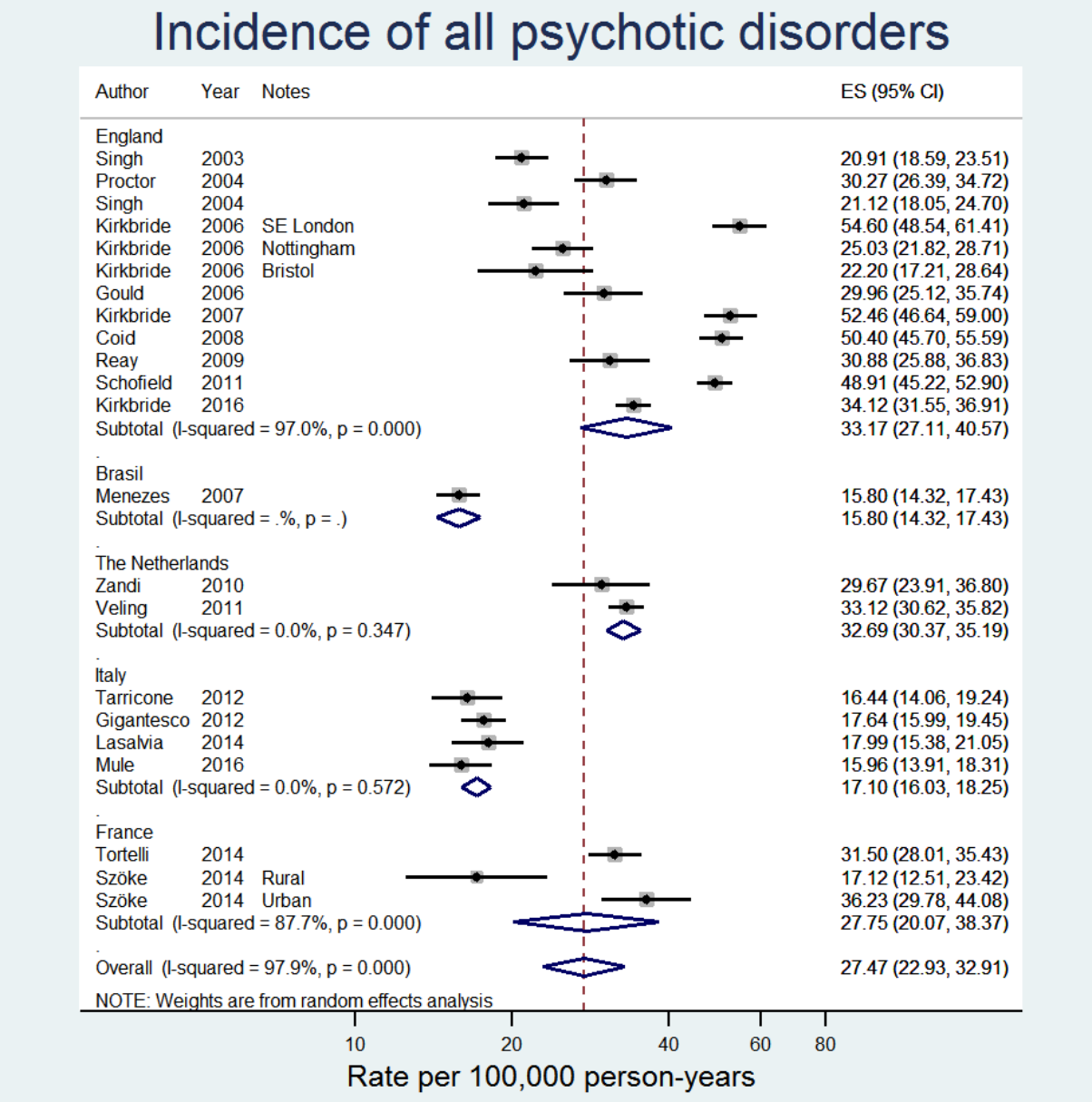
Some citations provided data from more than one centre or diagnostic outcome in the same category, and I included these centres in any meta-analyses as separate estimates of incidence (Cocchi et al., 2015; Kirkbride et al., 2006; Kroon et al., 2013). The sample size for some analyses therefore exceeded the number of studies included (all psychoses, non-affective psychosis, affective psychosis, and bipolar disorder).

2.4.5.1 All psychotic disorders, overall rates

This meta-analysis included seventeen citations across 20 settings (Coid et al., 2008; Gigantesco et al., 2012; Gould et al., 2006; Kirkbride et al., 2007; Kirkbride et al., 2006; Lasalvia et al., 2014; Menezes & Scazufca, 2007; Proctor et al., 2004; Reay et al., 2010; Schofield et al., 2011; Singh et al., 2004; Singh et al., 2003; Szöke et al., 2014; Tarricone et al., 2012; Tortelli et al., 2014; Veling et al., 2011; Zandi et al., 2010). Rates varied close to four-fold, from 15.8 (95%CI: 14.3-17.4) per 100,000 person years in São Paulo (Menezes & Scazufca, 2007) to 54.6 (95%CI: 48.5-61.4) in Southeast London (Kirkbride et al., 2006). The overall pooled incidence rate of all psychotic disorders was 28.3 (95%CI: 23.4-34.3) per 100,000 person-years, but heterogeneity remained extremely high (I^2: 97.4%), and this appeared to be particularly driven by England and France (Figure 2.4). No evidence of small-study effects was found (β: -6.90; p=0.11).

When displaying results per country, the pooled incidence in England (33.0 per 100,000 person-years; 95%CI: 26.2-41.7) and the Netherlands (32.7; 95%CI: 30.4-35.2) was slightly higher than in France (27.8; 95%CI: 20.1-38.4) (Figure 2.4). The pooled incidence from Italy was lowest at 18.9 (95%CI: 15.6-23.0) per 100,000 person-years. Incidence rates varied substantially within countries.

Figure 2.4: Forest plot of the incidence of all psychotic disorders, grouped by country.



As can be seen in Table 2.12 below, only freedom and urbanicity were significantly associated with the incidence of all psychotic disorders in a univariable meta-regression. For every 1 (out of 10) unit increase in self-reported freedom the IRR increased by 1.71 (95%CI: 1.08-2.71). Fore very rank increase in urbanicity, the IRR increased by 1.03 (95%CI: 1.00-1.05). The multivariable meta-regression model for all psychotic disorders reached significance when only urbanicity and latitude were remaining. The overall model explained 46.0% of the between-study variance (adjusted R^2). When adjusting for latitude, the magnitude of the association between incidence of psychosis and urbanicity did not change (Table 2.12) and every one degree further from the equator was independently associated with a higher IRR of psychotic disorders after adjusting for level of urbanicity (IRR: 1.06, 95%CI: 1.00-1.11) (Table 2.12).

Table 2.12: Meta-regression: univariable and multivariable results for all psychotic disorders (n=20)

Variable	Univariable IRR (95% CI)	Univariable p-value	Multivariable IRR (95%CI)	Multivariable p-value
Mid-year	0.99 (0.95-1.03)	0.58	n/a	n/a
Quality	1.06 (0.95-1.18)	0.33	n/a	n/a
GINI-index	1.02 (0.95-1.08)	0.60	n/a	n/a
Freedom	**1.71 (1.08-2.71)**	**0.03**	n/a	n/a
Trust	1.00 (0.99-1.01)	0.66	n/a	n/a
Urbanicity[1]	**1.03 (1.00-1.05)**	**0.02**	**1.03 (1.01-1.05)**	**<0.01**
Latitude	1.04 (0.97-1.10)	0.25	**1.06 (1.00-1.11)**	**0.04**
Type	1.08 (0.91-1.28)	0.36	n/a	n/a

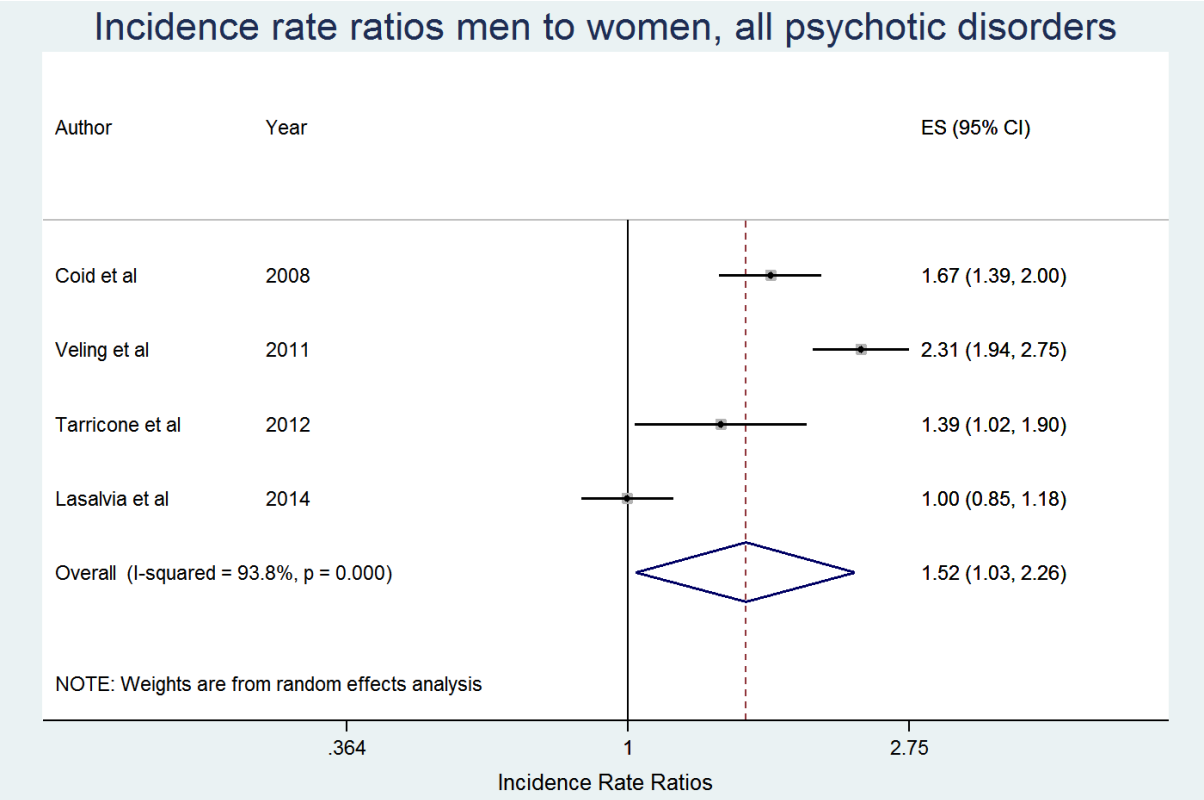
IRRs in **bold** are significant (p<0.0.5)

[1] IRRs for urbanicity are reversed (1/IRR) to aid interpretability

2.4.5.2 All psychotic disorders, rates by sex

Only four citations reported the incidence of all psychotic disorders for men and women separately (Coid et al., 2008; Lasalvia et al., 2014; Tarricone et al., 2012; Veling et al., 2011). In general, rates of psychosis were higher in men than in women. IRRs ranged from 1.00 (95%CI: 0.85-1.18) in the Veneto region (Lasalvia et al., 2014) to 2.31 (95%CI: 1.94-2.75) in The Hague (Veling et al., 2011), with an overall pooled IRR of 1.52 (95%CI: 1.03-2.26) (Figure 2.5). Variance was high (I^2: 93.8%)

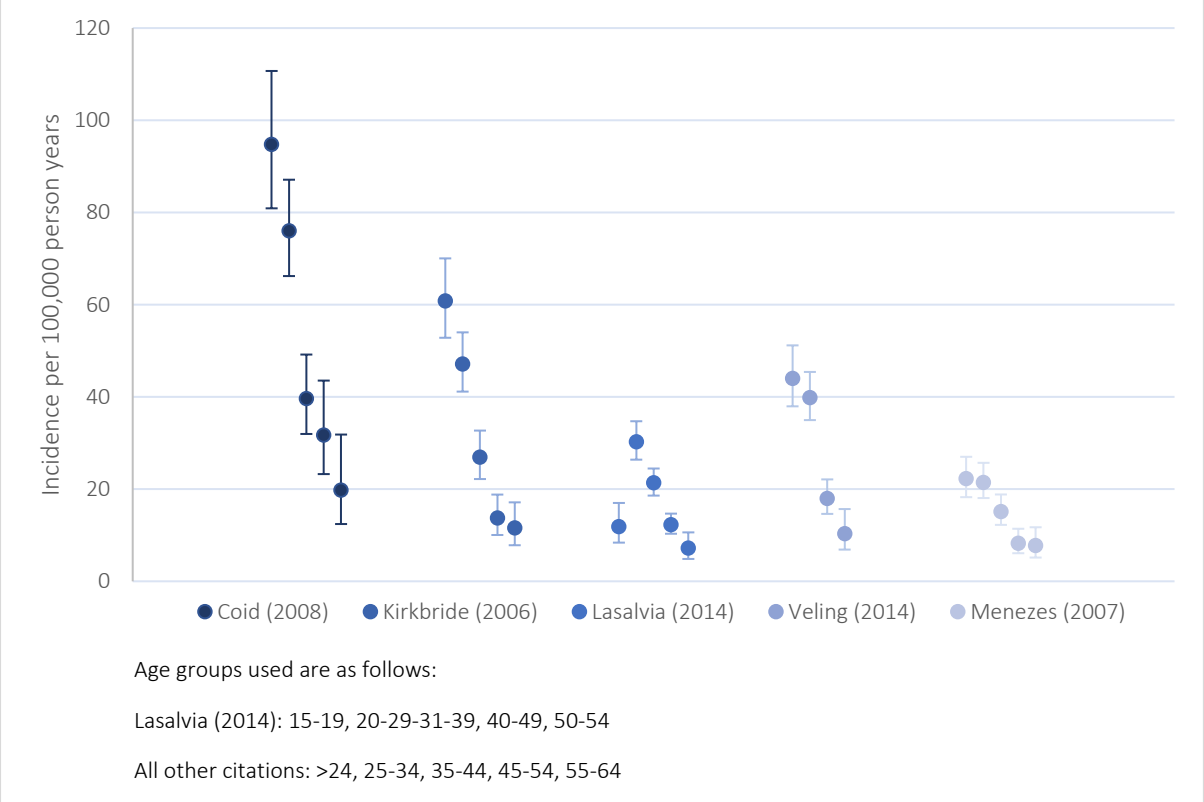
Figure 2.5: Forest plot of incidence rate ratios men-women of all psychotic disorders.



2.4.5.3 All psychotic disorders, rates by age group

Four citations reported the incidence of psychosis by age group (Coid et al., 2008; Kirkbride et al., 2006; Menezes & Scazufca, 2007; Veling, Susser, Selten, & Hoek, 2014) (with JBK providing the additional data needed for (Coid et al., 2008)). Unfortunately, it was impossible to recode all data to the same age-groups. Three citations (Coid et al., 2008; Kirkbride et al., 2006; Veling et al., 2015) could all be coded in the same ten-year age groups (under 24, 25-34, 35-44, 45-54 and 55-64) and one citation (Lasalvia et al., 2014) computed rates across different age groups (15-19, 20-29, 30-39, 40-49 and 50-54). Figure 2.6 below shows the incidence by age group, for each citation. The age groups are aligned from youngest to oldest, for each study. In all citations (apart from Lasalvia et al) incidence peaked in younger age groups, and declined over the life course.

Figure 2.6: Incidence by age group (youngest to oldest) and citation.



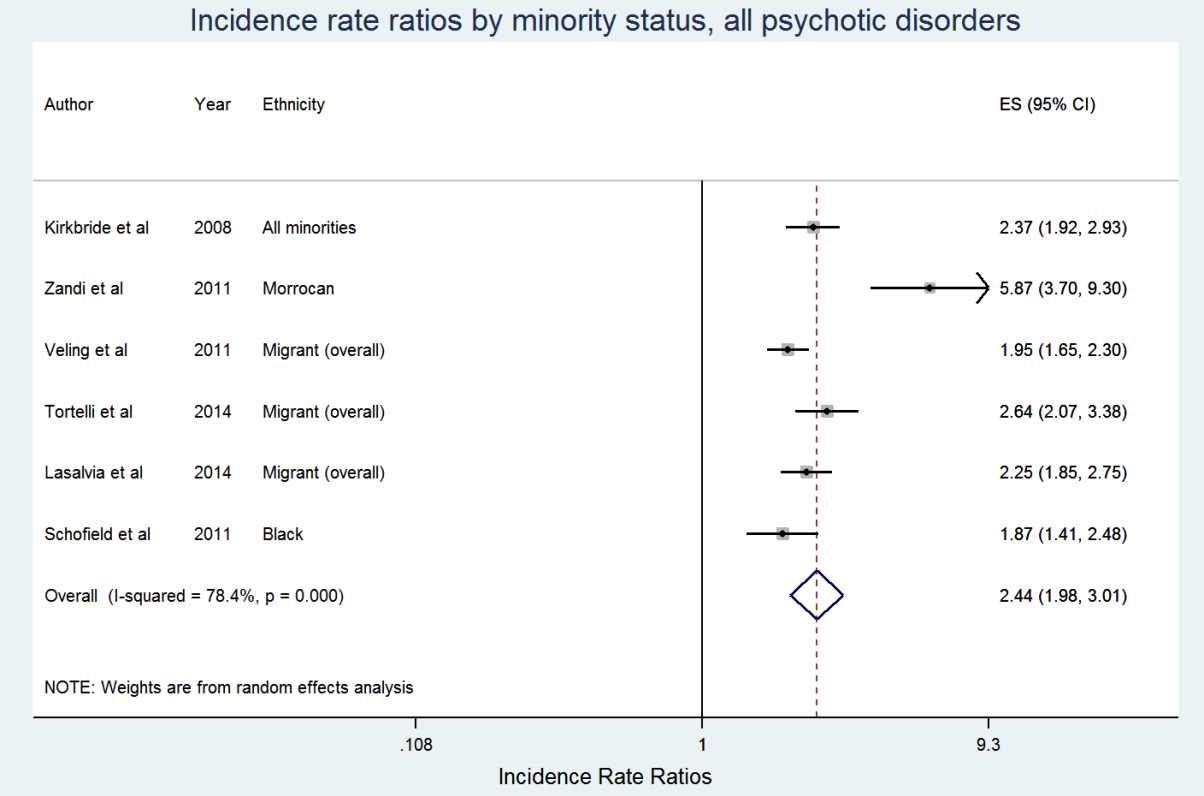
2.4.5.4 All psychotic disorders, rates by ethnicity

Six citations reported the incidence of all psychotic disorder by ethnicity (Kirkbride et al., 2008; Lasalvia et al., 2014; Schofield et al., 2011; Tortelli et al., 2014; Veling et al., 2011; Zandi et al., 2010). One citation reported the rates for the White British majority and various ethnic minorities (Kirkbride et al., 2008), one GP-database study for the White British majority and Black minorities (Schofield et al., 2011), a further study compared the White Dutch majority with various ethnic minorities (Veling et al., 2011) and a further citation looked at Moroccan minorities in the Netherlands specifically (Zandi et al., 2010). One Italian

citation looked at rates in migrant groups overall (Lasalvia et al., 2014), and a French citation examined rates of psychosis in various migrant groups (Tortelli et al., 2014).

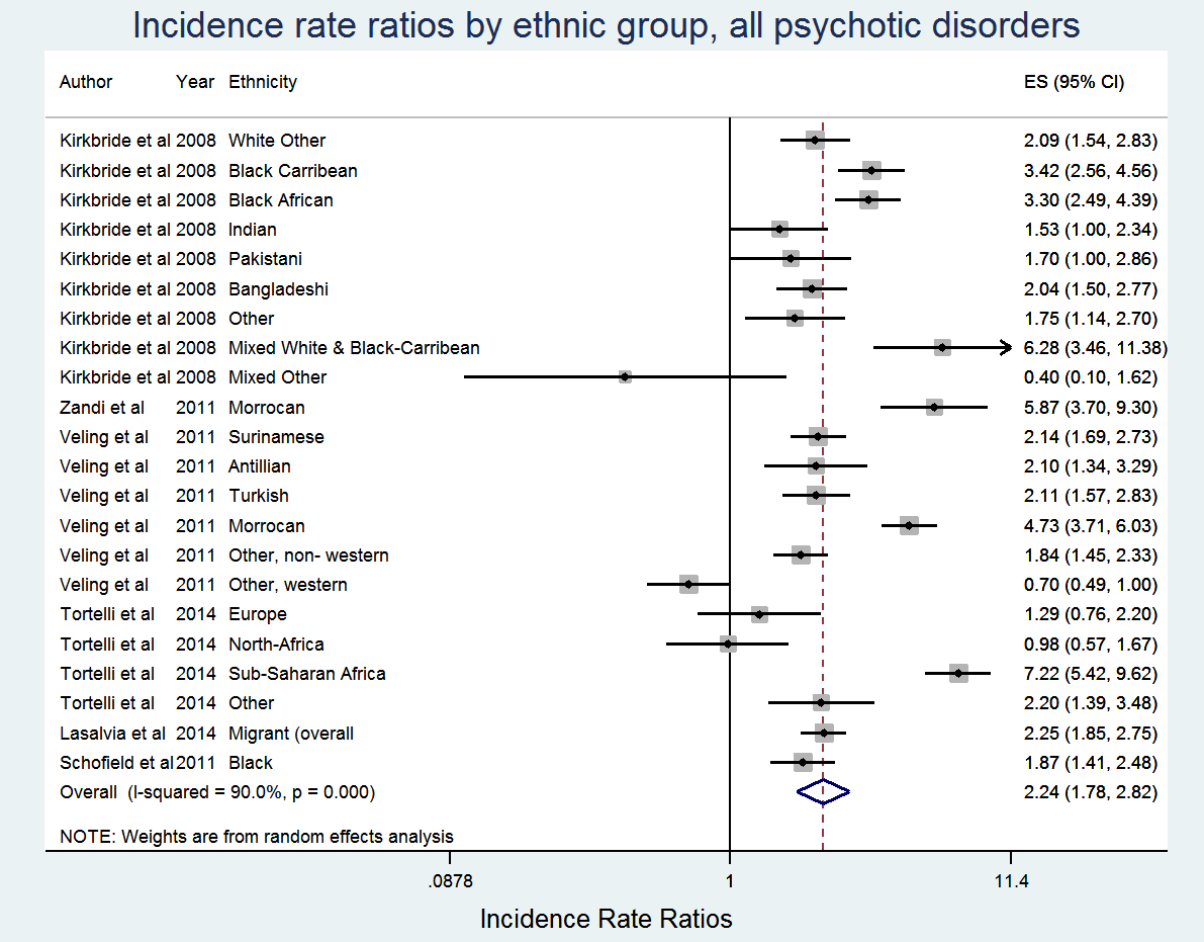
The pooled IRR of all ethnic minorities compared with the ethnic majority was 2.44 (95%CI: 1.98-3.01) (Figure 2.7). There was substantial variance in IRRs, ranging from 1.87 (95%CI: 1.41-2.48) for the Black group in Southeast London (Schofield et al., 2011) to 5.87 (95%CI: 3.70-9.30) in the Moroccan group in Utrecht (Zandi et al., 2010).

Figure 2.7: Forest plot of incidence rate ratios by broad ethnic group, all psychotic disorders



When unpacking these results further (see Figure 2.8 below), it emerged that in English studies, those of mixed Black Caribbean and White background and Black-African or -Caribbean background experienced the highest excess risk of psychotic disorders (albeit with wide confidence intervals) (Kirkbride et al., 2008; Schofield et al., 2011), whereas in Dutch studies the Moroccan groups experienced the highest excess risk (Veling et al., 2011; Zandi et al., 2010). In France, those from Sub-Saharan Africa were at the highest risk (Tortelli et al., 2014). The Italian citation included in this meta-analysis found an IRR of 2.25 (95%CI: 1.85-2.75) for all migrant groups compared with the Italian majority (Lasalvia et al., 2014).

Figure 2.8: Forest plot of incidence rate ratios by ethnic group, all psychosis

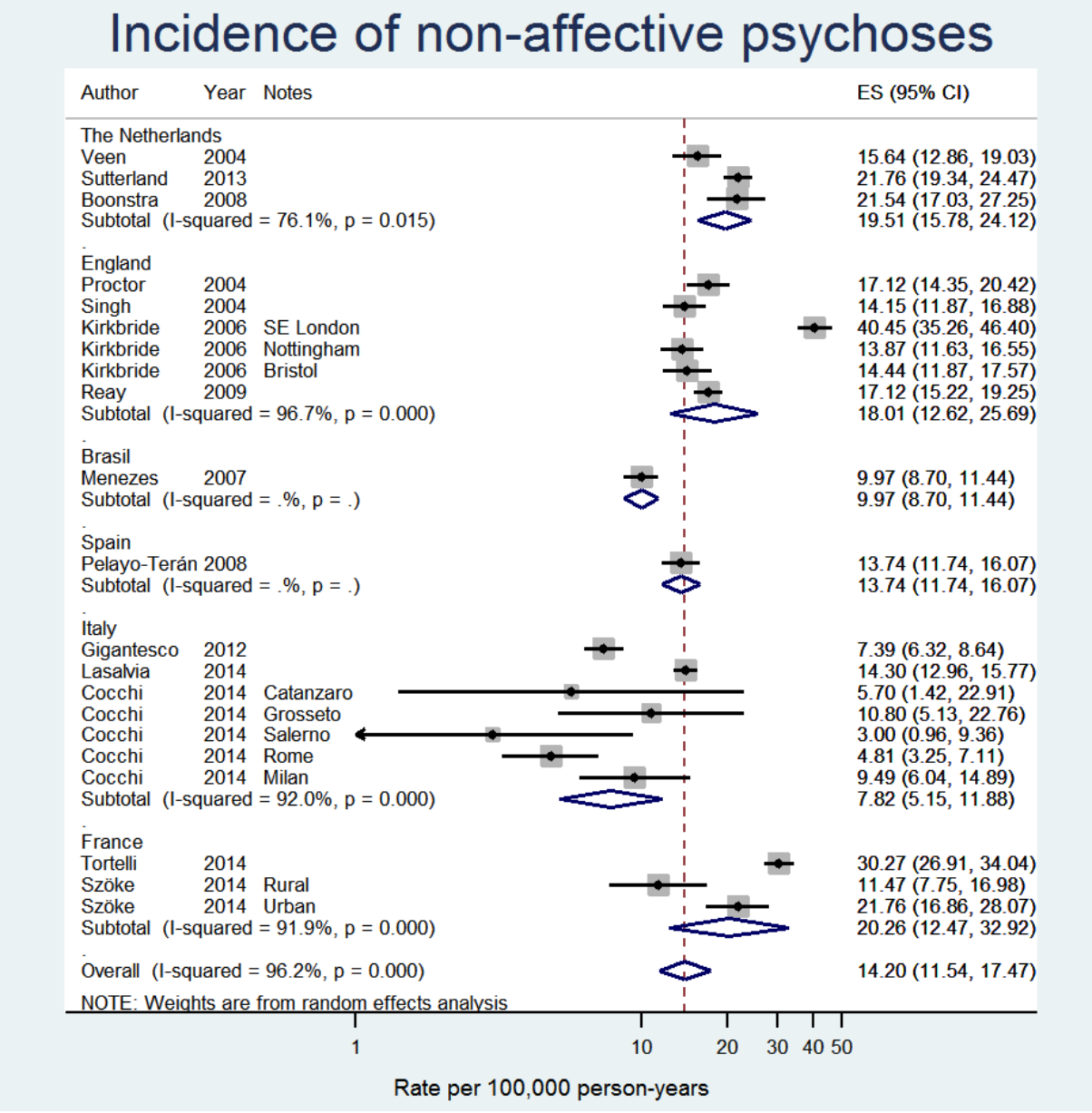


2.4.5.5 Non-affective disorders, overall rates

Fourteen citations reported rates of non-affective psychotic disorders across 21 settings (Boonstra, Wunderink, de Wit, et al., 2008; Cocchi et al., 2015; Gigantesco et al., 2012; Kirkbride et al., 2006; Lasalvia et al., 2014; Menezes & Scazufca, 2007; Pelayo-Terán et al., 2008; Proctor et al., 2004; Reay et al., 2010; Singh et al., 2004; Sutterland et al., 2013; Szöke et al., 2014; Tortelli et al., 2014; Veen et al., 2004). The pooled estimate of the incidence rates for non-affective psychoses was 14.2 (95%CI: 11.5-17.5) per 100,000 person-years, but heterogeneity was high (I^2: 96.2%) and variance between incidence rates more than thirteen-fold: from 3.0 (95%CI: 1.0-9.4) per 100,000 person-years in Salerno (Cocchi et al., 2015), to

40.5 (95% CI: 35.3-46.4) reported in Southeast London (Kirkbride et al., 2006). In general, rates were highest in France, and higher in England and the Netherlands than in Brazil, Spain and Italy (Figure 2.9). No evidence of small study effects was found using Egger's test (β: -3.49; p=0.15).

Figure 2.9: Forest plot of incidence of non-affective psychoses, grouped by country.



When assessing the variance in incidence rates in a univariable meta-regression, both freedom and latitude were associated with the incidence of non-affective psychoses (Table 2.14 below). However, in the multivariable model, none of the variables reached statistical significance. Insufficient studies reported data by sex, age or ethnicity to synthesise data by subgroup.

Table 2.13: Meta-regression: univariable results for non-affective psychoses (n=20)

Variable	Incidence Rate Ratio (95 %CI)	P-value
Mid-year	0.97 (0.92-1.00)	0.21
Quality	1.12 (0.99-1.25)	0.07
GINI-index	0.97 (0.90-1.05)	0.42
Freedom	**2.26 (1.24-4.12)**	**0.01**
Trust	1.01 (0.99-1.02)	0.41
Urbanicity[1]	1.01 (0.97-1.05)	0.47
Latitude	**1.07 (1.02-1.13)**	**0.01**
Type	1.15 (0.93-4.42)	0.20

IRRs in **bold** are significant (p<0.05)
[1]IRRs for urbanicity are reversed (1/IRR) to aid interpretability

2.4.6.6 Schizophrenia, overall rates

Eight citations reported the incidence of schizophrenia across ten settings (Coid et al., 2008; Drukker et al., 2006; Kirkbride et al., 2006; Lasalvia et al., 2014; Reay et al., 2010; Singh et al., 2004; Sutterland et al., 2013; Turola et al., 2012). Rates varied eight-fold: from 4.4 (95%CI: 3.4-5.7) (Reay et al., 2010) to 32.5 (95%CI: 28.9-36.5) per 100,000 person years (Coid et al., 2008), with a pooled estimate of 11.9 (95%CI: 8.3-17.1) per 100,000 person-years and high heterogeneity (I^2: 98.3%; Figure 2.10 below). No evidence of small study effects was found using Egger's test (β: -5.39; p=0.15). Heterogeneity was not explained by meta-regression (Table 2.14 below).

Figure 2.10: Forest plot of incidence of schizophrenia.

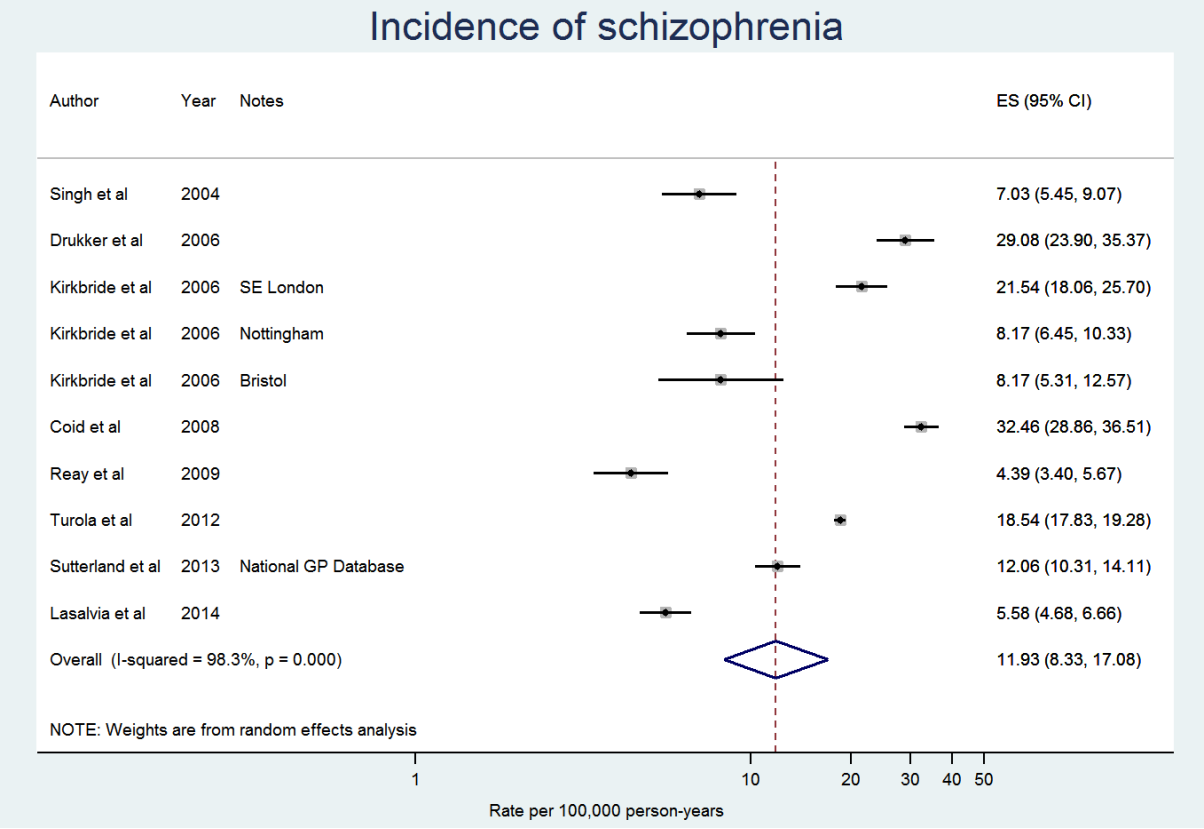


Table 2.14: Meta-regression: univariable results for schizophrenia (n= 10)

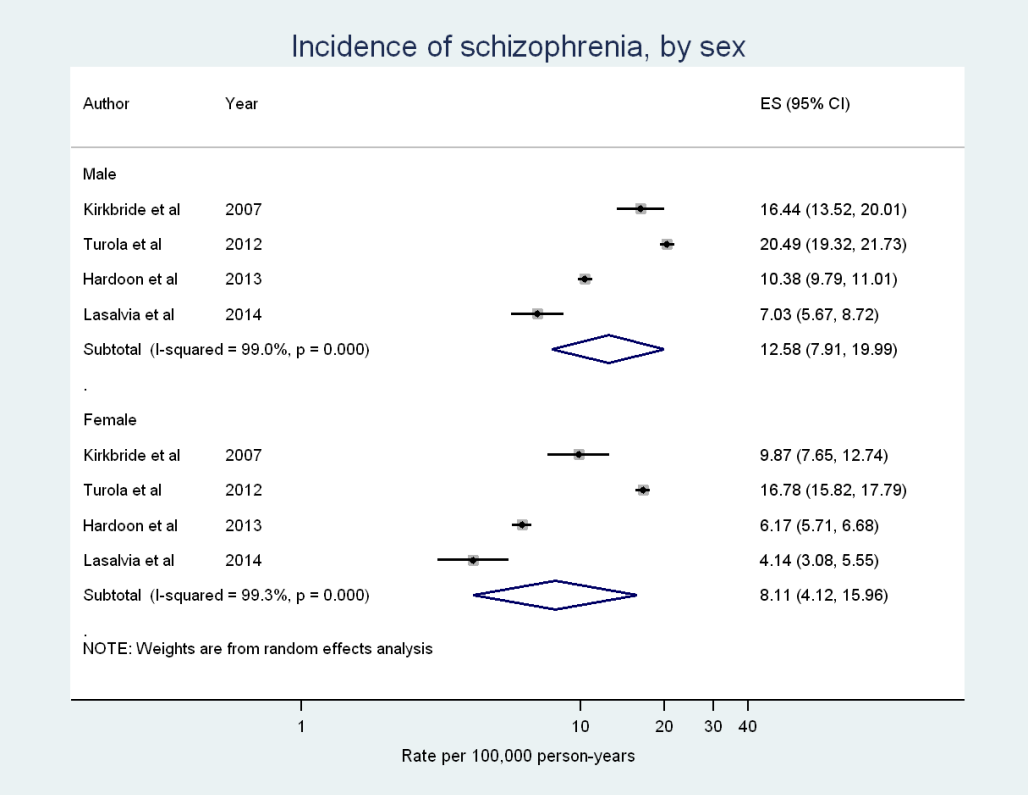
Variable	Incidence Rate Ratio (95% CI)	P-value
Mid-year	0.93 (0.81-1.07)	0.23
Quality	0.98 (0.70-1.38)	0.91
GINI-index	0.94 (0.81-1.09)	0.37
Freedom	0.97 (0.22-4.45)	0.97
Trust	1.02 (0.98-1.05)	0.34
Urbanicity[1]	1.06 (0.96-1.16)	0.18
Latitude	0.94 (0.77-1.15)	0.47
Type	1.01 (0.55-1.84)	0.98

[1]IRRs for urbanicity are reversed (1/IRR) to aid interpretability

2.4.6.7 Schizophrenia, rates by sex

The only sub-group for which sufficient data was available to synthesise results was sex, insufficient citations were available to do so for age and ethnicity. Four citations reported rates of schizophrenia for men and women separately (Hardoon et al., 2013; Kirkbride et al., 2007; Lasalvia et al., 2014; Turola et al., 2012) (Figure 2.11 below). Incidence rate ratios could not be computed due to unavailability of denominator data. The pooled incidence for men was higher than for women at 12.6 (95%CI: 7.9-20.0) and 8.1 (95%CI: 4.1-16.0) per 100,000 person-years respectively, as was the case in every citation included in this analysis.

Figure 2.11: Forest plot of incidence of schizophrenia, by sex.



2.4.6.8 Affective psychoses, overall rates

Nine citations reported the incidence of affective psychoses across twelve settings (Kirkbride et al., 2006; Lasalvia et al., 2014; Menezes & Scazufca, 2007; Proctor et al., 2004; Reay et al., 2010; Singh et al., 2004; Szöke et al., 2014; Tortelli et al., 2014; Veen et al., 2004). Incidence of affective psychoses was similar in Southeast London (Kirkbride et al., 2006) and outer Paris (Szöke et al., 2014) (14.0; 95% CI: 11.3-17.4 and 14.3; 95%CI: 10.5-19.6 per 100,000 person-years respectively) but lower in all other settings, including inner Paris (1.4; 95%CI: 0.8-2.4) (Tortelli et al., 2014), see Figure 2.12 below. The latter is a study of first admissions as an inpatient, whereas other studies are first contact studies, and will include both individuals who will become inpatients, and those who remain in the community. The pooled estimate of the incidence was 6.2 (95%CI: 4.6-8.3) per 100,000 person-years, but heterogeneity remained high at 93.9%. This heterogeneity was not explained by meta-regression (Table 2.15 below), and insufficient citations were available to pool rates by age, sex or ethnicity. No evidence of small study effects was found using Egger's test (β: -3.08; p=0.37).

Figure 2.12: Forest plot of incidence of affective psychoses.

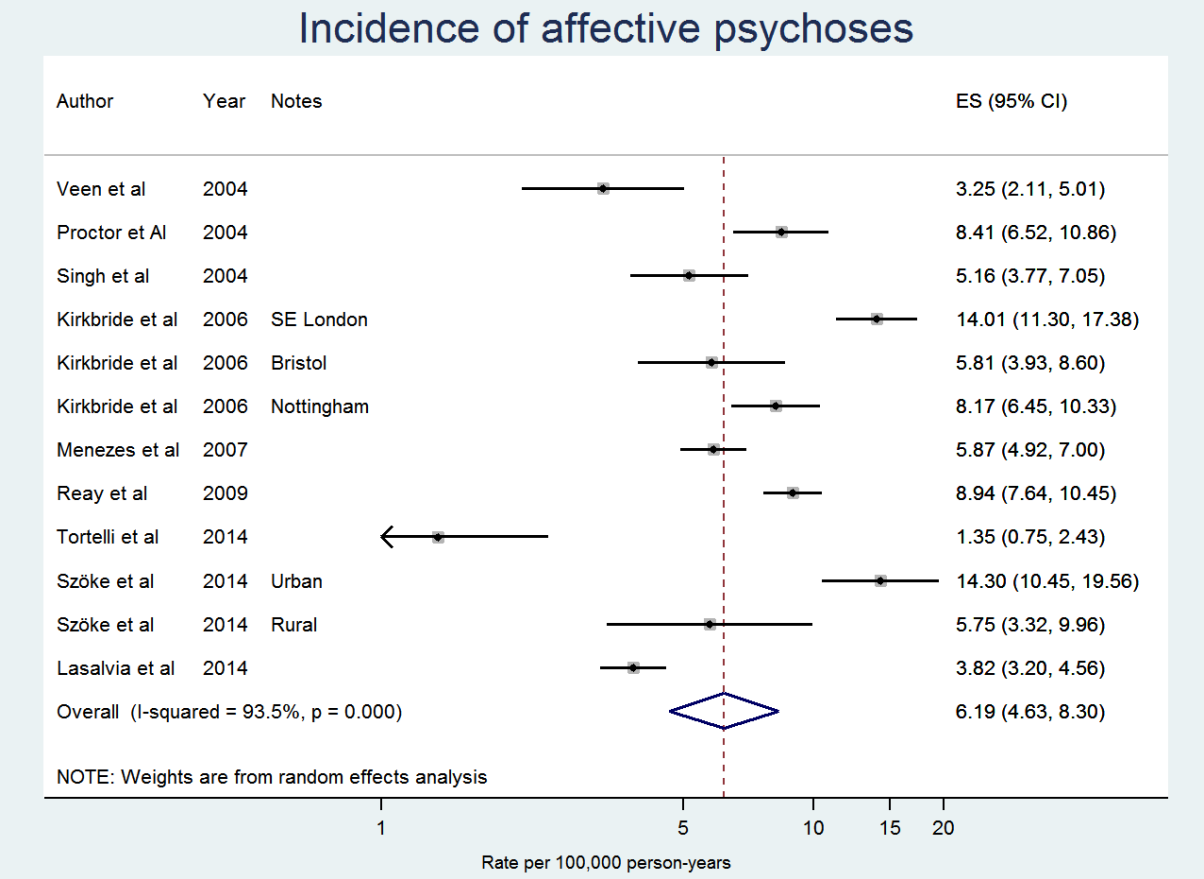


Table 2.15: Meta-regression: univariable results for affective psychosis (n=11)

Variable	Incidence Rate Ratio (95% CI)	P-value
Mid-year	0.99 (0.91-1.08)	0.84
Quality	1.09 (0.85-1.39)	0.45
GINI-index	1.09 (0.96-1.24)	0.18
Freedom	2.28 (0.62-8.34)	0.19
Trust	0.99 (0.95-1.03)	0.59
Urbanicity[1]	0.98 (0.91-1.05)	0.63
Latitude	1.06 (0.92-1.22)	0.41
Type	0.93 (0.63-1.35)	0.66

[1]IRRs for urbanicity are reversed (1/IRR) to aid interpretability

2.4.6.9 Bipolar disorder, overall rates

Six citations reported seven estimates of the incidence of bipolar disorder (Gigantesco et al., 2012; Hardoon et al., 2013; Kroon et al., 2013; Lasalvia et al., 2014; Reay et al., 2010; Selten et al., 2003; Veling, Hoek, et al., 2008), with one citation reporting rates of bipolar 1 and bipolar 2 separately (Kroon et al., 2013). The pooled incidence of bipolar disorder was the lowest of all pooled rates, at 3.9 (95%CI: 2.2-6.8) per 100,000 person-years (Figure 2.13 below). The highest incidence was reported at 13.6 (95%CI: 13.1-14.1) per 100,000 person-years in a citation covering an English national GP database (Hardoon et al., 2013). This subgroup further included a study based on a national surveillance system (Gigantesco et al., 2012) and a study based on a national psychiatric register (Selten et al., 2003). Both of these reported relatively high rates: 5.4 (95%CI: 4.5-6.6) and 9.8 (95%CI: 8.7-11.0) per 100,000 person-years. This diversity of methodology was also reflected in the extremely high heterogeneity (I^2: 99.0%) and was not explained by meta-regression: in the univariable analyses none of the variables reached statistical significance, and there were insufficient observations to carry out a multivariable analysis (Table 2.16). Insufficient citations were available to synthesise rates by age, sex or ethnicity. Some evidence of small study effects was suggested by the funnel plot (Figure 2.14 below), with small studies tending to report lower incidence rates. There were insufficient citations to conduct a formal Egger's test.

2.4.6.10 Other diagnostic outcomes

A number of citations reported a diagnostic outcome that could not be synthesised. Psychotic depression was reported by four citations (Fearon et al., 2006; Gigantesco et al., 2012; Reay et al., 2010; Veling, Susser, Van Os, et al., 2008), and ranged from 1.4 in The Hague (Veling, Susser, Van Os, et al., 2008) to 5.4 (95%CI: 4.2-6.7) per 100,000 person-years in Northumberland (Reay et al., 2010). A further two citations reported substance-induced psychosis, which had an incidence of 1.6 per 100,000 person years in Nottingham (Singh et al., 2004), and of 2.6 in Bologna (Tarricone et al., 2012).

Figure 2.13: Forest plot of the incidence of bipolar disorder.

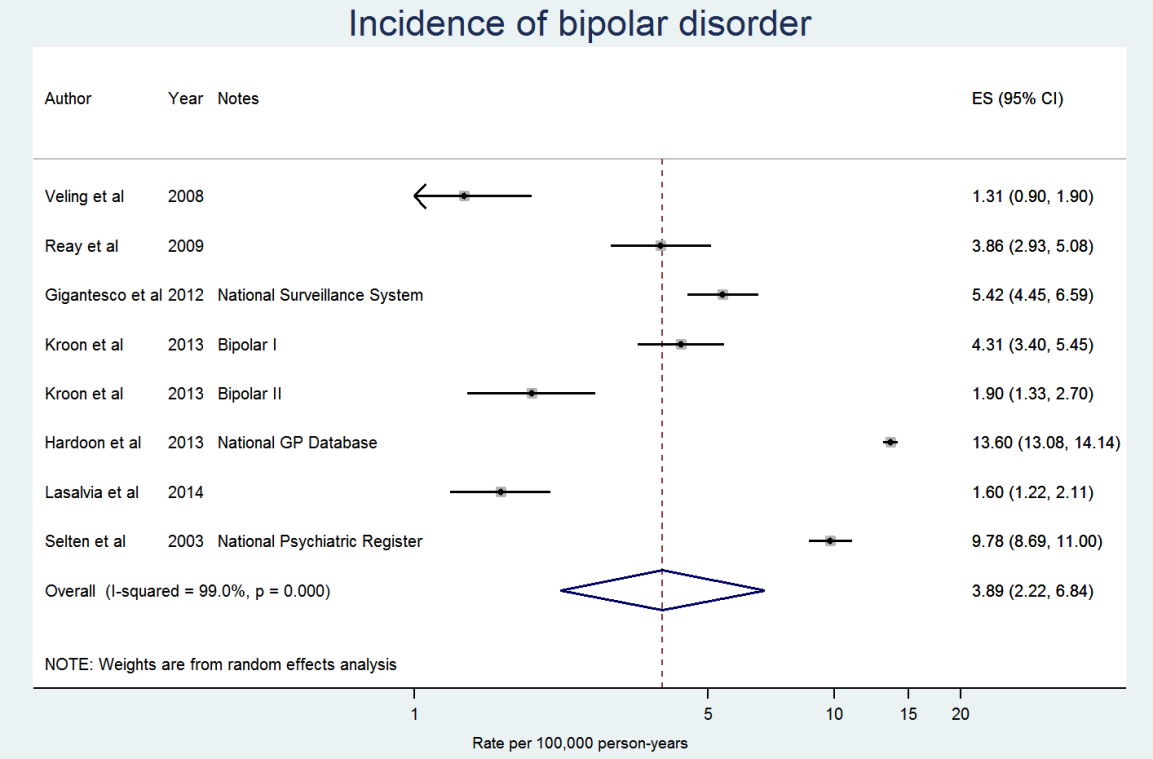


Table 2.16: Meta-regression: univariable results for bipolar disorder (n=9)

Variable	Incidence Rate Ratio (95% CI)	P-value
Mid-year	0.96 (0.72-1.28)	0.72
Quality	0.42 (0.14-1.26)	0.10
GINI-index	1.06 (0.87-1.31)	0.50
Freedom	2.39 (0.28-22.23)	0.38
Trust	0.99 (0.95-1.04)	0.65
Urbanicity[1]	0.81 (0.67-1.11)	0.12
Latitude	1.08 (0.34-6.47)	0.56
Type	1.30 (0.88-1.94)	0.16

[1]IRRs for urbanicity are reversed (1/IRR) to aid interpretability

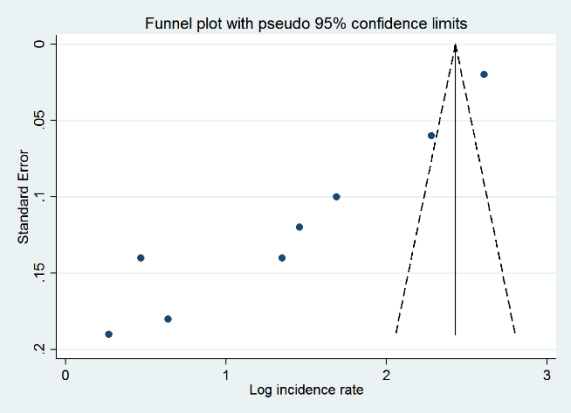


Figure 2.14: Funnel plot of small study effects, bipolar disorder

2.5 Discussion

2.5.1 Summary of main findings

Following a systematic literature of four databases, I identified 50 citations reporting the incidence of psychotic disorders published between 2002 and 2014 conducted in the six countries included in the EU-GEI study. Overall, the incidence of psychosis varied substantially between citations, but heterogeneity was poorly explained by the predictor variables included in the meta-regression (hypotheses 1 and 2, Table

2.17 below). There was no association between incidence of any outcome and study quality (hypothesis 3, Table 2.17). The incidence of all psychotic disorders (n=20) varied close to four-fold and variance was associated with urbanicity and latitude. Rates were typically higher for men than for women, and tended to decline across the life course. Ethnic minorities had a higher incidence compared with the majority population, although there was substantial variation between minority groups. Incidence of non-affective disorders (n=21) varied thirteen-fold, and this variation was not explained by a multivariable meta-regression model. The incidence of schizophrenia (n=10) varied eightfold. This variation was also not explained by meta-regression, although men had higher rates of schizophrenia than women. The incidence of both affective disorders broadly (n=12) and bipolar disorder (n=6) varied ten-fold and this could not be explained by meta-regression. Psychotic depression, substance-induced psychosis, psychotic symptoms and the 'other' category yielded too few citations to synthesise findings.

Table 2.17: Reappraisal of hypotheses

Hypothesis		Outcome
1.	Incidence varies between studies	Confirmed
2.	Incidence is higher in:	
	More urban areas	Partially confirmed
	Settings at a higher latitude	Partially confirmed
	More economically unequal countries	Rejected
	Countries with lower self-reported freedom	Rejected
	Countries with lower self-reported trust	Rejected
3.	No association between incidence and study quality	Confirmed

No evidence of small-study effects such as publication bias was found via formal Egger's tests for those outcomes with sufficient citations (all psychotic disorders, non-affective psychoses, schizophrenia or affective psychosis). Bipolar disorder was assessed by visual interpretation of the funnel plot. This was difficult due to sparsity of data points, but some evidence for small study effects appeared to be present with a lack of small studies with a high incidence rate.

2.5.2 Strengths and limitations

The methodological rigour of this systematic review and meta-analysis has been assessed on the basis of the AMSTAR criteria (see Section 2.3.12) (Table 2.18).

Table 2.18: AMSTAR Checklist

AMSTAR Item	Done?	Section
A priori design	Yes	2.3
Duplicate study selection and data extraction	No	2.3.4 / 2.3.5
Comprehensive literature search	Yes	2.3.3
Status of publication as inclusion criterion	Yes	2.3.3
List of studies included and excluded	Yes	2.4.1 / Appendix 2B
Characteristics included	Yes	2.4.1
Scientific quality assessment	Yes	2.3.8 / 2.4.1
Method of combining data items appropriate	Yes	2.3.9 / 2.4.6
Publication bias assessed	Yes	2.3.8
Conflict of interest included	Yes	Table 2.5

The only item on the AMSTAR checklist that this review does not meet is the duplicate data extraction. The majority of English citations (all those published until 31/12/2009) were included in Kirkbride et al's meta-analysis and as such data was extracted by this study team. HEJ independently extracted the data for the remaining English citations and for all citations from other settings. This was closely supervised by JBK and overseen by PBJ. HEJ and JBK had regular meetings to assess progress and this included cross-checks of the spreadsheet used for data-extraction.

There are however some further limitations arising from this work. The first is that the present review only covered six countries, whereas McGrath et al's global systematic review included citations from 31 countries (McGrath et al., 2004). As the countries covered in the present review accounted for almost half (42.3%) of citations included in McGrath's review, this is predominantly a limitation in terms of generalisability: the countries covered in this review are predominantly European and it is therefore difficult to extrapolate beyond this context. The countries covered in this review are also very similar in terms of latitude (with the exception of Brazil), and are on a relatively narrow range of income equality and self-reported freedom.

Some citations provided data from more than one catchment area or diagnostic outcome, and these centres were included in any meta-analyses as separate estimates of incidence (Cocchi et al., 2015; Kirkbride et al., 2006; Kroon et al., 2013). If either of these studies is biased, it might have skewed the pooled incidence rates in the respective analyses. As the emphasis of interpretation is on assessment of variance rather than pooled incidence, this isn't a major limitation. Heterogeneity can also be artificially lowered by treating these centres as different studies when they in fact used the exact same methodology, but this appears not to have played a major role (heterogeneity remained high).

Imperfect measurement threatens the validity of meta-regression, and may have contributed to the null-findings. Study quality, type of study, the median year of recruitment and latitude could be obtained accurately, but the measurement of urbanicity is less reliable. The number of inhabitants of a city or population density of an area is subject to change over time. This was researched in 2015, and changes in urbanicity between recruitment into the included studies and this systematic review are possible (although there is no direct evidence for this). Furthermore, data was obtained from the public domain, and not from official governmental sources (as the exact catchment areas were unknown), and the reliability of this data is unknown. Similarly, not all national-level variables (Gini index, self-reported levels of trust, self-reported levels of freedom) have been collected at the same time (see Section 2.4.2). All data on inequality was retrieved from the World Bank for all years since 2002. For the meta-regression the most recent year was used, but for Brazil an increase in inequality of more than ten percent was observed in this time period, and for all other countries in the analysis inequality had increased too. It is therefore possible that these indicators do not accurately reflect the reality at the time when the incidence data was collected.

In interpreting the results from the meta-regression, attention should also be paid to the ecological fallacy as data from a larger geographical area is applied to a smaller catchment area and it is not possible to say how typical these catchment areas are of the wider geographical area. This is most obvious with the data collected at the national level, but it also pertains to the urbanicity measure: this is a relatively crude indicator obtained for an entire city or area, and the catchment area might not be representative of this.

This review also benefits from a number of methodological strengths. The research question, search strategy, data extraction procedure, inclusion criteria and statistical methodology had all been specified and published before the systematic review was started. The search strategy was inclusive and based on a previously used strategy (Kirkbride et al., 2012). The literature search was carried out after consultation with a librarian, and was comprehensive. It incorporated four electronic databases (Embase, PsycINFO, PubMed and Web of Science), and two supplementary strategies: references of all included studies were checked manually, and all lead authors were approached to identify missing studies. Studies that met the inclusion criteria were included regardless of their publication status. Some unpublished and preliminary data were found from conference proceedings (Appendix 2B) but this could not be included in this Chapter as these citations covered the EU-GEI study, and one of the aims of this Chapter was to aid in the contextualisation of findings from this study (as presented in Chapter 4).

I presented details on included and excluded studies to aid transparency, and reporting in this review was transparent in general, in line with the AMSTAR criteria (Table 2.19). The method of combining data items was also appropriate, I assessed small study effects and included information on conflicts of interest.

2.5.3 Interpretation of findings

This systematic review presents a very varied epidemiological landscape: incidence rates appear higher in England, the Netherlands and France compared with Spain, Italy and Brazil. Any such comparison is hampered by the lack of studies in France, Spain and Brazil. Meta-regression was relatively unsuccessful at capturing the causes of this heterogeneity and this could be for a number of reasons. I might not have captured the setting-level variables of relevance to predicting psychosis incidence. A further contribution to this null-finding might be that a number of important known risk factors such as genetic risk (Schizophrenia Working Group of the Psychiatric Genomic Consortium, 2015), childhood trauma (Varese et al., 2012), familial risk (Esterberg et al., 2010) and cannabis use (Manrique-Garcia et al., 2012) could not be included as they aren't routinely available at a population-level. Measurement of environmental risk factors included in this study is also imperfect (see above).

This review includes studies covering a wide range of study quality, study methodologies pertaining to case finding, diagnostic classifications and even study types. The effects of study quality and study type on incidence rates are partially captured in the meta-regression, but only for those citations where an estimate of uncertainty was available. No evidence of a relationship between study type and incidence is

established in this review, but many of the case registers' incidence rates are at the high end of the spectrum. In a direct comparison between an incidence study and a case register, Hogerzeil and colleagues (2004) demonstrated an almost three-fold higher number of incident cases using the case register. Citations reporting data from case registers don't differ in quality from those who did not (Mann-Whitney U-test: -0.12, p=0.90), but case registers were only used in England and the Netherlands. It is therefore possible that the lower rates in remaining countries were partially explained by their reliance on first-contact or first-admission studies, which might underestimate true incidence. There is no evidence that lower rates in Spain and Italy could be explained by poorer quality of studies, as study quality doesn't appear to differ by country (Kruskal-Wallis χ^2: 8.5, 5df; p=0.11). However, considering the small number of studies in Spain and Brazil (n=1 each) in particular, this evidence is tentative.

A final potential contributor to the observed variance that can't be accounted for in the present methodology is differences in health care systems and help-seeking behaviour between countries. In England and the Netherlands, early intervention services for psychosis are widespread and well-established. They act as the centralised referral point for new cases, and may have contributed to improved case-detection in these countries. No such services exist in other countries so case-detection might be more difficult. Studies were conducted on service as usual and therefore accurately represented healthcare burden, but this might not reflect true incidence in the absence of additional efforts or training to identify all first episode psychosis (FEP) cases (such as a leakage study). It is also possible that in some of the countries included in this analysis, individuals are more willing, able, or forced to access the care system whereas in other countries more care might be provided by the family or direct social environment initially. This is however purely speculative.

2.5.4 Comparison with existing literature

Comparison with existing literature is tentative, due to differences in inclusion criteria and methodologies. The most recent international systematic review of the incidence of schizophrenia reported a median incidence of 15.2 per 100,000 person-years, and a 10^{th} to 90^{th} percentile range of 7.7 to 43.0 (McGrath et al., 2004). I reported a slightly lower pooled rate of 11.9 (95%CI: 8.3-17.1), and the overall range was also slightly lower than the 10^{th} to 90^{th} percentile reported (4.4 to 32.5 per 100,000 person-years). Our pooled rate of all psychotic disorders was more than twice as high at 28.3 per 100,000 person-years (95%CI: 23.4-34.3), indicating an important role for the other non-affective and for the affective disorders in determining overall FEP incidence. The only systematic review pertaining to mood disorders only synthesises incidence of major depressive disorder (Waraich et al., 2004), and as such isn't directly comparable to the present review. Similar to the present review, McGrath and colleagues (2004) found that the study features of quality, case finding method and diagnostic confirmation and criteria were not associated with incidence.

The increased rates of schizophrenia and all psychotic disorders in men were in line with previous literature. McGrath and colleagues (2004) estimated a median IRR of 1.4 (IQR: 1.1-1.8), and a 2003 meta-analysis of sex differences in schizophrenia demonstrated an pooled IRR of 1.4 (95%CI: 1.3-1.6) (Aleman, Kahn, & Selten, 2003). A collaborative recalculation of incident cases of schizophrenia did not only establish an increased risk in men (IRR: 1.2; 95%CI: 1.0-1.3) but also a different trajectory across the life-course: men were at higher risk of psychosis until age 39, and after age 50 this risk-profile was reversed (van der Werf et al., 2014). The increased risk of all psychosis in ethnic minorities is also in line with the literature. A systematic review and meta-analysis including 21 studies demonstrated that both first (IRR: 2.3; 95%CI: 2.0-2.7) and second-generation migrants (IRR: 2.1; 95%CI: 1.8-2.5) were at increased risk (Bourque et al., 2011). Only two studies included in the present review (Kirkbride et al., 2008; Veling et al., 2011) were also included in Bourque's review (albeit published in different citations), increasing confidence that the present study isn't purely a replication of this study and that findings are generalisable more broadly.

Support in the literature for the findings relating to the environmental risk factors is mixed. A strong role for urbanicity was anticipated on the basis of previous literature (McGrath et al., 2004; Vassos et al., 2012), but the present review only found a positive association between urbanicity and incidence for all psychotic disorders (in multivariable analyses). A previous meta-analysis had found a positive association between latitude and incidence of schizophrenia for men, but not for women or all persons (Saha et al., 2006), and the present review only found a positive association for all psychotic disorders in the multivariable model, after accounting for urbanicity. The included variables relating to the social environment (income inequality, self-reported trust and self-reported freedom) did not yield any significant results in the present review. There are a number of limitations that might contribute to this null-finding (see previous Section). However, some studies included in this review found that neighbourhood-level indicators of deprivation (Bhavsar et al., 2014; Kirkbride et al., 2007), or social fragmentation (Allardyce et al., 2005; Drukker et al., 2006; Kirkbride et al., 2007; Veling et al., 2015; Zammit et al., 2010) were associated with increased incidence. This has been found in other settings too (Anderson, Fuhrer, Abrahamowicz, & Malla, 2012; Burns & Esterhuizen, 2008) and suggests that the present review either used inappropriate indicators or country-level is not the appropriate level of analyses.

2.5.5 Conclusion

Considering the strengths and limitations of this review, the main conclusion is that there is considerable variation in the incidence of psychosis. Incidence varies by the classical epidemiological dimensions of place and person: heterogeneity between studies is high, and incidence appeared to be higher in France, England and the Netherlands, and in more urban and northern (for all psychotic disorders) settings, as well as in ethnic minority groups, young people and in men. It was difficult to determine what causes this variance, whether it is an artefact of different study methodologies of whether it reflects a real variance in

treated incidence across geographical locations. Meta-regression did not suggest a role for study methodology, and in Chapter 4 I account for this methodological heterogeneity by examining variance in incidence in a multinational study using comparable methodology across seventeen catchment areas in six countries. In the next Chapter (3), I will first introduce this study and its' methodology.

Chapter 3 - Methodology of the EU-GEI study

3.1 Introduction

Psychotic disorders affect between 0.7 and 3.5 per cent of the general population over the life course (see Section 1.2). There are a number of well-established environmental risk factors associated with the development of psychotic disorders, such as childhood trauma (Matheson et al., 2012; Varese et al., 2012), cannabis use (Arseneault et al., 2002; Manrique-Garcia et al., 2012; Moore et al., 2007), urban birth and upbringing (Vassos et al., 2012), and being of ethnic minority background (Bourque et al., 2011), operating on a background of familial risk (Esterberg et al., 2010) and genetic vulnerability (Schizophrenia Working Group of the Psychiatric Genomic Consortium, 2015). Until recently (Colodro-Conde et al., 2017; Hartz et al., 2017; Sariaslan et al., 2015), the environment risk factors and genetic vulnerability were often studied separately, and their interplay was rarely tested (Van Os et al., 2014). In order to remedy this, the EUropean network of national schizophrenia networks studying Gene-Environment Interactions study (EU-GEI) study was established. The study aimed to investigate the role of these environmental and genetic determinants, and their interactions, in the development, severity and outcomes of psychotic disorders.

The EU-GEI study is a multi-disciplinary, multi-national research collaboration that was funded for a period of five years, from 1 May 2010 until 30 April 2015. The work was divided into eleven work packages, including a prodromal work package, a genetic work package, an experimental work package and the work package that is used in the remainder of my PhD: functional enviromics. A full overview of work packages can be found in Table 3.1 below. Work packages 2-8 all consisted of different participating centres: a total of 26 institutions participated (European Network of National Schizophrenia Networks Studying Gene-Environment Interactions, 2009a). Most participating centres are located in Europe, though collaboration extended to Hong Kong, Melbourne and Ribeirão Preto (Brazil). A total of over 9,000 participants enrolled during the first four years of the study including patients, their siblings and general population controls, as well as number of twins through a dedicated twin study (European Network of National Schizophrenia Networks Studying Gene-Environment Interactions, 2009b). The remainder of this Chapter is dedicated to detailing the study methodology of work package 2 (WP2). This work package was designed to investigate the role of environmental risk factors, and elucidate their underlying factors, as well as construct an instrument to index 'environmental load'.

Table 3.1: Work packages of the EU-GEI study

Work package	Name	Objectives
1	Management	Guarantee implementation of EU-GEI as laid down in consortium agreement
2	Functional Enviromics	Validate reported environmental risks; elucidate factors underlying proxy environmental risks; construct an environmental load indexing instrument
3	Discovery Genetics	Identify genetic variation associated with schizophrenia diagnosis, intermediate phenotypes, severity and course; develop genetic hypothesis for GxE
4	Experimental GxE	Assess neurobiological and behavioural substrates mediating GxE
5	GxE Prodrome	Identify interactive environmental genetic and clinical causes of transition from being at-risk to clinical disorder; develop predictive translational tools
6	GxE Vulnerability and Severity	Identify interactive causes of cognition and MRI vulnerability and severity phenotypes
7	GxE Course	Identify interactive causes of course and outcome of schizophrenia; develop predictive translational tools
8	GxE Data and Statistics	Develop statistical tools for GxE; test for GxE in combined EU-GEI samples of WP2, WP5-7; develop Risk Assessment Chart methodology for use in clinical practice
9	Ethics	Develop an ethical context for translational GxE research in schizophrenia
10	Dissemination	Make sure results of EU-GEI have lasting impact beyond those directly involved
11	Training	Provide multidisciplinary and ethical training environment within EU-GEI

3.2 Study design and settings

WP2 consisted of both an incidence and a case-sibling-control study. The general methodology shared by both arms of the study will be described in this Chapter, and details pertaining specifically to the incidence arm and the case-sibling-control arm can be found in the methods sections of Chapters 4 and 6 respectively.

Seventeen centres across six countries participated in WP2 (Figure 3.1). These catchment areas were purposefully selected to represent the full urban-rural spectrum. In England, the study took place in Cambridgeshire and Peterborough, and in Southeast London. In the Netherlands, it took place in central Amsterdam, and in a rural area around the cities of Gouda and Voorhout. A total of six Spanish centres participated: Madrid, Barcelona, Valencia, Oviedo, Santiago and Cuenca. In France, participants were recruited across the 20th *arrondissement* of Paris (cases only), in a Parisian suburb (Val-de-Marne) and in a rural part of central France (Puy de Dôme). In Italy, participants were recruited from the city of Palermo (on Sicily) and the municipality of Bologna. Incidence data from an earlier study (2005-2007) in the Veneto region in Northern Italy was also included in the EU-GEI study, as the methodology was comparable, and full cases and controls in this catchment area were recruited in line with the remaining catchment areas. Finally, the Brazilian data was collected in Ribeirão Preto, which is located in the state of of São Paulo.

Figure 3.1: Map of EU-GEI settings.

3.3 Sampling and recruitment

Recruitment varied from 12 months (London) to 48 months (Val-de-Marne and Bologna) (Table 3.2) with a median of 25 months (IQR: 24-36). Case ascertainment predominantly took place between 2010 and 2013 with the exception of recruitment of incidence cases in Veneto. Recruitment of incidence cases involved trained researchers making regular contact with all secondary and tertiary mental healthcare providers to identify any potential cases. Research teams were overseen by a psychiatrist with experience in epidemiological research, and included trained research nurses and clinical psychologists. Teams received training in epidemiological principles and incidence study design to minimise non-differential

ascertainment bias across different local and national healthcare systems (for the incidence arm of the study). Training was made available online via WP11.

Table 3.2: Recruitment period and duration per setting

Setting	Recruitment start date	Recruitment end date	Recruitment duration in months
England			
Southeast London	01/05/2010	01/05/2011	12
Cambridgeshire	01/10/2010	30/09/2013	36
The Netherlands			
Amsterdam	01/10/2010	01/10/2013	36
Gouda & Voorhout	01/12/2010	01/12/2013	36
Spain			
Madrid	23/02/2011	31/12/2012	22
Barcelona	20/12/2010	31/12/2012	25
Valencia	22/12/2010	31/12/2012	24
Oviedo	13/12/2010	31/12/2012	25
Santiago	13/12/2010	31/12/2012	25
Cuenca	08/02/2011	31/12/2012	23
France			
Paris	01/06/2012	01/06/2014	24
Val-de-Marne	01/06/2010	01/06/2014	48
Puy de Dôme	01/09/2010	31/08/2012	24
Italy			
Bologna	01/01/2011	31/12/2014	48
Veneto	02/01/2005	31/12/2007	36
Palermo	02/10/2010	31/05/2014	44
Brazil			
Ribeirão Preto	01/04/2012	01/04/2015	36

Written informed consent was obtained from all participants prior to taking part in the full case-sibling-control study. Where subjects lacked the necessary literary skills, verbal consent was obtained with a witness present. For incidence-only cases who declined participation in the full study ethical approval was obtained to extract basic demographic and clinical details from patient records. Ethical approval was obtained in each catchment area individually.

3.3.1 Case recruitment

The aim of the incidence study was to identify all individuals aged 18 – 64 years who made contact with mental health services for an FEP in one of the clearly defined catchment areas for the duration of the study period for the relevant catchment area. All incidence cases presenting to mental health services were asked to participate in the full study via the equivalent of their care coordinator or consultant psychiatrist. The aim of the case-sibling-control study was to recruit 1,200 cases, 600 siblings and 1,200 population-based controls.

3.3.2 Sibling recruitment

Cases who had consented to participate in the full study were asked if they had any eligible siblings, and if they would be happy to approach them to take part in the study. Siblings were contacted by the case in the first instance. If cases were happy for their sibling to be included in the study, but unwilling to invite them personally, they were approached by a member of the research team. When cases were unwilling for their siblings to be contacted, no contact was made.

3.3.3 Control recruitment

The overall control sampling strategy was to recruit population-based controls using quota sampling to ensure representativeness in terms of age, sex and ethnicity. The exact methods were based on locally available sampling strategies and thus differed per setting. In Cambridgeshire and Peterborough, a stratified random sampling strategy was used, with GP practices sampled in the first stage and individuals within that practice in the second. Using sampling frames, a predefined number of GP practices were randomly selected, and approached for participation in the study. In all practices, GPs were asked to remove any individuals from their patient list who were deemed not suitable or appropriate to contact, due to for instance current severe illness or recent bereavement. A set number of eligible controls were then randomly selected from an anonymised patient list and contacted by the practice. During a staggered recruitment period, an initial invitation letter was followed up by one reminder for non-responders after two weeks, and a maximum of three phone calls.

A secondary strategy was designed to ensure accurate representation of hard-to-reach groups (young men and ethnic minority groups) through purposely over-sampling certain groups. In Southeast London, for instance, young Black men were oversampled.

3.3.4 Inclusion and exclusion criteria

Cases were included in the study if they:

- were aged 18-64;
- made contact with mental health services for a first episode of psychosis (even if longstanding);
- were resident in one of the clearly defined catchment areas, and
- made contact during the study period in the relevant catchment area.

Cases were excluded if:

- they had received treatment for an episode of psychosis prior to the study period;
- there was evidence their psychotic symptoms were precipitated by an organic cause, or
- they experienced transient psychotic symptoms resulting from acute intoxication (ICD10: F1X.5).

Siblings were included if they were:

- A full or half-sibling of a case;
- Aged 18 or over;
- There was no evidence of a current of past psychotic disorder;
- There was no evidence of current or past treatment with antipsychotic medication.

Controls were included if:

- They were aged 18 to 64;
- They were resident within one of the clearly defined catchment areas;
- There was no evidence of a current or past psychotic disorder;
- There was no evidence of current or past treatment with antipsychotic medication.

3.4 Diagnostic outcomes

For all probable FEP cases, research diagnoses were ascertained using the Operational Criteria Checklist (OPCRIT) algorithm. This is a 90-item questionnaire regarding psychopathology and background information, originally designed to facilitate a polydiagnostic approach in molecular genetics (Azevedo et al., 1999; Craddock et al., 1996). The EU-GEI study used 77 of these items to avoid duplication, as the remaining 13 items were collected elsewhere. OPCRIT assessment was based on a semi-structured clinical interview, or review of case notes and other relevant information. The clinical interview schedule used at each site followed local expertise, including the Schedules for Clinical Assessment in Neuropsychiatry [SCAN] (World Health Organization, 1992b) in England and Italy, the Comprehensive Assessment of Symptoms and History [CASH] (Andreasen, Flaum, & Arndt, 1992) in the Netherlands, the Structured Interview for DSM-IV [SCID] (First, Spitzer, Gibbon, & Williams, 1996) in Brazil, and the Diagnostic Interview for Genetic Studies [DIGS] (Nurnberger et al., 1994) in France. Where OPCRIT assessment was not possible, we relied on clinical diagnoses for the incidence study, and cases without an OPCRIT diagnosis were excluded from case-control analyses.

OPCRIT has been shown to have high inter-rater reliability generally (Azevedo et al., 1999; Craddock et al., 1996), and in our study following training (kappa=0.7). Using OPCRIT was preferred over relying on clinical diagnoses as it ensured that individuals with the same psychopathology were given the same diagnoses for the purposes of this study, despite potential local differences in clinical practice. OPCRIT was derived as soon as possible after first presentation.

In this thesis, I examined broad diagnostic categories as epidemiological outcomes. Currently, when individuals present to mental health services with an FEP in the UK, they often receive a broad diagnosis. This is partially to avoid stigma, but also because there is a degree of uncertainty associated with assigning a precise diagnosis, given that symptoms may develop over time following an initial presentation (Heslin et al., 2015). This appears to be particularly apparent in disorders other than schizophrenia (Bromet, Naz, Fochtmann, Carlson, & Tanenberg-Karant, 2005). To avoid spurious diagnostic accuracy, the main

outcome of interest in the remainder of this thesis is all clinically relevant psychotic disorders (ICD10: F20-33). This also accurately reflects how many people are diagnosed and treated for a psychotic disorder in mental health services, and is still relevant for the study of putative environmental risk factors. I have also analysed data by the non-affective/affective dichotomy as secondary outcomes to examine if they have a different aetiology. There is evidence to suggest that whilst non-affective disorders are spatially patterned, this isn't the case for affective disorders (March et al., 2008) and that whilst schizophrenia and bipolar disorder have some shared genetic risk (see Section 1.3), there are also neurodevelopmental differences between the two (Murray et al., 2004).

So, the analyses presented in Chapters 4 and 6 in this thesis were conducted for the following diagnostic outcomes:

- All psychotic disorders (F20-F33)
- All non-affective disorders (F20-F25)
- All affective disorders (F30-33)

Those diagnosed with psychosis NOS (F28/29) will be retained for analysis within the all psychotic disorders category, but not analysed separately due to small numbers. In Chapter 7, due to the limited sample size (n=443), I have only analysed data for the overall psychotic disorder category (F20-F33).

3.5 Data collection, entry and management

A full set of instruments administered to cases, controls and siblings can be found in Table 3.3 below. Individual level data-collection was done in clinical facilities in each of the catchment areas and, where possible, at participants' homes for non-invasive procedures such as questionnaires.

For the incidence arm of the study (Chapter 4), sociodemographic details (age, sex, ethnicity) were obtained from the MRC Sociodemographic Schedule. In the analysis of the case-sibling-control arm of the study (Chapters 6 and 7), I used data from the MRC Sociodemographic Schedule, the Childhood Trauma Questionnaire, and the cannabis questionnaire, as well as genetic data. The specific data-items will be explained in more detail in the methods section in the relevant Chapters. Because I made extensive use of the MRC Sociodemographic Schedule, the relevant sections are attached as an appendix (Appendix 3A).

All data was collected and housed securely at each of the participating centres, and was entered locally onto a secure, encrypted database system, based on commercial software but adapted specifically for EU-GEI purposes. Data was entered once with field codes restricted to logical values where possible, to minimise data entry errors. For this PhD, pseudo-anonymised data identifiable by EU-GEI and subject identifiers only was requested from the central database using a standardised form. Data was managed in Stata and stored on a secure drive which is backed up daily and only accessible to local EU-GEI researchers (PBJ, JBK, HEJ).

Table 3.3: Full list of instruments administered during the EU-GEI study

Instrument	Purpose	Administered to
Clinical instruments		
Nottingham Onset Schedule (NOS)	Assess duration of untreated psychosis	Cases
OPCRIT	Obtain research diagnosis	Cases
Schedule Deficit Syndrome	Assess if cases with schizophrenia also have deficit syndrome	Cases
Community Assessment of Psychic Experience (CAPE)	Asses psychopathology in general population / exclude controls with psychoses	Controls, sibling
Structured Interview for Schizotypy-Revised (SIS-R)	Assess schizotypy in general population	Controls, siblings
Venepuncture	Extract DNA	Cases, controls, siblings
Family Interview for Genetic Studies (FIGS)	Gather family history of psychosis	Cases, controls, siblings
Premorbid Adjustment Scale (shortened)		Cases, controls, siblings
Past and present medication		Cases, controls, siblings
Psychological instruments		
Global Assessment of Function (GAF)	Assess symptoms /impairment of function	Cases, controls, siblings
Shortened Wechsler Adult Intelligence Scale (WAIS)	Assess neuropsychology	Cases, controls, siblings
Brief Core Schema Scale	Assess attributional bias	Cases, controls, siblings
Beads task	Assess probabilistic reasoning bias	Cases, controls, siblings
Degraded facial recognition task		Cases, controls, siblings
White noise task	Assess attributional bias to random events	Cases, controls, siblings
Benton Facial Recognition Task		Cases, controls, siblings
Sociodemographic instruments		
Schedules for the Assessment of Social Contexts and Experiences (SASCE). This included: • MRC Sociodemographic Schedule • Childhood Experiences of Care and Abuse (CECA); • Amended Bullying Questionnaire; • Interview for Recent Life Events; • Discrimination scale; • Harvard Trauma Questionnaire	Gathering sociodemographic information (date of birth, sex, ethnicity, individual and parental place of birth, individual and parental social class, migration history, housing and living circumstances, current and past addresses, employment history, relationships, social networks) and individual-level environmental risk factors across the life course	Cases, controls, siblings
Childhood Trauma Questionnaire	Assess childhood trauma	Cases, controls, siblings
Cannabis questionnaire	Assess usage of cannabis, including other drugs and alcohol	Cases, controls, siblings
Social Environment Assessment Tool	Rating of the environment in terms of neighbourhood trust and cooperation	Cases, controls, siblings
Other instruments		
Stigma scale	Assess experiences of stigma between service users, siblings and controls (user-led)	Cases, controls, siblings
Researcher Checklist	Used at initial telephone interview to check against inclusion criteria	Cases, controls, siblings

Missing data was treated differently for the incidence and case-sibling-control arm of the study, and as such is discussed in more detail in Chapters 4 and 6 respectively.

3.6 Local variations in the protocol

Methodology was designed to be identical across settings, however practical variations in healthcare provision meant that health service contact points varied between countries, so local adaptation was necessary. Deviations from the protocol are noted in Table 3.4 below. In Ribeirão Preto, Paris and Val-de-

Marne only, a leakage study was conducted, which led to the identification of additional incidence cases (details in Chapter 4). The Verona data was derived from a previous study which used comparable methodology (Lasalvia et al., 2014), but had a lower upper-age limit of 54. The limitations resulting from this mainly pertain to the incidence study, and as such are discussed in Chapter 4. Due to the exact nature of the ethics approval in London, incidence data could only be analysed while physically at the South London and Maudsley NHS Trust. This was a logistic issue which didn't affect the quality of the data. In Gouda & Voorhout, no ethical approval was granted for researchers to retrospectively obtain additional clinical information from case notes (for incidence-only cases only).

Table 3.4: Local variations in protocol

Catchment area	Variation	Threat to validity	What I did about it
Ribeirão Preto	Leakage study conducted (Incidence study only)	Bias	Caution in interpretation of results: rate might be higher than in other centres.
Gouda & Voorhout	No ethical approval to retrieve case notes for OPCRIT (Incidence study only)	Potential for bias	Clinical diagnoses were obtained.
Verona	Used data from an earlier study, with a lower upper age limit (54).	Residual confounding	Caution in interpreting results.
Paris & Val-de-Marne	Impossible to identify second generation migrants in denominator population (Incidence study only)	Residual confounding	Exercise caution when interpreting results from these settings: likely to be an underestimate of the rate in minority populations.
	Leakage study conducted (Incidence study only)	Bias	Caution in interpretation of results: rate might be higher than in other centres.
Puy de Dôme	Ethnicity only available for full cases (Incidence study only)	Bias	Excluded from age-sex-ethnicity standardisation.
London	Data had to be obtained separately (incidence study only)	None	Obtained data separately and carried out analyses in London

3.7 Recruitment per catchment area

As can be seen in Table 3.5 below, a total of 2,774 incidence cases was identified during the study period across our catchment areas, of which 41.43% (1,148 cases) agreed to participate in the full study. A total of 1,499 controls were recruited, as well as 272 siblings. This meant that the recruitment target for controls was exceeded, was met by 96% for cases and only 45% of the targeted number of siblings was recruited. Almost all catchment areas recruited cases, siblings and controls with the exception of Paris where only cases were recruited (and one sibling), and London where only cases and controls were recruited.

Due to the limited number of siblings, particularly when allowing for the clustering of participants at the catchment area level, I decided early on to restrict the analyses of the case-sibling-control arm of the study to cases and controls only (for the purposes of this thesis). It proved to be very difficult to recruit siblings, as cases were often unwilling to disclose to their siblings that they had experienced a psychotic episode. This level of self-stigma was not foreseen in the study protocol, and excluding siblings is an unfortunate limitation of the analyses of the case-control study.

Table 3.5: Recruitment of incidence cases, full cases, siblings and controls per catchment area

Catchment area	Incidence cases	Full cases	Siblings	Controls
England				
Southeast London	262	202	0	230
Cambridgeshire	266	45	5	108
The Netherlands				
Amsterdam	292	96	25	101
Gouda & Voorhout	167	102	51	110
Spain				
Madrid	89	43	21	38
Barcelona	108	31	13	37
Valencia	58	48	15	32
Oviedo	82	44	11	37
Santiago	36	28	9	38
Cuenca	27	18	4	38
France				
Paris	120	36	1	0
Val-de-Marne	212	55	4	100
Puy-de-Dome	42	15	1	47
Italy				
Bologna	165	70	4	65
Veneto	104	59	6	115
Palermo	179	58	3	100
Brazil				
Ribeiãro Preto	565	198	99	303
Total	**2,774**	**1,148**	**272**	**1,499**

3.8 Representativeness of the sample

I assessed if cases who agreed to participate in the case-control study were representative of the incidence sample and, using the population at-risk identified in the incidence study, if controls were representative of the population at-risk in terms of age, sex and ethnic minority status. The population at-risk was ascertained using official government statistics in each country and was categorised by sex (binary), minority status (binary) and age group (18-24, then five-year bands). Full details on the population at-risk can be found in Section 4.3. For the purposes of ascertaining representativeness by age group, the following age categories were used: 18-24, 25-24, 35-44, 45-54 and 55-64. Representativeness was ascertained using Chi^2 tests or, when the sample size was too small, Fisher's exact test.

3.8.1 Full case sample

Cases aged 18-24 were over-represented in the case-control sample (n=410, 35.9%) compared with the incidence sample (n=808, 29.1%), and those aged 45-54 (n=106, 9.3%) and 55 or over (n=38, 3.3%) were under-represented (compared with 13.8% and 5.5%; χ^2: 35.5, p<0.01, Table 3.6 below). This reflected heterogeneity by setting; in London (χ^2: 31.4, p<0.01), Amsterdam (χ^2: 24.1, p<0.01) and Ribeirao Preto (χ^2: 13.0, p=0.02) this broad pattern held, whereas in other centres full cases were representative of the incidence sample in terms of age (Table 3.7 below).

Table 3.6: Representativeness of the full case sample compared with the incidence sample

	Incidence sample		Full case sample			
	n	Percentage	*n*	Percentage	χ^2	p-value
Age						
18-24	808	29.2	410	359	35.51	<0.01
25-34	868	31.4	382	33.4		
35-44	558	20.2	207	18.1		
45-54	382	13.8	106	9.3		
55-64	152	5.5	38	3.3		
Sex						
Male	1,578	57.0	707	61.6	7.12	<0.01
Female	1,192	43.0	441	38.4		
Minority status*						
Majority	1,639	60.1	645	57.0	3.13	0.08
Majority	1,088	39.9	486	43.0		

* This does not include Puy-de-Dome, as ethnicity data was not available for incidence-only cases

Overall, male incidence cases were more likely to participate in the case-control study (61.6% compared to 57.0% of the incidence sample; χ^2: 7.1, <0.01, Table 3.6 above) than female cases, despite the fact that in all centres apart from London the sex-distribution appeared to reflect the incidence sample (Table 3.7 below), indicating low power at catchment-area level to detect small differences. Incidence cases from the ethnic majority were not more or less likely to participate in the full study than their ethnic minority counterparts (χ^2: 3.1, p=0.08). In Paris and Val-de-Marne however, incidence cases of ethnic minority background were over-represented in the full sample (χ^2: 9.9, p<0.01 and χ^2: 2.2, p<0.01 respectively; see Table 3.7 below).

Table 3.7: Representativeness of the full case sample compared with the incidence sample, by catchment area

	Age		Sex		Minority status	
	χ^2	p-value	χ^2	p-value	χ^2	p-value
England						
Southeast London	**31.4**	**<0.01**	**4.3**	**0.04**	0.3	0.57
Cambridgeshire	5.4	0.25	0.0	0.88	0.1	0.72
The Netherlands						
Amsterdam	**24.0**	**<0.01**	23.0	0.08	0.0	0.96
Gouda & Voorhout	0.4	0.99	0.1	0.53	1.3	0.26
Spain						
Madrid	0.9	0.92	0.0	1.00	0.1	0.75
Barcelona	1.2	0.89	2.9	0.09	0.0	0.93
Valencia	0.3	0.99	0.6	0.45	0.0	0.94
Oviedo	0.3	0.97	0.0	0.90	[1]	0.64
Santiago	1.4	0.84	0.5	0.48	[1]	1.00
Cuenca	0.3	0.96	0.0	1.00	0.08	0.78
France						
Paris	1.7	0.79	0.1	0.78	**9.9**	**<0.01**
Val-de-Marne	3.7	0.45	2.0	0.16	**22.6**	**<0.01**
Puy-de-Dome	1.7	0.32	0.2	0.64	n/a	n/a
Italy						
Bologna	4.2	0.38	0.1	0.77	[1]	0.76
Veneto	0.9	0.83	0.1	0.80	0.0	0.85
Palermo	5.7	0.23	0.1	0.71	1.1	0.30
Brazil						
Ribeirão Preto	**12.2**	**0.02**	0.6	0.43	2.1	0.15

Estimates in **bold** are significant (p<0.05)
[1]Fisher's exact test was performed, no estimate given.

3.8.2 Control sample

Overall, controls were not strictly representative of the population-at-risk in terms of age, sex, and ethnic minority status. Those aged 18-34 were over-sampled, whereas those aged 35 and over were under-sampled (χ^2: 212.4, p<0.01, Table 3.8 below). When examining this by setting, only controls in Cambridgeshire and Peterborough (χ^2: 0.37, p=0.99), Gouda and Voorhout (χ^2: 5.79, p=0.22) Valencia (χ^2: 5.40, p=0.27) and Oviedo (χ^2: 5.20, p=0.27) were representative of the population at-risk in terms of age (Table 3.9 below).

Table 3.8: Representativeness of the control sample compared with the population-at-risk.*

	Population at-risk		Controls			
	n	Percentage	*N*	Percentage	χ^2	p-value
Age						
18-24	1,828,075	14.1	323	21.7	212.42	<0.01
25-34	3,057,640	23.6	511	34.3		
35-44	3,058,837	23.7	323	15.6		
45-54	2,856,614	21.9	253	17.0		
55-64	2,152,499	16.6	172	11.5		
Sex						
Male	6,337,783	49.5	672	46.0	7.11	<0.01
Female	6,464,653	50.5	788	54.0		
Minority status						
Majority	9,881,660	77.2	1,072	72.1	21.65	<0.01
Majority	2,917,823	22.8	414	27.9		

* This does not include Paris, as no controls were recruited here

The control samples for each setting appeared representative in terms of sex, with only Bologna (χ^2: 4.32, p=0.04) and Ribeirão Preto (χ^2: 6.31, p<0.01) over-sampling women (Table 3.9 below). Nonetheless, across the whole sample, women were disproportionately more likely to participate as controls compared with men: 54.0% (n=788) of the control sample was women, compared to 50.5% (n=6,464,653) of the population at risk (χ^2: 7.1, p<0.01, Table 3.8 above).

Similarly, even though overall controls from ethnic minorities appeared to be over-represented in the sample (27.9%, n=414 compared to 22.8%, n=2,917,823 of the population at-risk, Table 3.8 above), this was not typically the case for individual centres. Only in Val-de-Marne ethnic minorities were over-sampled (χ^2: 11.6, p<0.01), and in Gouda and Voorhout (χ^2: 7.7, p<0.01) and Bologna (Fisher's exact p=0.04) they were under-sampled (Table 3.9 below).

Table 3.9: Representativeness of the control sample compared with the population-at-risk, by catchment area.

	Age		Sex		Minority status	
	χ^2	p-value	χ^2	p-value	χ^2	p-value
England						
Southeast London	**33.3**	**<0.01**	0.1	0.81	2.3	0.13
Cambridgeshire	0.4	0.99	0.4	0.52	0.0	1.00
The Netherlands						
Amsterdam	**21.6**	**<0.01**	0.6	0.43	3.4	0.07
Gouda & Voorhout	5.8	0.22	0.3	0.60	**7.7**	**<0.01**
Spain						
Madrid	**22.2**	**<0.01**	0.2	0.66	0.9	0.33
Barcelona	**21.6**	**<0.01**	0.4	0.54	0.1	0.75
Valencia	5.4	0.27	0.1	0.76	0.0	0.87
Oviedo	**13.6**	**0.01**	0.0	0.95	[1]	0.18
Santiago	**21.0**	**<0.01**	0.1	0.76	[1]	1.00
Cuenca	5.2	0.27	0.0	1.00	0.0	0.90
France						
Paris						
Val-de-Marne	**13.2**	**<0.01**	0.0	0.10	**11.6**	**<0.01**
Puy-de-Dome	**12.6**	**0.01**	0.6	0.46	[1]	0.34
Italy						
Bologna	**57.9**	**<0.01**	**4.3**	**0.04**	[1]	**0.04**
Veneto	**21.4**	**<0.01**	1.1	0.30	1.8	0.18
Palermo	**47.1**	**<0.01**	0.0	0.99	0.9	0.35
Brazil						
Ribeirão Preto	**117.3**	**<0.01**	**6.3**	**0.01**	0.2	0.63

Estimates in **bold** are significant (p<0.05)

[1] Fisher's exact test was performed, no estimate given.

In the near future, sampling weights will become available, to ensure that the control sample is fully representative of the population at-risk. They were not yet available for analyses presented in this thesis. However, sampling weights are mainly useful for as survey weights for prevalence estimates (Porta, 2014) and less so for case-control comparisons. As I am not trying to infer any population values, but merely approximate the odds of a risk factor in the case sample vis-à-vis the control sample, I believe sampling weights are unlikely to change the conclusions presented from any analyses in this thesis.

Chapter 4 - Substantial variation in the treated incidence of psychotic disorders: findings from the multinational EU-GEI study.

This Chapter has resulted in the following peer-reviewed publication:

Jongsma, H.E., Gayer-Anderson, C., Lasalvia, A., et al (2017). Treated Incidence of Psychotic Disorders in the Multinational EU-EI Study. *JAMA Psychiatry.* DOI:10.1001/jamapsychiatry.2017.3554

4.1 Background

Psychotic disorders are characterised by abnormal perception, beliefs and thought processes. In the systematic review in Chapter 2, I estimated a pooled incidence rate of all psychotic disorders of 28.3 (95%CI: 23.4-34.3) per 100,000 person-years, although variance was high. Psychotic disorders contribute substantially to global burden of disease (Whiteford et al., 2015) because of their frequent chronicity over the life course (Owen et al., 2016), and the poorer physical (Daumit et al., 2006; Leucht et al., 2013), social (Hare, 1956; Marwaha & Johnson, 2004) and lifestyle (Vancampfort et al., 2012) outcomes faced by many people with psychosis, culminating in an average reduced life expectancy of 14.5 years (Hjorthøj et al., 2017).

Until a decade ago, the prevailing view was that psychotic disorders, particularly schizophrenia, were distributed homogeneously worldwide. This belief arose from (mis-)interpretation of the most recent international study of the incidence of psychotic disorders, conducted in the early 1980s and published as the landmark WHO "10-country" study (Jablensky et al., 1992). Although rates of schizophrenia varied from 7 to 14 per 100,000 person-years across eight international sites, this effect was narrowly outside statistical significance, despite statistically-robust evidence of a 2.5-fold variation in a more broadly-defined outcome of non-affective psychoses ($p<0.01$).

In the decades which followed, important empirical research (see Bourque et al., 2011, Kirkbride et al., 2012, McGrath et al., 2004) has demonstrated that psychotic disorders vary across replicable social and environmental gradients, including raised rates amongst men, younger adults (Thorup, Waltoft, Pedersen, Mortensen, & Nordentoft, 2007), ethnic minority groups (Bourque et al., 2011), and with urban birth and upbringing (Vassos et al., 2012), potentially in tandem with genetic susceptibility (Schizophrenia Working Group of the Psychiatric Genomic Consortium, 2015). This built on earlier epidemiological studies, from the USA (Dohrenwend & Dohrenwend, 1969; Faris & Dunham, 1939; Hollingshead & Redlich, 1958) and Europe (Giggs & Cooper, 1987; Hare, 1956; Maylaih et al., 1989), which revealed strong associations between neighbourhood social deprivation and greater rates of psychosis. Nevertheless, no international comparison of incidence of psychotic disorders using a consistent methodology has been carried out since the WHO study. To address this, I will estimate the incidence of psychotic disorders across 17 catchment

areas in six countries using a comparable methodology, as part of EU-GEI study. Specifically, I will test whether any observed variance in incidence could be attributed to putative social and environmental factors, including individual age, sex and ethnicity, and setting-level latitude (McGrath, Saha, Chant, & Welham, 2008), population density (Vassos et al., 2012), unemployment, proportion of single-person households and owner-occupied homes as markers of social disadvantage.

4.2 Hypotheses

In this Chapter, I tested the following hypotheses:

1. There would be substantial variation in incidence of all psychoses, non-affective psychoses and affective psychoses across the catchment areas included in the EU-GEI study.
2. This variance would not be accounted for by standardisation for age, sex, and minority status.
3. Latitude, population density, unemployment, proportion of single-person households and owner-occupied homes would partially account for this variance.
4. A higher latitude, population density, unemployment and percentage of single households would be associated with increased incidence, as will be a lower percentage of owner-occupied homes.

4.3 Methods

The study design, sampling and recruitment strategy, case ascertainment and data collection and management of the EU-GEI study were detailed in Chapter 3. This Section only contains additional information relevant to this Chapter.

4.3.1 Population at risk

I estimated the population at-risk, aged 18-64 years, in each catchment area from the most accurate local or national routine demographic data available (Table 4.1), stratified by age group (18-24, then 5-year bands), sex and minority status (see below). I multiplied the population by case ascertainment duration (in years, see Section 3.3) to estimate person-years at-risk.

Table 4.1: Denominator and majority status characteristics, by catchment area

Country	Denominator source (year)	Denominator type	Ethnic majority	Ethnic minorities
England	Office for National Statistics (2011)	Census	White British	Any other ethnicity
The Netherlands	Statistics Netherlands (2014)	Yearly estimates	Individual and both parents born in the Netherlands	Any other ethnicity
Spain	Instituto Nacional de Estadística (2012)	Yearly estimates	Born in Spain	Born abroad
France	Institute Nationale de la statisique et des études économiques (2011)	Yearly estimates	Born in France and territories*	Born abroad
Italy	L'Instituto nazionale di statistica (yearly)	Yearly estimates	Born in Italy	Born abroad
Brazil	Sistema IBGE de Recuperação Automática (2010)	Yearly estimates	White	Any other skin colour

* French overseas territories are: Guadeloupe, French Guiana, Martinique, Réunion, Mayotte, French Polynesia, Saint Pierre and Miquelon, Wallis and Futuna, Saint Martin, Saint Bartélemy, New Caledonia and French Southern Antarctica.

4.3.2 Measures

The primary outcome was an OPCRIT-confirmed ICD-10 diagnosis of any clinically-relevant psychotic disorder (ICD-10: F20-33). This broad phenotype was considered alongside two secondary outcomes: non-affective psychoses (F20-25) and affective psychoses (F30-33).

Data on age group (as above), sex and minority status were collected at baseline on all participants using the MRC Sociodemographic Schedule (Mallett, 1997) and case notes. Here, I defined a binary variable to distinguish between the ethnic majority population in each catchment area, and all other minority groups. In each country, the majority population was classified as the majority (White) ethnic group, following national conventions (Table 4.1). In the UK, the Netherlands and Brazil, an ethnicity-based distinction was made between the White British/Dutch/Brazilian groups and all minority ethnic groups. In Spain and Italy, Spanish-born and Italian-born groups were defined as the majority population, respectively, with all foreign-born groups classified as the minority group. Practically, both definitions lead to the identification of a White majority group, since large proportions of adult-aged second- and later-generation groups do not yet exist in Italy or Spain, given substantial immigration is a recent phenomenon (see Section 1.4). France recognises all people born in France or its territories as 'French-born', with no further provision for ethnicity, and I followed that definition here.

I estimated environmental risk factors in each catchment area. Latitude was estimated in degrees from the equator (Encyclopaedia Brittanica, 2017). Population density was derived as the number of inhabitants per square kilometre, based on official total population estimates. I also derived three broad measures of the social environment (unemployment, owner-occupied housing, single-person households) from the 2011 European Household and Population Census (European Commission, 2013), a decennial census which provides comparable data at a provincial level (Nomenclature of Territorial Units for Statistic [NUTS]-2 regions). Equivalent data for Ribeirão Preto were derived from the 2010 National Census of Brazil (Instituto Brasileiro de Geografia e Estatística, 2017). Duration of untreated psychosis (DUP; in weeks) was estimated for descriptive purposes, assessed via the Nottingham Onset Schedule (Singh et al., 2005), and based on time from onset of symptoms to first contact with secondary mental health services for suspected psychosis.

4.3.3 Missing data

Seven cases (0.3%) were missing data on age or sex, and were excluded from direct standardisation and statistical modelling, but retained for crude incidence rate estimation. Except for Puy-de-Dôme (where the majority of cases was missing data on ethnicity, see Section 3.6), I coded any participants missing minority

status data (N=5; 0.2%) to the majority group. Data from Puy-de-Dôme was retained for estimation of crude incidence rates, but excluded from standardisation and statistical modelling.

4.3.4 Statistical analyses

For each outcome, I estimated crude incidence rates per 100,000 person-years and 95% confidence intervals, by sociodemographic characteristics and by catchment area. Next, I used direct standardisation for (i) age-band and sex, and (ii) age-band, sex, and minority status to investigate variation in rates between catchment areas (Kirkwood & Sterne, 2003). I used the total population of England & Wales from the 2011 Census (Office for National Statistics, 2013b) as the standard population, and estimated standardised incidence ratios (SIR) using the overall sample incidence rate as the reference category. Finally, I used random effects (intercepts) Poisson regression to investigate variance in incidence by sociodemographic and environmental factors. Random effects (intercepts) were modelled at catchment area-level to account for the hierarchical structure of the dataset. Age, sex, their interaction and minority status were treated as *a priori* confounders. I entered catchment area-level variables into our models using a forward-fitting strategy, per their strength of association with incidence in univariable analyses, assessed via Akaike's Information Criterion (AIC; lower scores indicate better model fit). Model building was assessed via likelihood ratio test (LRT).

In sensitivity analyses, I inspected the extent of bias introduced into our results due to diagnoses for a small proportion of cases being from clinical notes rather than OPCRIT. I also tested whether population density was associated with FEP incidence *within* countries in *post hoc* sensitivity analyses, stratifying by country, given the previous literature (Vassos et al., 2012).

4.4 Results

4.4.1 Sample description

I identified 2,774 people presenting with a first episode of ICD-10 psychotic disorder during 12.93 million person-years, corresponding to a crude incidence of 21.4 (95%CI: 19.4-23.4) per 100,000 person-years. A total of 1,578 of FEP participants (56.9%) were men, although this varied by catchment area (χ^2: 34.3, p<0.01), from 48.8% (40 of 82; Oviedo) to 77.8% (21 of 27; Cuenca) (Table 4.2). A total of 1,091 (39.8%) of FEP participants were from a minority background, varying (χ^2: 45.8, p<0.01) from just 2.8% (1 of 36; Santiago) to 75.6% (198 of 262; Southeast London). By comparison, almost 50% of the population at-risk were men, and 23% were from a minority group (Table 4.2).

Median age-at-first-contact was 30.5 years (IQR: 23-41), but this varied from 26 (IQR:21-37) in Cuenca to 35.5 (IQR: 28-42) in Veneto (Kruskal-Wallis χ^2:51.3; 16df, p<0.01) (Table 4.2). Median age-at-first-contact was earlier in men (28, IQR: 22-38) than women (34, IQR: 26-45; Mann-Whitney U-test: -11.1, p<0.01). There was no difference in median age-at-first-contact by minority status (Mann-Whitney U-test: 1.0,

p=0.31). Median DUP was 8 weeks (IQR: 2-35 weeks), varying from 2.5 weeks (IQR: 1-7) in Madrid to 26 (IQR:2-77) in Cuenca (Kruskal-Wallis χ^2: 119.8; 16df, p<0.01) (Table 4.2).

Overall, 78.7% (2,183 of 2,774) of participants received a non-affective diagnosis (Table 4.2) (crude incidence: 16.9 per 100,000 person-years, 95%CI: 16.2-17.6). A further 19.9% (551 of 2,774) received a diagnosis of affective psychosis (incidence: 4.3 per 100,000 person-years; 95%CI: 3.9-4.6). Remaining participants (1.4%; 40 of 2,774) were diagnosed with psychotic disorder, not otherwise specified. Median age-at-first-contact was younger for non-affective (30 years; IQR: 23-41) than affective psychoses (32 years; IQR: 24-45; Mann-Whitney U-test: -2.5; p=0.01); a higher proportion of women (25.0% vs. 19.7%; χ^2: 30.7; p<0.01) and minority groups (41.1% vs. 36.2%; χ^2: 4.2; p=0.04) were diagnosed with affective psychoses.

4.4.2 Variation by demographic variables

The age pattern of FEP incidence differed between men and women (Figure 4.1; LRT-χ^2 on 8df =119.3, p<0.01). Crude rates peaked for men between 18-24 years old (61.0 per 100,000 person-years; 95%CI: 59.0-63.1) and decreased steeply thereafter. For women, incidence also peaked in the youngest age group at 27.0 per 100,000 person-years (95%CI: 24.9-29.1), but decreased more gradually, with a small secondary peak between 50-54 years. Rates were higher in minority groups (IRR: 1.59, 95%CI: 1.46-1.72), after multivariable adjustment for age, sex, their interaction and relevant catchment-area level characteristics (see below).

Incidence of non-affective psychoses peaked in men aged 18-24 at 50.3 per 100,000 person-years (95%CI: 48.3-52.4) and declined steeply until age 35, after which decline slowed (Figure 4.2). Incidence in women also peaked in the youngest age group at 19.8 per 100,000 person-years (95%CI: 17.7-21.9), but declined only gradually across the life course (Figure 6.2). Rates were higher in minority groups (IRR: 1.63, 95%CI: 1.49-1.79) following adjustment for age, sex, their interaction and relevant catchment-area characteristics. Incidence of affective disorders followed a less distinct pattern for both genders. It peaked between ages 18-24 at 9.7 per 100,00 person-years for men (95%CI: 7.6-11.9) and 7.1 for women (95%CI: 4.9-9.3) and appeared to be lowest in the oldest age groups, but decline was non-linear and a slight increase in middle age was observed (although there was overlap of confidence intervals, Figure 4.3). Rates were higher in minority groups (IRR: 1.47, 95%CI: 1.22-1.76) following adjustment for age, sex, their interaction and relevant catchment-area characteristics.

Table 4.2: Population and sample characteristics, by catchment area

Setting	Total person-years	Men (%)	Ethnic majority (%)	Total cases	Non-affective psychoses (%)	Affective psychoses (%)	Men (%)	Ethnic Majority (%)	Median age of first contact (IQR)	Median DUP in weeks (IQR)
England										
Southeast London	426,453	212,981 (49.9)	175,706 (41.2)	262	245 (93.5)	17 (6.5)	141 (53.8)	64 (24.4)	32 (24-43)	10 (2-50)
Cambridgeshire	1,554,423	782,607 (50.4)	1,238,172 (79.7)	266	185 (69.6)	77 (29.0)	151 (56.7)	164 (61.7)	28 (22-37)	9 (3-52)
Netherlands										
Amsterdam	621,141	313,287 (50.4)	293,709 (47.3)	292	264 (90.4)	27 (9.3)	188 (64.4)	89 (30.5)	31 (24-42.5)	9.5 (2-68)
Gouda & Voorhout	766,770	384,975 (50.2)	651,786 (85.0)	167	122 (73.5)	39 (23.4)	101 (60.8)	127 (76.2)	29 (22-38)	4 (1-19)
Spain										
Madrid	414,786	205,367 (49.5)	329,425 (79.4)	89	72 (80.9)	12 (13.5)	58 (63.8)	76 (86.4)	30 (23-40)	2.5 (1.7)
Barcelona	883,894	426,258 (48.2)	688,283 (77.9)	108	96 (88.9)	8 (7.4)	62 (57.4)	82 (75.9)	28 (21.5-35.5)	7.5 (2-52)
Valencia	364,192	180,698 (49.6)	299,983 (82.4)	58	51 (87.9)	5 (8.6)	32 (55.1)	48 (82.7)	28 (24-39)	6 (3.5-17)
Oviedo	462,624	226,890 (49.1)	428,483 (92.6)	82	66 (80.5)	12 (14.6)	40 (48.8)	67 (81.7)	32 (24-43)	5.5 (2-32.5)
Santiago	574,944	286,767 (49.9)	556,192 (96.7)	36	30 (83.3)	5 (13.9)	20 (55.6)	35 (97.2)	33 (25-43.5)	13 (4-79)
Cuenca	195,074	102,697 (52.6)	160,724 (82.4)	27	26 (96.3)	0 (0.0)	21 (77.8)	20 (74.1)	26 (21-37)	26 (2-77)
France										
Paris	268,362	128,162 (47.8)	179,220 (66.8)	120	108 (90.0)	12 (10.0)	83 (69.2)	66 (55.0)	30.5 (22.5-40.5)	10.5 (5-25)
Val-de-Marne	510,632	242,334 (47.5)	342,091 (77.0)	212	134 (63.2)	76 (35.9)	107 (51.2)	142 (67.9)	30 (23-42)	8.5 (2-71)
Puy-de-Dôme	226,545	113,579 (50.1)	213,784 (94.4)	42	28 (66.7)	14 (33.3)	28 (66.7)	n/a	31 (22-46)	4 (2-10)
Italy										
Bologna	931,746	453,320 (48.9)	789,474 (85.1)	165	130 (78.8)	35 (21.2)	86 (52.1)	116 (70.3)	30 (23-41)	4 (1-15)
Veneto	505,508	259,282 (51.3)	446,523 (88.3)	104	82 (78.9)	14 (13.5)	56 (53.9)	83 (79.8)	35.5 (28-42)	n/a
Palermo	1,594,882	781,002 (49.0)	1,493,857 (93.7)	179	155 (86.6)	23 (12.9)	100 (55.9)	158 (88.3)	30 (24-40)	3 (1-13)
Brazil										
Ribeirão Preto	2,631,689	1,299,112 (49.4)	1,745,638(66.3)	565	389 (68.9)	175 (31.0)	304 (53.8)	302 (53.5)	32 (25-43)	13.5 (4-39)
Total	12,933,670	6,401,911 (49.5)	9,971,270 (77.1)	2,774	2,183 (78.7)	551 (19.9)	1,578 (57.0)	1,639 (60.1)	30.5 (23-41)	8 (2-35)
χ^2; p-value		$4.4*10^3$; <0.01	$1.4*10^6$; <0.01		172.6; <0.01	189.9; <0.01	34.3; <0.01	453.0; <0.01	51.3; <0.01	119.7; <0.01

Figure 4.1: Crude incidence and cumulative percentage of all FEP, by age and sex.

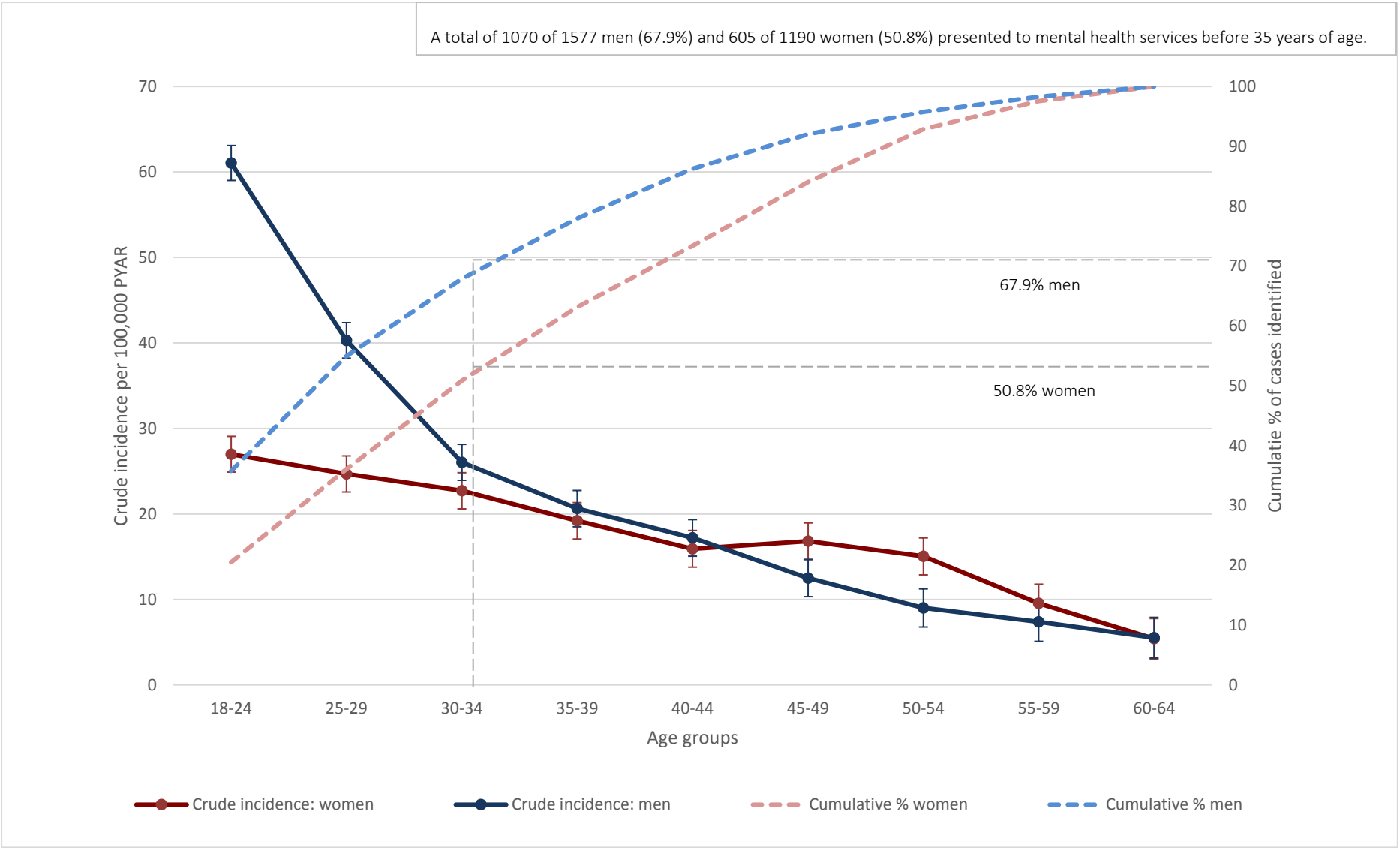


Figure 4.2: Crude incidence of non-affective psychoses and percentage of cases identified, by age and sex

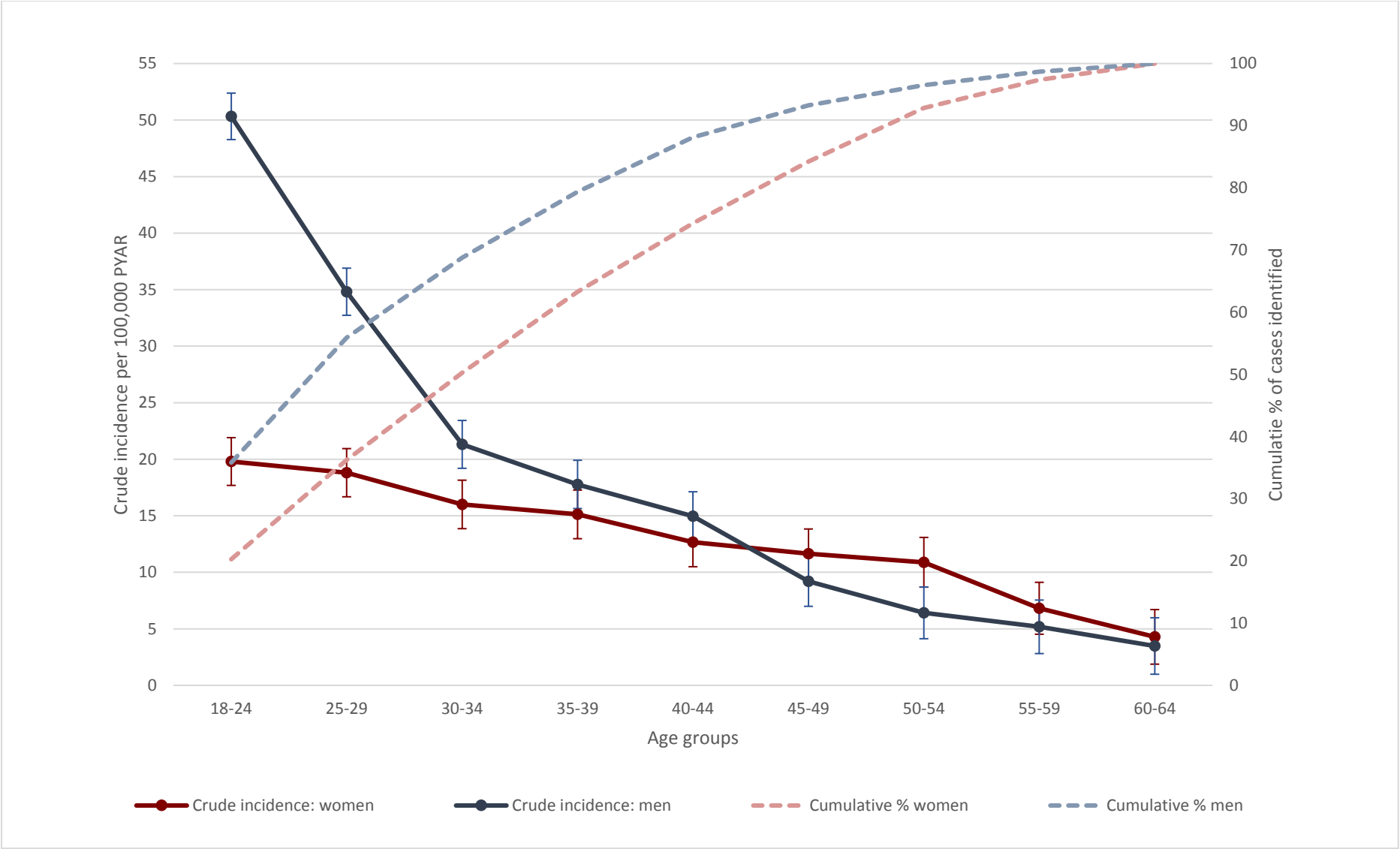
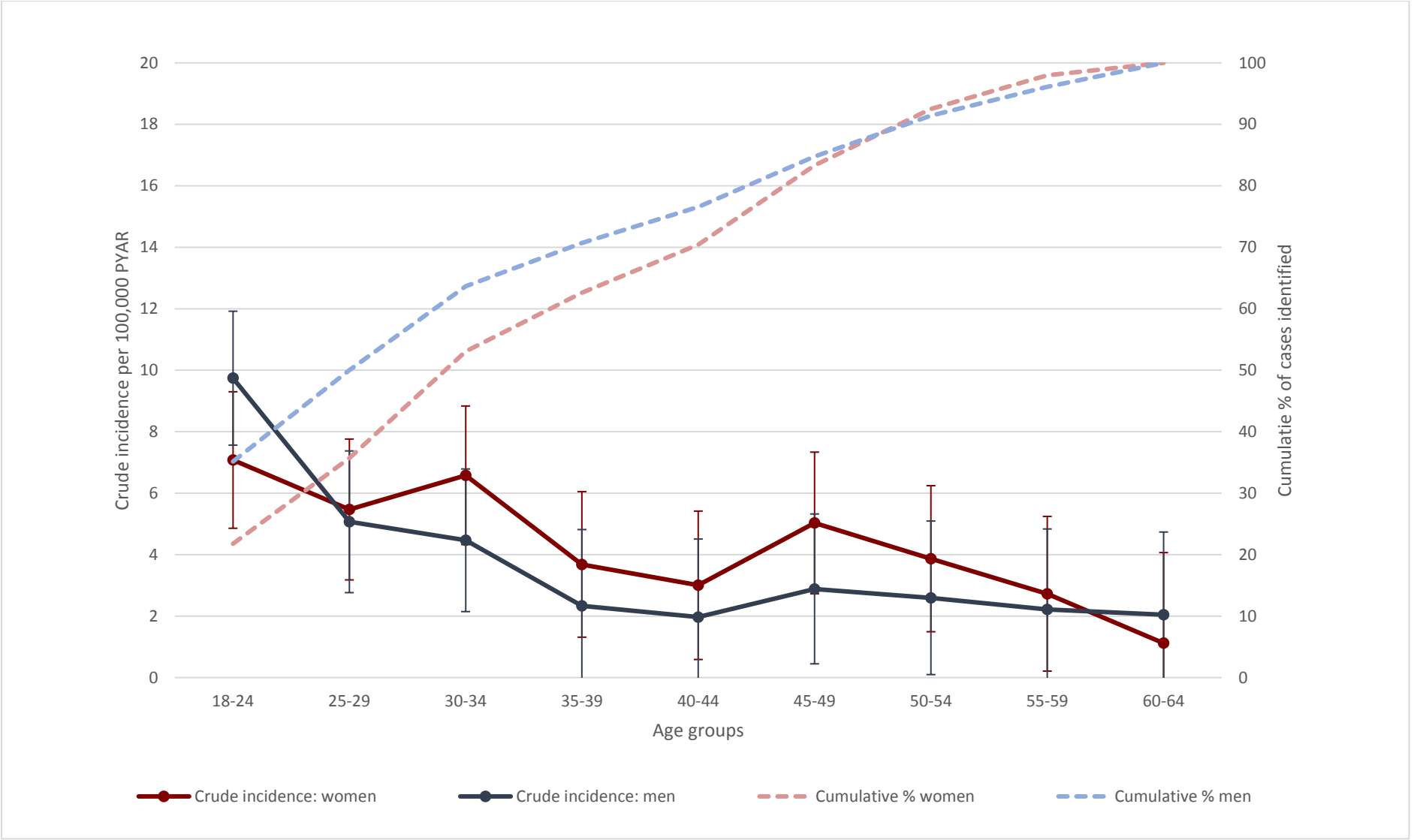


Figure 4.3: Crude incidence of affective psychoses and percentage of cases identified, by age and sex



4.4.3 Variation by catchment area

I observed ten-fold variation in crude FEP incidence across our catchment areas (Figure 4.4/Table 4.3), from 6.3 per 100,000 person-years (95%CI: 3.9-8.6) in Santiago to 61.4 (95%CI: 59.4-63.5) per 100,000 person-years in Southeast London. Age-sex standardisation had a negligible impact on this variation (Figure 4.4). Additional standardisation for minority status attenuated rates, although an almost eight-fold variation remained; compared with the overall EU-GEI incidence rate, SIRs varied from 0.29 (95%CI: 0.21-0.40) in Santiago to 2.21 (95%CI: 1.84-2.65) in Paris (Table 4.3).

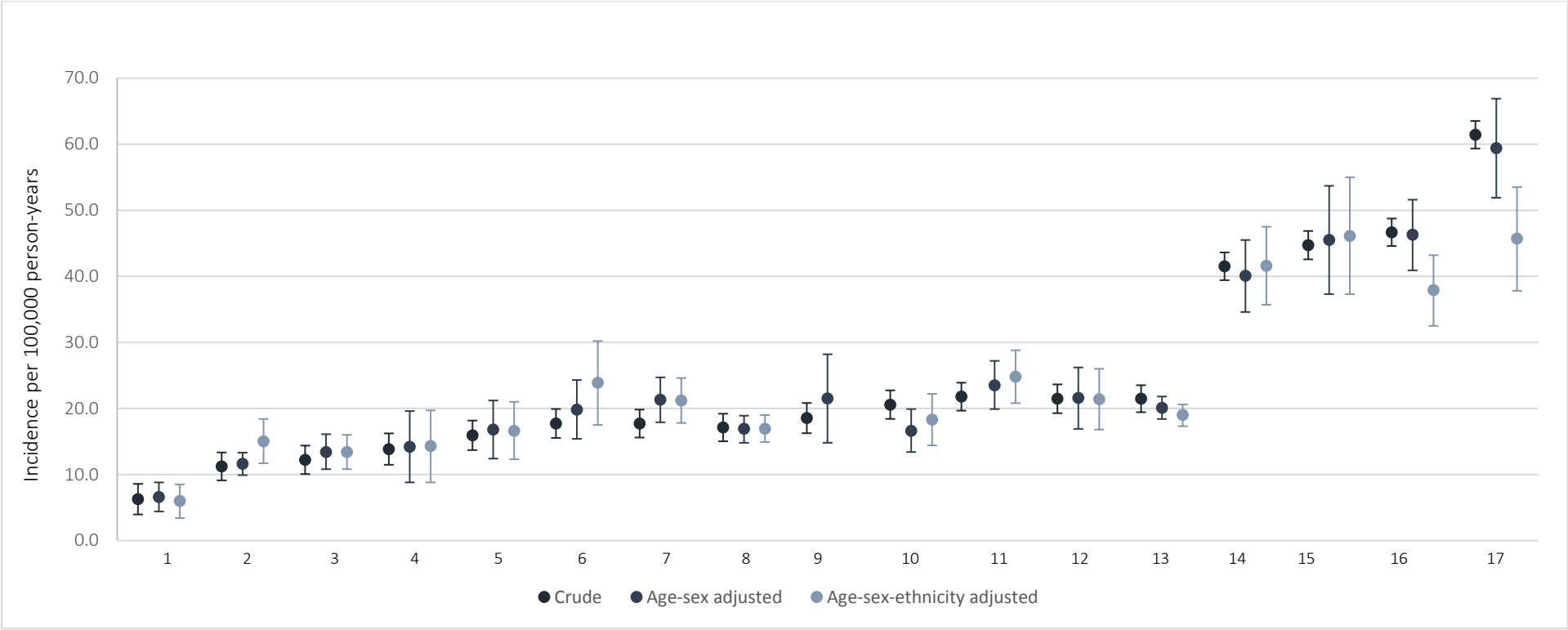
The crude and directly standardised incidence of non-affective and affective psychoses varied independently by setting (Table 4.3). We observed over a 10-fold variation in the crude rate of non-affective psychoses, from 5.2 new cases per 100,000 person-years (95%CI: 3.6-7.5) in Santiago to 57.5 (95%CI: 50.7-65.1) in Southeast London. Standardisation attenuated rates, although an eight-fold variation remained. Compared with the overall EU-GEI incidence rates, SIRs varied from 0.30 (95%CI: 0.21-0.43) in Santiago to 2.50 (95%CI: 2.19-2.85) in Southeast London. Crude rates of affective psychoses also varied by setting, from 0.9 per 100,000 person-years in Santiago (95%CI: 0.4-2.1) and Barcelona (95%CI: 0.5-1.8) to 14.9 in Val-de-Marne (95%CI: 11.9-18.6), more than a 17-fold difference. This difference remained following standardisation: SIRs ranged from 0.19 (0.08-0.46) in Santiago to 3.50 (2.75-4.45) in Val-de-Marne.

Table 4.3: Crude incidence rates and directly age-sex-minority standardised incidence ratios of all FEP, by catchment area

	All psychotic disorders		Non-affective psychoses		Affective psychoses	
Setting	Crude incidence rate (95%CI)	ASM standardised ratio (95%CI)[1]	Crude incidence rate (95%CI)	ASM standardised ratio (95%CI)[1]	Crude incidence rate (95%CI)	ASM standardised ratio (95%CI)[1]
England						
Southeast London	61.4 (59.4-63.5)	2.19 (1.93-2.48)	57.5 (50.7-65.1)	2.50 (2.19-2.85)	4.0 (2.5-6.4)	1.07 (0.66-1.74)
Cambridgeshire	17.1 (15.0-19.2)	0.81 (0.71-0.92)	11.9 (10.3-13.7)	0.71 (0.61-0.82)	5.0 (4.0-6.2)	1.19 (0.94-1.51)
Netherlands						
Amsterdam	46.7 (44.6-48.7)	1.81 (1.61-2.05)	42.2 (37.5-47.8)	2.03 (1.79-2.31)	4.3 (3.0-6.3)	1.00 (0.68-1.47)
Leiden	21.8 (19.7-23.9)	1.19 (1.01-1.39)	15.9 (13.3-19.0)	1.13 (0.94-1.36)	5.1 (3.7-7.0)	1.29 (0.93-1.78)
Spain						
Madrid	21.5 (19.3-23.6)	1.05 (0.83-1.26)	17.4 (13.8-21.9)	1.08 (0.86-1.37)	2.9 (1.6-5.1)	0.60 (0.34-1.05)
Barcelona	12.2 (10.1-14.4)	0.64 (0.53-0.78)	10.9 (8.9-13.3)	0.73 (0.59-0.89)	0.9 (0.5-1.8)	0.21 (0.11-0.43)
Valencia	15.9 (13.7-18.2)	0.79 (0.61-1.03)	14.0 (10.6-18.4)	0.88 (0.67-1.17)	1.4 (0.6-3.3)	0.36 (0.15-0.86)
Oviedo	17.7 (15.5-19.9)	1.14 (0.92-1.42)	14.3 (11.2-18.2)	1.15 (0.90-1.47)	2.6 (1.5-4.6)	0.83 (0.47-1.48)
Santiago	6.3 (3.9-8.6)	0.29 (0.21-0.40)	5.2 (3.6-7.5)	0.30 (0.21-0.43)	0.9 (0.4-2.1)	0.19 (0.08-0.46)
Cuenca	13.8 (11.5-20.8)	0.68 (0.17-1.00)	13.3 (9.1-19.6)	0.83 (0.56-1.22)	n/a	n/a
France						
Paris	44.7 (42.6-46.9)	2.21 (1.84-2.65)	40.2 (11.7-16.6)	2.45 (2.02-2.97)	4.5 (2.5-7.9)	1.38 (0.78-2.45)
Val-de-Marne	41.5 (39.4-43.6)	1.99 (1.73-.29)	26.2 (22.2-31.1)	1.63 (1.37-1.94)	14.9 (11.9-18.6)	3.50 (2.75-4.45)
Puy-de-Dôme	18.5 (16.3-20.8)	n/a	12.4 (8.5-17.9)	n/a	n/a	n/a
Italy						
Bologna	17.7 (15.6-19.8)	1.01 (0.87-1.19)	14.0 (11.7-16.6)	1.02 (0.85-1.22)	3.8 (2.7-5.2)	1.05 (0.74-1.47)
Veneto	20.6 (18.4-22.7)	0.88 (0.72-1.06)	16.3 (13.1-20.1)	0.87 (0.70-1.09)	2.8 (1.6-4.7)	0.60 (0.35-1.01)
Palermo	11.2 (9.1-13.3)	0.72 (0.62-0.83)	9.7 (8.3-11.4)	0.81 (0.69-0.96)	1.4 (1.0-2.2)	0.38 (0.25-0.58)
Brazil						
Ribeirão Preto	21.5 (19.5-23.5)	0.91 (0.83-1.00)	14.8 (13.4-16.3)	0.81 (0.72-0.90)	6.6 (5.7-7.7)	
Total	21.5 (19.4-23.4)	Reference	6.9 (16.2-17.6)	Reference	4.3 (3.9-4.6)	Reference

[1] ASM: age-sex-minority status directly standardised rates to the 2011 population structure of England and Wales

Figure 4.4: Crude, age-sex standardised and age-sex-minority standardised incidence rates of all FEP per catchment area



1. Santiago	5. Valencia	9. Puy-de-Dôme*	13. Madrid	17. Southeast London
2. Palermo	6. Bologna	10. Veneto	14. Val-de-Marne	
3. Barcelona	7. Cambridgeshire & Peterborough	11. Ribeirão Preto	15. Paris	
4. Cuenca	8. Oviedo	12. Gouda & Voorhout	16. Amsterdam	

* Data by ethnicity was not available for Puy-de-Dôme

4.4.4 Variation by putative environmental risk factors

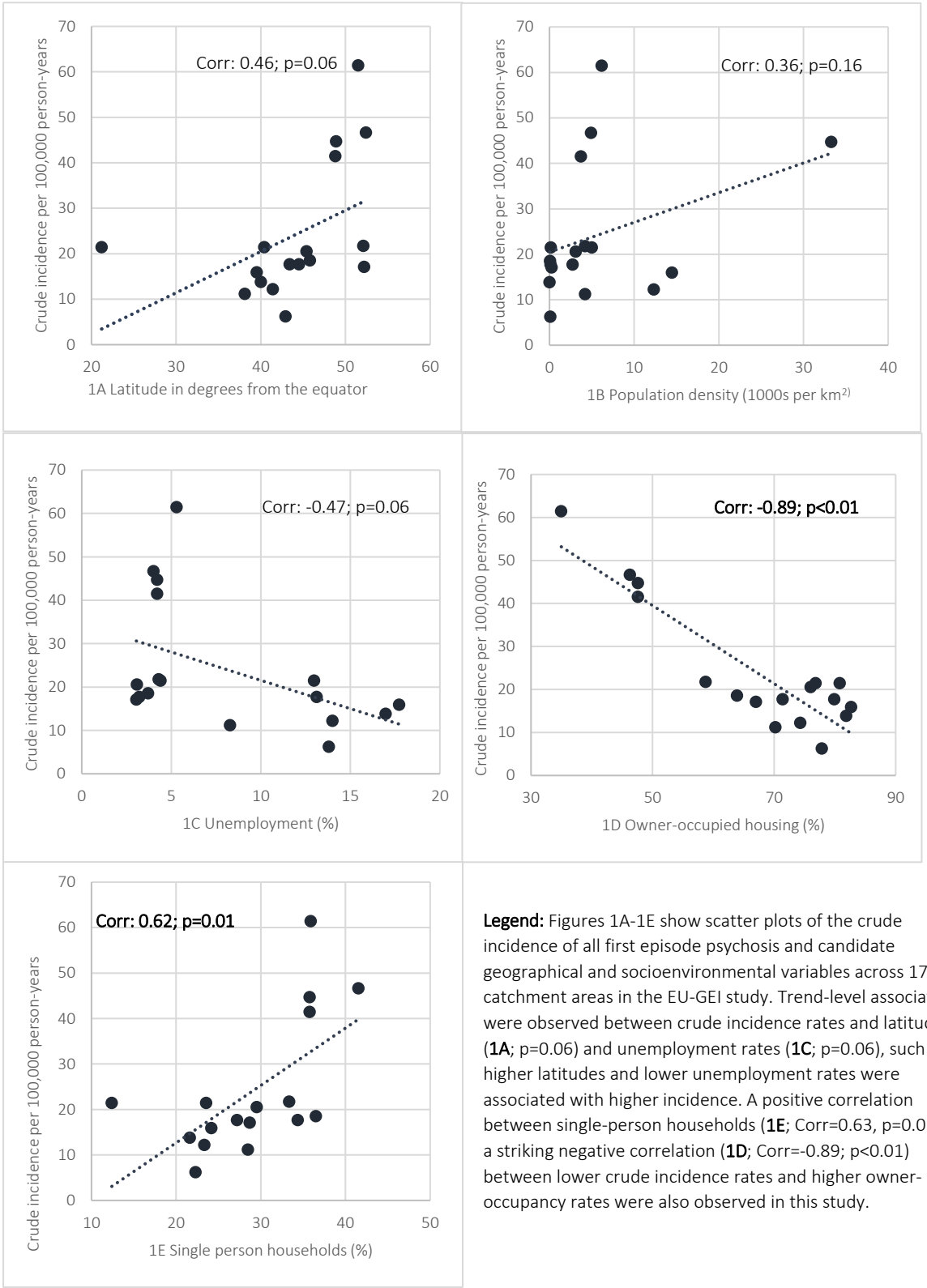
The distribution of catchment area-level exposure by catchment area are shown in Table 4.4 below, and their associations with crude FEP incidence are shown in Figure 4.5. Population density varied from 11.6 people per square kilometre in Cuenca to 33,260 in Paris. The range of latitude represented in this study is narrow: with the exception of Brazil (21 °south) all settings are located between 38 and 52 degrees north of the equator. Unemployment is particularly high in all Spanish settings, and the percentage of single-person household varies from 12.4 in Ribeirão Preto to 41.5 in Amsterdam. The percentage of houses that is owner-occupied is lowest in Southeast London at 35.0 and highest in Cuenca at 81.9 (Table 4.4).

Table 4.4: Distribution of geographical and socio-environmental exposures, by catchment area

Setting	Population density in people/km²	Latitude	Unemploy-ment (%)	Single household (%)	Owner occupied homes (%)
England					
Southeast London	6,162.3	51.5 °N	5.3	35.9	35.0
Cambridgeshire	241.5	52.2 °N	3.0	26.7	67.0
Netherlands					
Amsterdam	4,908.0	52.4 °N	4.0	41.5	46.3
Gouda & Voorhout	4,208.0	52.1 °N	4.3	33.3	58.7
Spain					
Madrid	4,997.2	40.4 °N	13.0	23.6	76.8
Barcelona	12,362.5	41.4 °N	14.0	23.3	74.3
Valencia	14,467.9	39.5 °N	17.8	24.1	82.7
Oviedo	141.9	43.4 °N	13.1	27.2	80.0
Santiago	102.3	42.9 °N	13.8	22.3	77.9
Cuenca	11.6	40.0 °N	17.0	21.6	81.9
France					
Paris	33,260.0	48.9 °N	4.2	35.8	47.6
Val-de-Marne	3,721.2	48.8 °N	4.5	35.8	47.6
Puy-de-Dôme	68.5	45.8 °N	3.7	36.5	63.9
Italy					
Bologna	2,744.0	44.5 °N	3.2	34.4	71.4
Veneto	3,100.0	45.4 °N	3.2	29.5	76.0
Palermo	4,200.0	38.1 °N	8.1	28.5	7.02
Brazil					
Ribeirão Preto	145.2	21.1°S	4.4	12.4	80.8

A positive correlation (0.63, p=0.01) was observed between percentage of single-person households and crude FEP incidence, and a striking negative correlation (-0.82, p<0.01) between percentage of owner-occupied houses and crude FEP incidence (Figure 4.5).

Figure 4.5: Correlation between crude incidence of all FEP and geographical and socioenvironmental variables



Legend: Figures 1A-1E show scatter plots of the crude incidence of all first episode psychosis and candidate geographical and socioenvironmental variables across 17 catchment areas in the EU-GEI study. Trend-level associations were observed between crude incidence rates and latitude (**1A**; p=0.06) and unemployment rates (**1C**; p=0.06), such that higher latitudes and lower unemployment rates were associated with higher incidence. A positive correlation between single-person households (**1E**; Corr=0.63, p=0.01) and a striking negative correlation (**1D**; Corr=-0.89; p<0.01) between lower crude incidence rates and higher owner-occupancy rates were also observed in this study.

Univariable random intercepts Poisson regression showed that greater owner-occupancy (IRR for a 10% increase: 0.73, 95%CI: 0.65-0.81) and unemployment (IRR for a 10% increase: 0.54, 95%CI: 0.34-0.84) were associated with lower incidence of all psychotic disorders, while percentage of single-person households (IRR for a 10% increase: 1.68, 95%CI: 1.24-2.27) was associated with higher incidence (Table 4.5).

Table 4.5: Univariable and multivariable random intercepts Poisson regression of all FEP

Variable	Univariable IRR (95% CI)	Univariable Wald p-value	Multivariable IRR (95% CI)[1]	Multivariable LRT p-value
Individual-level				
Minority status (vs majority)	1.69 (1.56 – 1.84)	<0.01	1.59 (1.46 – 1.72)	<0.01
Setting-level				
Distance from equator (degrees)	1.03 (1.00 – 1.07)	0.07	0.99 (0.97 – 1.01)	0.46
Population density (per 1000 people per km^2)	1.02 (0.99 – 1.05)	0.15	1.01 (0.99 – 1.02)	0.44
Owner-occupancy (10%)	0.73 (0.65 – 0.81)	<0.01	0.76 (0.70 – 0.83)	<0.01
Single-person households (10%)	1.68 (1.24 – 2.27)	<0.01	1.06 (0.78 – 1.43)	0.73
Unemployment (10%)	0.54 (0.34 – 0.84)	<0.01	0.90 (0.66 – 1.23)	0.51

IRR: Incidence rate ratio; LRT: likelihood-ratio test

[1]Models adjusted for age, sex, their interaction and, for setting-level variables, ethnicity. IRR for non-significant setting-level variables obtained from a model after additional adjustment for owner-occupancy.

A null random intercepts Poisson model confirmed substantial variation in incidence by catchment area (σ: 0.32, p<0.01), which persisted after adjustment for age, sex, their interaction, and minority status (σ: 0.23, p<0.01). In multivariable analyses, FEP incidence was 1.59 (95%CI: 1.46-1.72) times higher in minority groups compared with the majority population, and lower in catchment areas with owner-occupied homes (IRR for a 10% increase in owner-occupancy: 0.76, 95% CI: 0.70-0.83) after adjustment for age, sex and their interaction. No other setting-level variables, including latitude (IRR: 0.99; 95% CI: 0.97-1.01), improved the final model (Table 4.5), where residual variance by catchment area remained, albeit attenuated (σ: 0.06, p=0.02).

Multivariable Poisson regression revealed that, as for all FEP, owner-occupancy was associated with incidence of non-affective psychoses (IRR: 0.76; 95%C: 0.69-0.83) although residual variance by catchment-area remained (σ: 0.06, p=0.02). For the affective psychoses, unemployment was associated with incidence (IRR: 0.30; 95%C: 0.17-0.53), and residual variance by catchment area remained (σ: 0.20, p=0.02). In multivariable regression, elevated rates of both disorders were associated with minority status to a similar extent (Table 4.6).

Table 4.6: Univariable and multivariable random intercepts Poisson regression of non-affective and affective psychotic disorders

Variable	Non-affective psychoses		Affective psychoses	
	Univariable IRR (95% CI)	Multivariable IRR (95% CI)[1]	Univariable IRR (95% CI)	Multivariable IRR (95% CI)[1]
Individual-level				
Minority status (vs majority)	1.75 (1.59 – 1.92)	1.63 (1.49 – 1.79)	1.54 (1.28 – 1.85)	1.47 (1.22 – 1.76)
Setting-level				
Distance from equator (degrees)	1.03 (1.00 – 1.06)[2]	0.99 (0.97 – 1.02)	1.02 (0.97 – 1.08)	1.00 (0.97 – 1.03)
Population density (per 1000 people per km^2)	1.03 (1.00 – 1.07)[3]	1.01 (1.00 – 1.03)[4]	1.00 (0.95 – 1.05)	1.00 (0.96 – 1.03)
Owner-occupancy (10%)	0.72 (0.65 – 0.80)	0.76 (0.69 – 0.83)	0.74 (0.58 – 0.93)	0.95 (0.77 – 1.16)
Single-person households (10%)	1.71 (1.26 – 2.32)	1.10 (0.81 – 1.49)	1.58 (0.93 – 2.68)	0.99 (0.67 – 1.47)
Unemployment (10%)	0.60 (0.37 – 0.97)	1.07 (0.79 – 1.47)	0.27 (0.15 – 0.48)	0.30 (0.17 – 0.53)

IRR: Incidence rate ratio; IRR in bold are statistically significant at p<0.05

[1]Models adjusted for age, sex, their interaction and, for setting-level variables, ethnicity. IRR for non-significant setting-level variables obtained from a model after additional adjustment for owner-occupancy.
[2]p=0.07
[3]p=0.06
[4]p=0.12

4.4.5 Sensitivity analyses

A proportion of cases were diagnosed using clinical rather than research diagnoses (N=367; 13.2%), given insufficient data to complete an OPCRIT. More women (14.8% vs. 11.9%; χ^2 on 1df=4.4, p=0.03) and participants from minority backgrounds (14.9% vs. 12.2%; χ^2 on 1df=4.4, p=0.04) were diagnosed via clinical ratings, though no differences by age group were observed (χ^2 on 8df=10.2, p=0.25). A higher proportion of affective psychoses were obtained via clinical diagnoses (31.4% vs. 18.5%; χ^2 on 1df=31.7, p<0.01). Excluding people with a clinically-based diagnosis from our analyses did not substantially alter findings (Table 4.7).

Table 4.7: Multivariable random intercepts Poisson regression excluding 367 participants with clinically-based diagnoses

Variable	All FEP IRR (95% CI)[1]	Non-affective psychoses IRR (95% CI)[1]	Affective psychoses IRR (95% CI)[1]
N participants (% full sample)	*2,407 (86.8)*	*1,943 (89.0)*	*441 (80.1)*
Individual-level			
Minority status (vs majority)	1.55 (1.42 – 1.69)	1.57 (1.42 – 1.73)	1.38 (1.12 – 1.70)
Setting-level			
Distance from equator (degrees)	0.99 (0.97 – 1.02)	0.99 (0.96 – 1.02)	1.01 (0.97 – 1.05)
Population density (per 1000 people per km^2)	1.01 (0.99 – 1.03)	1.01 (0.99 – 1.03)	1.00 (0.97 – 1.04)
Owner-occupancy (10%)	0.75 (0.68 – 0.83)	0.75 (0.68 – 0.84)	0.86 (0.68 – 1.10)
Single-person households (10%)	1.08 (0.77 – 1.52)	1.07 (0.74 – 1.54)	1.20 (0.75 – 1.94)
Unemployment (10%)	0.96 (0.68 – 1.37)	1.10 (0.76 – 1.59)	0.31 (0.16 – 0.60)

Legend: Sensitivity analysis to inspect possible bias introduced due to 367 participants diagnoses from clinical diagnoses rather than OPCRIT-based diagnoses
[1]Models adjusted for age, sex, their interaction and, for setting-level variables, ethnicity. IRR for non-significant setting-level variables obtained from a model after additional adjustment for owner-occupancy

In *post hoc* multivariable models (Table 4.8), population density was positively associated with FEP incidence in England (IRR: 1.17; 95%CI: 1.13-1.21) and the Netherlands (IRR: 1.89; 95%CI: 1.40-2.56), but not Spain or France, while a negative association was observed in Italy (IRR: 0.72; 95%CI: 0.62-0.83).

*Table 4.8: Effect of population density on incidence of all FEP from multivariable random intercepts Poisson regression, by country**

Country	Number of settings	Multivariable IRR (95%CI)[1]	Wald p-value
England	2	1.17 (1.13-1.21)	<0.01
The Netherlands	2	1.89 (1.40-2.56)	<0.01
Spain	6	1.01 (0.96-1.06)	0.61
France	3	1.01 (1.00-1.03)[2]	0.14
Italy	3	0.72 (0.62-0.83)	<0.01

IRR: incidence rate ratio
* Brazil excluded from these analyses as only a single setting was part of the study here
[1] Adjusted for age, sex, their interaction, minority status and owner-occupancy
[2] Adjusted for age, sex, their interaction and owner-occupancy only (data on majority status not available for Puy-de-Dôme)

4.5 Discussion

4.5.1 Main findings

In the largest international study of the epidemiology of psychotic disorders for 25 years, I observed substantial variation in FEP incidence across 17 catchment areas in six countries, confirming differential risk by place and person (hypotheses 1-3, Table 4.9). In line with previous studies, I observed higher rates of all psychotic disorders in minority groups (Bourque et al., 2011) and amongst young people (Thorup et al., 2007), particularly for men (van der Werf et al., 2014). I confirmed a small, but robust secondary peak in risk of all FEPs for women after 45 years old. Catchment areas with higher owner-occupancy levels were associated with lower incidence rates, implicating socioeconomic factors in the presentation of psychotic disorders, in line with previous research (Allardyce et al., 2005; Kirkbride et al., 2014) (Hypothesis 4, Table 4.9).

Table 4.9: Reappraisal of hypotheses

	Hypothesis	Outcome
1.	Substantial variation in incidence of all FEP, non-affective psychoses and affective psychoses	Confirmed
2.	Variance not accounted for by age-sex-minority standardisation	Confirmed
3.	Latitude, population density, unemployment, single-person household, owner occupied homes partially accounted for variance	Partially confirmed
4.	Incidence was higher at higher latitude, population density, unemployment, single-person households and lower owner-occupied homes	Partially confirmed

4.5.2 Strengths and limitations

Findings should be interpreted in light of the strengths and limitations of this study. The large sample size allowed me to estimate three psychotic outcomes in 17 settings with a high degree of precision. To

minimise ascertainment bias, all researchers received training via face-to-face epidemiological training sessions, regular teleconferencing, online training manuals and inter-rater reliability protocols. Nonetheless, some limitations of the multinational design need to be acknowledged.

Detection of patients who never present to services is an issue for all epidemiologic studies, and rate estimates should be interpreted as the treated incidence. While the overarching case ascertainment methodology was similar across all settings, adaptation to local health care systems was necessary. For example, primary care in each catchment area may have referred different proportions of patients with FEP to secondary mental health care services, but referral guidelines were very similar across national settings; these guidelines all urge prompt referral of anyone with FEP. That said, I did not assess whether referral practices were consistent within and between catchment areas. Difference in the average timing of referral may have affected the case mix within the FEP category, but not the overall number of referrals; each centre was in a steady state.

Differences in the organisation of secondary mental health care services across localities may also have influenced detection of patients. In England and the Netherlands, for example, the widespread commissioning of Early Intervention in Psychosis services as centralised referral points for young people with psychosis may have led to improved detection of new cases. The leakage study in Brazil revealed a substantial number of new cases at this site (n=279; 49%) while similar approaches in two French sites (Paris and Val-de-Marne) identified far fewer missed cases (n=7; 6% and n=28; 13% respectively), in line with previous studies (Kirkbride et al., 2006). In all settings, comprehensive and regular contact with services should have helped to minimise under-ascertainment, although it is possible that a handful of cases, including those treated privately, may have been missed. I believe, however, that such a bias would not have fully accounted for the eight-fold variation between catchment areas.

Validated semi-structured interview schedules (SCAN, CASH, SCID, DIGS, see Section 3.4) were used, where possible, to obtain standardised research-based OPCRIT diagnoses close to the patient's first presentation. I have no reason to believe use of different schedules by setting biased estimates; indeed, this choice was adapted to local expertise to minimise bias, which may have otherwise arisen from using unfamiliar interview schedules. The incidence of all FEP was considered the primary outcome, as this is useful for, and consistent with contemporary practice in psychosis management and treatment, to allow symptoms to evolve at first presentation and minimise stigma. This practice is also consistent with some evidence of diagnostic instability in the early course of disorder (Heslin et al., 2015), which is particularly apparent for psychotic disorders other than schizophrenia (Bromet et al., 2005). Although I relied on clinical diagnoses in a small proportion of cases, where some biases were evident, this did not alter interpretation of the main findings.

I classified ethnic minority status as a binary variable, following official definitions used in each country to distinguish majority and minority groups. This simple approach was valid, except perhaps in France where I may have misclassified some minority groups in the majority population. France does not differentiate between people born in mainland France versus its overseas territories, nor is able to identify second-generation (French-born) migrants. This misclassification would likely have led to conservative bias in IRRs with respect to minority status, as would have our decision to code participants with missing ethnicity data (n=5; 0.2%) to the majority group, given strong previous evidence of raised rates in minority groups (Bourque et al., 2011). Using a binary variable to control for putative effects of minority may have resulted in residual confounding; psychosis risk by ethnicity will be studied in greater detail in future EU-GEI publications, and in the remainder of this thesis using case-control data.

I used a consistent methodology in European catchment areas to estimate measures of the social environment, with comparable data taken from the Brazilian census. European data could only be obtained at the NUTS-2 regional level – larger than the catchment areas. This may have led to exposure misclassification, although the effect of this ecological bias is difficult to determine.

While I controlled for several risk factors simultaneously (age, sex, minority status and catchment area-level factors), I was unable to investigate the role of other putative risk factors for psychosis, including cannabis use (Moore et al., 2007), urban birth (Vassos et al., 2012), family history of psychosis (Esterberg et al., 2010), childhood trauma (Varese et al., 2012) or genetic risk (Schizophrenia Working Group of the Psychiatric Genomic Consortium, 2015). Such risk factors are not routinely available in denominator estimates, but will be investigated in the case-control data from the EU-GEI study (see Chapters 6 and 7).

4.5.3 Comparison with the previous literature

Overall, incidence rates observed in this Chapter are consistent with the literature, although between-study heterogeneity in methodologies, inclusion criteria and diagnoses studied make direct comparisons difficult. For example, the incidence of broadly-defined schizophrenia in the WHO 10-country study (Jablensky et al., 1992) varies from 15 to 42 per 100,000 person-years, although that study used a different age range (15-54 years) and did not consider affective psychoses (Jablensky et al., 1992). In our study, comparable rates of non-affective psychoses vary from 5 to 41 per 100,000 person-years after standardisation for age, sex and migrant status. The systematic review presented in Chapter 2 observed a pooled crude incidence of all psychotic disorders of 28.3 (95%CI: 23.4-34.3) per 100,000 person-years, somewhat higher than the overall crude incidence rate I observed here (21.4). Such comparisons should be interpreted cautiously, given heterogeneity in estimation methods, and setting; few incidence studies have been conducted in southern Europe until recently (Lasalvia et al., 2014; Mulè et al., 2017; Tarricone et al., 2012), where rates appeared uniformly low, despite inclusion of urban catchment areas.

The higher rates of disorder observed in men (van der Werf et al., 2014), younger age groups (Thorup et al., 2007; van der Werf et al., 2014; Kirkbride, Hameed, Ankireddypalli, et al., 2017) ethnic minorities (Bourque et al., 2011), and for non-affective psychoses (Kirkbride et al., 2012), are also frequently reported in the literature. The present study provides further robust evidence of a secondary peak in psychosis risk for women between 45-60 years old, building on previous observations (Bromet et al., 1992; Jackson et al., 2013; Kirkbride et al., 2006). This effect warrants further investigation, with previous research hypothesizing a protective role for estrogen prior to menopause (Grigoriadis & Seeman, 2002), or the potential for an increase in psychosocial stressors experienced during this stage of the life course. Our findings add further evidence to the observation that early intervention services with an upper age limit of 35 years (or lower) may lead to gendered mental health inequalities (Lappin et al., 2016): only 50.8% of women with psychosis were identified before age 35 in our settings, compared with 67.9% of men (Figure 4.1).

Incidence not only varied by person, but importantly, by place, suggesting that the social environment may shape incidence patterns of FEP. The best-fitting models of all FEP and non-affective psychoses suggested that owner-occupancy levels predicted incidence, although residual variation at the setting-level was not explained by other catchment area-level measures. Although I can't exclude the possibility of reverse causality, owner-occupancy may also be a proxy for a variety of social exposures, most obviously socioeconomic position (Kirkbride et al., 2007), but extending to the social stability and cohesiveness, previously associated with psychosis (Allardyce et al., 2005). The incidence of FEP appeared to be lower in southern Europe, but I found no evidence of variation by latitude in multivariable models. Nevertheless, settings were located within a narrow band (38° to 53°north of the equator), except for Brazil (21° south of the equator). This location may have contributed to the null finding, and the absence of high rates of psychosis in our southern European settings, particularly in major urban centres, requires further investigation; in southern Europe incidence patterns with respect to population density appear to diverge from those observed in northern Europe (Table 4.8). Variation in the incidence of affective disorders, with lower rates in catchment areas with higher levels of unemployment is counterintuitive and unexpected. Whilst this might be explained by the inclusion of Spain, where rates of affective disorder were uniformly low and unemployment was a range of magnitude larger than in other countries, the finding may be a chance finding and further research is required to replicate and explain it.

Our study was predominantly based in Europe, and will be complemented by studies in other settings, including low and middle income [LAMI] settings (Morgan et al., 2017). Outside of Brazil, the only non-European country included in the present study, a dearth of high quality epidemiology data exists on psychotic disorders. Findings in Brazil were congruent with previous research (Menezes & Scazufca, 2007), and tentatively suggested that psychosis morbidity in LAMI countries could be considerable.

4.5.4 Conclusion

In this international, multi-centre study I found that treated incidence of psychotic disorders varied eight-fold between catchment areas after standardisation for age, sex and minority status. Rates were higher in younger people, in men, minorities, and areas with lower levels of owner-occupied housing, although substantial variation between catchment areas, and by broad diagnosis, remained. These results suggest that there is pronounced variation in the healthcare burden of psychosis worldwide.

Chapter 5 - Searching for the cause of higher rates of psychosis in ethnic and other minority groups

5.1 Introduction

The research synthesised and presented in Chapters 2 and 4 demonstrated that there is substantial variation in incidence of psychosis across person and place. One persistent finding is that ethnic minorities have a higher incidence of psychosis than the (White) majority. In the remaining Chapters of my thesis I aim to investigate why this is the case.

This Chapter is a temporary departure in my thesis, both in writing style and in content. This Chapter will adopt a philosophical perspective to outline a theory of causality suited to explain the proposed model of why rates of psychosis are higher in (ethnic) minorities. In the next Section, I will outline a theory of causality adequate to explain why rates of psychosis are higher in ethnic minorities. I will then critically review the epidemiological evidence as described in Section 1.4, and conclude that epidemiology alone is insufficient to solve our causal puzzle: we need both epidemiology and the social sciences. I subsequently examine literature from the social sciences and use this to form a background to Chapters 6 and 7, where I will return to epidemiology to formulate and empirically test the hypotheses as theorised in this Chapter.

5.2 A note on causality

This Chapter is concerned with causality. In essence, it seeks to answer the following question: what produces the higher risk of psychosis in ethnic minority groups? Epidemiologists uses many definitions of and criteria for causality, and there is no single well-articulated and widely-used definition (Parascandola & Weed, 2001). Therefore, in this Section, I outline an understanding of causality appropriate for this context, as well as some guidance on how to differentiate between causes and non-causally associated or enabling factors.

In line with Russo and Williamson, I take a cause to be a probabilistic association, underpinned by a mechanistic explanation (Russo & Williamson, 2007). For us to accept, for instance, infant exposure to cats as a cause for psychosis, we not only require a probabilistic association, but also a plausible mechanism. The mechanism (via *Toxoplasma gondii* infection) is indirect but not impossible, but as no probabilistic association appears to exist (Solmi, Hayes, Lewis, & Kirkbride, 2017), we do not readily accept causality for this claim.

By these criteria, there are many causes of psychosis, but none of them sufficient (guaranteed to produce) or necessary (must be present) to bring about disorder. For instance, lifetime cannabis use has a strong probabilistic association (a systematic review estimated a pooled adjusted OR of 1.41; 95%CI: 1.20-1.65)(Moore et al., 2007), and a plausible potential mechanistic association (the main active psychotropic component of cannabis caused transitory psychotic symptoms, particularly in those with an increased genetic risk (Van Os et al., 2010)) of psychosis. However, not all who smoke cannabis develop psychosis,

and not all those suffering with psychosis have ever smoked cannabis (see Chapter 6 for empirical evidence). Our theory of causation therefore needs to accommodate many causal contributions that are neither necessary nor sufficient. This means that monocausal accounts of disease, such as Koch's postulates (Koch, 1884; 1890), are inadequate for our causal question.

There are various multi-causal accounts that have gained popularity in epidemiology. One example is Rothman's idea of 'causal cakes' (Figure 5.1) (Rothman, 1976; 2005), which is very closely related to Mackie's original notion of INUS conditions, where any cause is an *Insufficient* but *Nonredundant* part of an *Unnecessary* but *Sufficient* condition (Mackie, 1965). Mackie explained this by referring to a house fire: it is not just the faulty electrical wiring that caused the house to burn down, but a combination of the faulty electrical wiring, the inflammable material nearby and the lack of adequate sprinkler installation. None of these three components are in themselves sufficient to burn down the house, only together they are. The faulty electrical wiring could have been replaced by an exploding gas canister, but the result would still have been the same (Mackie, 1965). This understanding of causality is well reflected in the epidemiology of psychosis, where there are a number of risk factors with both a strong probabilistic association and a convincing underlying mechanism, such as cannabis use (Moore et al., 2007) and childhood trauma (Varese et al., 2012) but none that is either necessary or sufficient.

Figure 5.1: Rothman's causal cakes (Rothman, 1976).

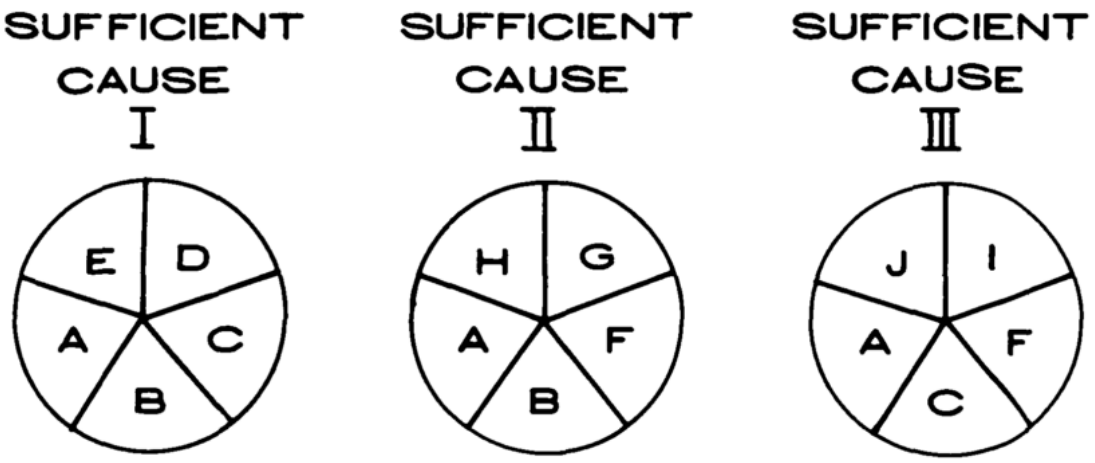
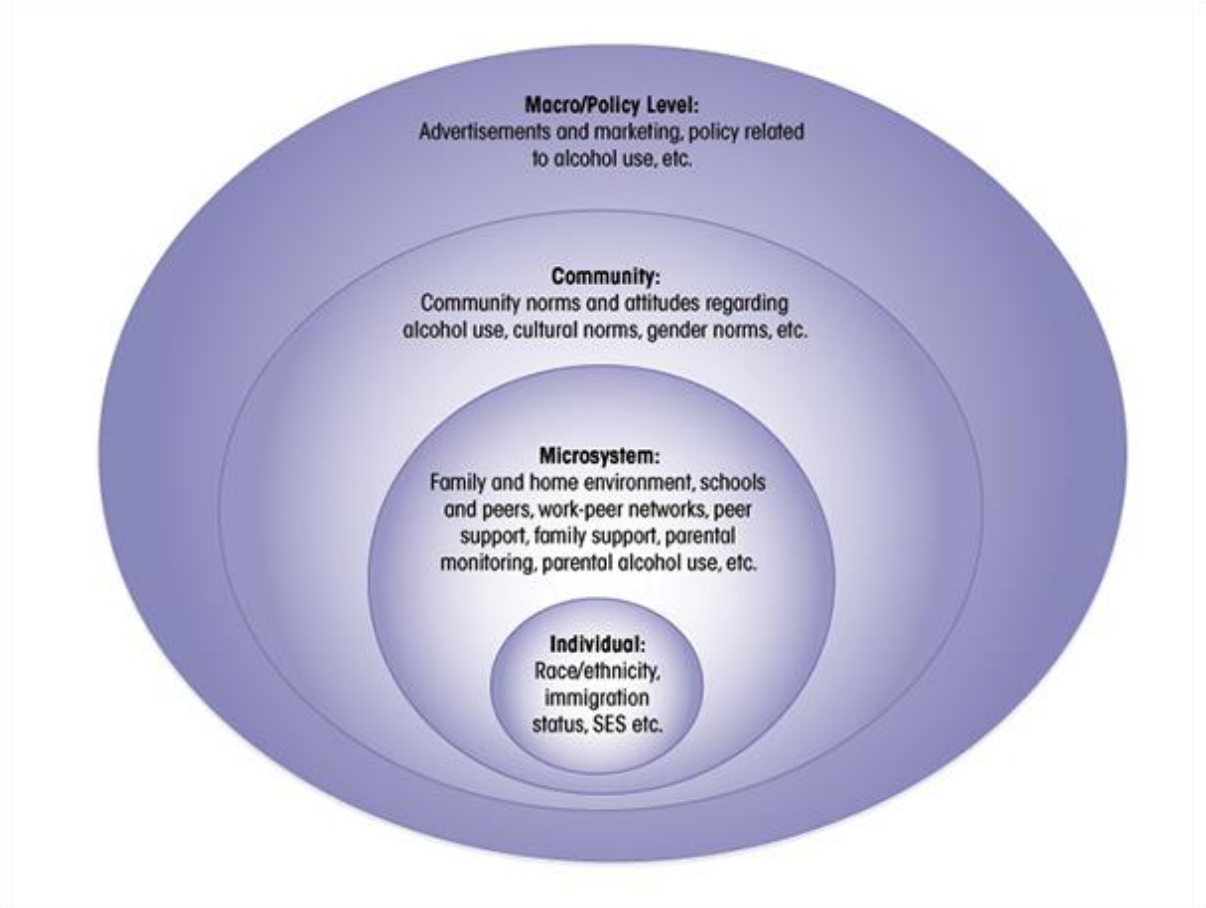


FIGURE 1. Conceptual scheme for the causes of a hypothetical disease.

Unfortunately, a multi-causal theory in this form does not provide us with method to assign particular salience to any single cause. The relevance of this becomes clear when considering dopaminergic dysfunction: this is hypothesised to be a 'final common pathway' in most individuals with psychosis (Howes & Kapur, 2009). However, in a multi-causal account as presented by Rothman and Mackie, it can only be represented as one element of the cake (element B in Figure 1, for example), and could not be considered different from the other causes, so we can't do justice to the unique salience of this potential cause. This philosophical problem is not unique to causality of psychosis, and Susser attempted to solve this by

proposing a multi-level multi-factorial theory of causality (Susser, 1973). He proposed nested levels of causality, determined by systems. A system is an abstract notion of connected factors in a somewhat coherent relationship. Figure 5.2 illustrates this by showing nested levels of socio-economic causes of alcohol use (Sudhinaraset, Wigglesworth, & Takeuchi, 2016).

Figure 5.2: Nested levels of socio-economic causes of alcohol use (Sudhinaraset et al., 2016).



In the above example, individual characteristics such as socio-economic status are one system, whereas national policy is another. According to Susser, causes are direct if their association is within the same system (such as women drinking more alcohol than men) and indirect if the association spans multiple systems (in societies with stricter laws licensing the sale of alcohol, there are lower individual levels of alcohol consumption). This however fails to account for potential interaction between levels (licencing laws impacting on men disproportionately, for example) but also, as Furman (2016) points out, to adequately account for indirect but salient causes, as it assumes that those causes that are direct are the most salient and those that are indirect are less salient.

We therefore need to be able to account for distant but salient causes, which we need to distinguish from 'enabling conditions' that also have to be in place, but aren't causal, such as the easy availability of cannabis paraphernalia, to return to our cannabis example. A common way to identify which risk factors are more likely to be causal in epidemiology is to assess them against the Bradford-Hill criteria of causation

(Table 5.1 below). No formal threshold is specified or intended but the criteria serve as a framework to guide interpretation of findings from observational data (Bradford Hill, 1965).

Table 5.1: The Bradford-Hill Criteria

Criterion		Meaning
1.	Strength	Effect size
2.	Consistency	Observed by different persons, in different places, circumstances and times
3.	Specificity	Cause specific to the effect
4.	Temporality	Does the cause precede the effect?
5.	Biological gradient	Is there a dose-response relationship?
6.	Plausibility	Is there a biologically plausible mechanism?
7.	Coherence	It can't seriously conflict with generally known facts.
8.	Experiment	Is there (semi-) experimental evidence?
9.	Analogy	Occasionally analogous risk factors are available?

Alternatively, Furman proposes to use Woodward's two criteria to enable us to differentiate between causes and enabling factors: stability and specificity (Furman, 2016). A risk factor is more likely to be causal if it is stable in the sense that it remains unchanged against changing background conditions, and if it is specific to the effect at hand. So, for the probabilistic association between cannabis and psychosis to be stable, it would need to hold true for both men and women, and across different countries, for example. For it to be specific, smoking cannabis would only be associated with an increased risk of psychosis and not with, for instance, obsessive-compulsive disorder. Whilst Woodward's criteria are parsimonious and philosophically elegant, this thesis alone is by definition insufficient to assess stability (as I only have data from the EU-GEI study at my disposal) and specificity (as the EU-GEI study was only concerned with psychotic disorders)[3]. Therefore, in Chapter 8 I will return to the Bradford-Hill criteria in order to aid in judging the likely causality of the empirical findings presented in Chapters 6 and 7.

In the remainder of this thesis, I search for a mechanistic explanation for the increased risk of psychosis in ethnic minority groups, and assess its' probabilistic association. With 'mechanistic' I don't mean identifying the relevant neurobiological pathway, but rather identifying the process lying beneath the increased risk of psychosis in ethnic minorities, operating at the system level of the social environment.

5.3 Limitations of the epidemiological evidence

In Section 1.4, I addressed the epidemiology of the higher rates of psychotic disorders in ethnic minorities. I asserted this increased risk was neither an artefact of demography nor of racist diagnostic biases. I also asserted that ethnicity itself is not a reliable risk indicator, or stable cause, as the incidence of schizophrenia in Caribbean countries is far lower than the incidence in migrants originating from those countries in Western Europe, and this couldn't be explained by migrants inherently being more at risk of developing psychosis (the so-called 'unhealthy migrant' effect). Furthermore, rates of disorder for broad ethnic groups appear to differ across host countries. I also argued that whereas there is evidence that pre-

[3] Although specificity will be difficult to achieve in brain disorders more generally, due to overlapping phenotypes.

migratory exposure to stressors and migration itself are ingredients of the causal cake of psychosis, these can't explain the excess risk in second or subsequent generation. Post-migratory social circumstances are shared by generations of migrants, and as such form an important potential ingredient of the causal cake.

Despite this alleged importance of social circumstances of ethnic minority groups and evidence for it at a population or neighbourhood-level (Section 1.4), there is little epidemiological research exploring what it means to be in a minority: 'a small group of people within a community or country, differing from the main population in race, religion, language, or political persuasion' (Oxford University Press, 2016). Ethnicity is often treated as a fixed characteristic, or a category, rather than a complex construct (Section 1.3). This is perhaps not surprising. The academic discipline of epidemiology is concerned with establishing and quantifying the distribution and determinants of disease. As such, it inevitably utilises discrete, observable and measurable categories or variables, as these are most suited to this purpose. However, these might not be best suited to establish a plausible mechanism underpinning any observed probabilistic association. For instance, seeing ethnicity as a discreet variable doesn't do justice to the complex social nature in which someone's experience of their ethnicity arises. Ethnicity isn't just a categorical variable, but it reflects part of our social identities: how others see us, and how we see ourselves (Fenton & Charlsley, 2000; Kelly & Nazroo, 2008). In other words: my Dutch nationality is the same as that of my brother and sisters but because I am the only one who emigrated to England this same category has grown to have a different meaning to me.

This isn't an argument against use of categorical variables in epidemiology; it is merely intended to suggest that epidemiologists should account for their complexities and form new lines of inquiry on the basis of qualitative findings from other disciplines (Fenton & Charlsley, 2000). In other words: epidemiology alone has been insufficient to identify potential causes. Sociological approaches to mental health are plentiful, and have one element in common: they consider mental illness not just as a characteristic of individuals but also stemming from the various aspects of their social circumstances (Horwitz, 2017) and the social relationships between people (Fenton & Charlsley, 2000). As such, these sociological explanations aid understanding of the social patterning of disorder (Thoits, 2017). In the next Section, I will examine how the social sciences aid our understanding of these relationships and can inform our epidemiological hypotheses[4].

[4] An excellent example of how this can be done is Morgan's exploration of negative pathways to care for African-Caribbeans, where he uses a sociological model as a framework to interpret current and guide further research (Morgan et al., 2004).

5.4 Using social science to inform new hypotheses

I propose to examine the social circumstances of minorities as an ingredient of the causal cake explaining their excess risk of psychosis. Initially, this examination will focus on *ethnic* minority groups as they are the clearest identifiable population group with an epidemiologically well-established increased risk.

In short, in this Section I propose that minorities that are at a greater social and cultural distance from the majority population, will experience higher levels of what Michael Marmot names psychosocial disempowerment: lacking control over your life (Marmot, 2015). I argue that this, in turn, could increase one's risk of psychosis. In this Section, I first provide the background to this proposition and subsequently clarify the main components of it, and their relationship to each other. I will also address how this relates to existing theories of excess risk of psychosis in minorities.

5.4.1 Background

Akerlof and Kranton's theory of identity economics argues it is of fundamental importance to an individual's wellbeing (or utility) to be able to choose and express their identity (Akerlof & Kranton, 2011). People derive utility from doing what they think is in line with what they, and the social group they're part of, expect them to be doing. Everyone has an idea of who they are, and how they ought to behave in most circumstances. We infer this identity and those norms from both our past choices (Benabou & Tirole, 2011), and, importantly, from our social context (Akerlof & Kranton, 2011). Forming an identity is a fundamentally social process, and is largely facilitated by complex forms of behavioural inference, imitation and anticipation (Dijksterhuis, 2005). Our identities are fluid to a lesser or greater degree: they may change over time and across social circumstances. Everyone also exhibits multiple identities: I am a woman, an academic, Dutch, a vegetarian, a feminist, an atheist etc – but I have not always held all of these identities (I grew up a Protestant meat-eater) and they are not always equally important to me. Similarly, it is apparent that there can be no singular 'ethnic minority identity', as this importantly varies within and between ethnic groups (Nazroo & Karlsen, 2003). Amartya Sen argues strongly in favour of such a pluralist view of identity, and against viewing people on the basis of one singular identity such as Muslim or Dutch (Sen, 2006). Yet, often groups are delineated on the basis of one of their identities and this plurality isn't acknowledged, perhaps as a part of the 'unfamiliarity homogeneity effect': everything that is unencountered and unfamiliar becomes uniform (Malinowska, 2016). An example of this is a recent news report that Britain needs to repair its relationships with 'the Muslim community' (Townsend & Warsi, 2017), which fails to acknowledge that a religious identity is only one of many identities and within this group there are men, women, academics, carpenters, feminists, Labour-supporters and Conservatives who don't necessarily have anything other than their religion in common.

Being a fundamentally social process, forming, maintaining and expressing an identity is crucially dependent on others. This means that it can be harder for minority groups to experience a degree of autonomy in forming an identity. The borders of their identity, and which behavioural norms are

acceptable, are less autonomous and more externally determined by the majority population (Akerlof & Kranton, 2011; Hutnik, 1992). Furthermore, there is a risk of the majority viewing minorities as having a singular identity ('being Black'). By virtue of this one singular identity, they might never be able to meet certain ideals. For example, Cambridge attracted only 302 (3%) Black applicants in its 2016 admissions cycle with a much lower than average acceptance rate (12.9% compared to 26.4% in total) (University of Cambridge, 2017) suggesting there might be some racist assumptions about the identity of a 'Cambridge undergraduate', or that Black applicants are systematically disadvantaged in their access to education (long before they even apply to Cambridge). This is not just limited to ethnic minorities (or academia): for a long time, simply being a woman was sufficient to never be able to become an academic, as women were not admitted to universities. In Cambridge, for instance, the first college to admit women was founded in 1869 (Girton College, 2017) ,with the last college only admitting women in 1986 (Magdalene College, 2017).

Identity is important for people to derive utility from their actions. People who identify with what they do, or consider themselves 'insiders' derive utility simply from expressing their insider status (Akerlof & Kranton, 2011). An example of this would be a studious pupil deriving utility from doing well in a test. Making this insider status (implicitly) contingent on an unrelated identity, means that minorities remain outsiders, and as such are left to feel inadequate. If, implicitly, the prevailing idea of a 'good pupil' is of a White middle-class girl who is studious and obedient, the utility that boys, minorities, working-class students, and more rowdy pupils derive from doing equally well on the same test is lower because they can never meet the ideal, and this maintains the status quo. It also might make it harder for minorities to form an identity in the first place: it becomes harder to identify a stereotype to base their identity on if they are only presented with role models from the majority (Akerlof & Kranton, 2011), or if minority role models are portrayed negatively, such as the reinforcing of the idea that maths isn't for girls, which has been shown to impede performance (Shih, Pittinsky, & Ambady, 1999).

On a societal level, a shared sense of identity is crucial for the forming of what Paul Collier labels 'mutual regard'. This is a characteristic of social groups that is hard to define but which extends beyond mutual respect and is described as sympathy and a 'benign fellow-feeling' (Collier, 2013). This mutual regard is crucial for the trust that supports social cooperation, and a sense of empathy that supports (financial) redistribution (Collier, 2013), and is related to, for instance, increased willingness to contribute to public goods (Candelo, Croson, & Li, 2016). In Western societies, this mutual regard extends beyond the family or clan to a group of fellow citizens with whom there is least some sense of shared identity (Collier, 2013). The extent to which this is the case, and what this shared identity consists of is left open, but it ties in with the concept of 'bonding social capital' (see below). Unfortunately, it appears that, overall, an increase in (ethnic) diversity in societies undermines this sense of shared identity and, at least in the short term, leads to 'hunkering down': the withdrawal from public life (Putnam, 2007). As it is easier to form bonds with

individuals we have characteristics in common with (bonding social capital), than it is with individuals with whom we differ on an important aspect (bridging social capital) (Putnam, 2007), it is plausible that, in such a situation, minorities fall outside the majority population's category of 'fellow citizens', and that as such a sense of mutual regard doesn't extend to them (Collier, 2013). Collier doesn't address this explicitly, but presumably this applies to a different degree do different minorities in different host countries, depending on the social and historical context.

This 'us and them' thinking is particularly strong if we have a singular view of identity (Sen, 2006): we see someone first and foremost as 'Black' instead of a pupil applying to study at Cambridge. This means minorities are always to some extent regarded as outsiders (or, at best as good Black pupil, with no such need to specify skin colour for the White majority), and are more homogenous in the eye of the majority, and as such less free and able to form or express a plurality of identities, or to derive utility from their insider status. It isn't unreasonable to think that more culturally distant minorities or visibly different minorities are more easily excluded, as it is more difficult for the White majority to identify mutually shared identities. This social exclusion could plausibly lead to increased social distance, and increased levels of psychosocial disempowerment, which in turn increase the risk of adverse health outcomes, including psychotic disorders. In the next Section, I will introduce these terms in more detail, and then return to their relationship.

5.4.2 Concept clarification

5.4.2.1 Social distance

Social distance is a sociological construct describing the distance between two groups in society. It includes differences in social class, ethnicity, gender and other demographic characteristics (as opposed to physical or locational distance) (Karakayali, 2009). It is usually conceptualised in one of three ways: in terms of affectivity (how we feel towards another social group), as normative social distance (who is considered and insider or outsider) and as interactive social distance (how much members of different groups interact) (Karakayali, 2009). Since the conceptualisation of Bogardus' Social Distance Scale (Bogardus, 1926) it has been frequently studied through questionnaires or direct observations, but also more indirectly. An example of this is Wilson's 2010 book on the concentrated structural disadvantage faced by many inner-city African Americans and how this impacts on, and interacts, with culture. In this book, Wilson was evidently concerned with the complexities of social distance of African-Americans vis-à-vis White Americans, but the words social distance were not mentioned (Wilson, 2010).

Minority groups are evidently different or distant from the majority in their ethnicity, as that is the identity I choose to differentiate between groups, but their wider socio-economic circumstances contribute to this distance, too. It is too simple to consider, for example, socio-economic status or educational attainment as confounders; rather they are crucial elements of how ethnicity affects health (Williams et al., 2017). For this thesis, I conceptualise social distance as follows: the extent to which the socio-economic

circumstances of minority groups are different from those of the White majority. This is closest to normative social distance (as above), and exactly how this is operationalised will be described in Chapter 6.

5.4.2.2 Cultural distance

Cultural distance is the extent to which an individual's ideas, customs and social behaviour (their culture (Oxford University Press, 2016)) are different from that of the majority group. I am originally Dutch, which is relatively similar to British, so it has not been too difficult for me to settle here. However, had I moved to South Korea, I would have found it a great deal more difficult to acculturate to all the little intricacies of day-to-day interactions. A crucial element of this cultural distance is language: both linguistic difference and fluency in the majority language.

Linguistic distance is simply how different or similar two languages are to each other (Chiswick & Miller, 2004; Wichmann, Holman, Bakker, & Brown, 2010). Initially, I thought that linguistic distance would be a direct proxy for cultural distance (Collier, 2013), but I don't think the original evidence (Montalvo & Reynal-Querol, 2010) provides sufficient support for this. Montalvo and colleagues (2010) examined the extent of ethno-linguistic fragmentation (how many different languages or dialects are spoken) in a country, rather than the distance between these languages and as such don't address cultural distance specifically. Nonetheless, there does appear to be a crucial role for language in the forming and expressing of identity, and for social inclusion and exclusion (Clark, 2016). Majority language acquisition has been implicated in the forming of a social identity in migrants (Candelo et al., 2016; Koczan, 2016), and in decreased vulnerability to disease (Akiyama, 1996). Language skills were also included in Babiker's composition of cultural distance, which was predictive of study success in overseas students in Edinburgh (Babiker, Cox, & Miller, 1980). Furthermore, there is a large business literature on linguistic distance and its' close correlation with cultural values (West & Graham, 2004). I have therefore chosen to retain linguistic distance and fluency in the majority language as an, albeit indirect and imperfect, proxy for cultural distance.

5.4.2.3 Psychosocial disempowerment

I have taken Michael Marmot's definition of psychosocial disempowerment: a lack of control over your life (Marmot, 2015). Another way of looking at this might be to experience stressors you 'can't behaviourally avoid' (Fisher & Baum, 2010). This has been researched under a chronic social stress paradigm: more structural social factors such as inequalities in power, knowledge, influence and prestige have a negative impact on health outcomes (Link & Phelan, 1995). This chronic stress, or 'hassles of daily life' (Wheaton & Montazer, 2017) is almost as detrimental to wellbeing as the impact of major life events (Thoits, 2017). The concept of psychosocial disempowerment does not refer to the structural drivers of this chronic stress (or stressors), but it refers to the stress itself. As no direct stress measure was available from EU-GEI data, psychosocial disempowerment will be approximated using self-perceived discrimination (see Section 6.3).

Much as animals become stressed when they are further down the social ladder (Sapolsky, 2005), the immediate, daily, social environment has a profound effect on humans too (McEwen, 2012). Barbara Ehrenreich, reflecting on her journalistic experiment as a low-wage worker in the USA puts it as follows:

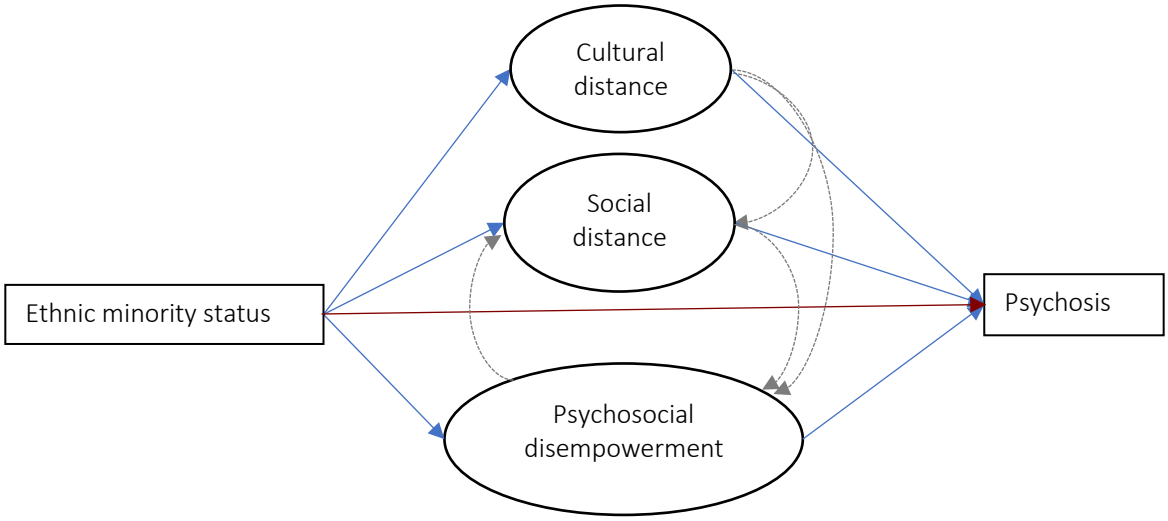
"*… As much as any other social animal, and more so than many, we depend for our self image on the humans immediately around us – to the point of altering our perceptions of the world so as to fit in with theirs. […] If you're made to feel unworthy enough, you may come to think that what you're paid is what you are actually worth*" (Ehrenreich, 2008).

5.4.3 A mechanistic explanation?

In the previous Sections I have laid out an argument to consider the social circumstances in which minority groups find themselves as a causal factor in explaining their higher rates of psychotic disorders. I have also introduced the concepts of social distance, cultural distance and psychosocial disempowerment. In this Section, I consolidate a suggested environmental pathway, which I will empirically test in subsequent Chapters.

Psychosocial disempowerment is associated with increased mortality and morbidity (Chandola, Brunner, & Marmot, 2006; Marmot, 2015; Siegrist & Marmot, 2004), and there is emerging evidence on the relationship between chronic stress and psychosis (Akdeniz, Tost, Streit, et al., 2014; Howes et al., 2017; Lederbogen et al., 2011). However, psychosocial disempowerment is not randomly distributed across society (Marmot, 2015), and this relates it to cultural and social distance. The broader organisation of society operates such that certain groups are more socio-economically disadvantaged than others, and consequently experience stress proliferation or high levels of psychosocial disempowerment. In sociology, this is studied under 'structural strain' theories (Thoits, 2017). This disadvantage tends to cluster in (ethnic) minority groups for a variety of reasons, not all of them explicitly racially motivated (Wilson, 2010), but partially because of a tendency to judge people on the basis of a singular identity which fosters 'us and them' thinking (Sen, 2006), and keeps those who are more culturally different from the majority from becoming insiders. This hypothesised relationship between all exposures is illustrated in Figure 5.3 below.

Figure 5.3: The hypothesised route from ethnic minority status to psychosis



The blue lines are hypothesised to explain the observed association represented by the red line. Ethnic minorities experience increased levels of cultural distance, which influences both their social distance and the higher levels of psychosocial disempowerment they experience, and the latter two also influence each other (grey dashed lines).

5.4.4 How this theory differs from existing theories

The inclusion of structural factors, such as lack of education, which put culturally distant groups at a greater social distance and increases their psychosocial disempowerment means this theory goes beyond 'black box' epidemiology by uniting broader social perspective with a way in which the environment gets under the skin (Hertzman & Boyce, 2010; McEwen, 2012), and could potentially increase risk for psychosis.

The broader social theory I propose therefore differs from cumulative social disadvantage (Morgan et al., 2008), which was later developed into a full socio-developmental model (Morgan, Charalambides, Hutchinson, & Murray, 2010). Cumulative social disadvantage is measured across six domains: education, employment, living arrangements, housing relationships and social networks (Morgan et al., 2008) and this concept isn't dissimilar to social distance. However, in this thesis social distance is only part of the cause, and not the main exposure of interest. Importantly, the socio-developmental model of psychosis in ethnic minorities, incorporating social distance, does not explicitly take into account cultural distance or psychosocial disempowerment, which I believe will be important determinants of minorities' excess psychosis risk.

Whereas the theory I propose has a similar focus to the social defeat theory (Selten & Cantor-Graae, 2005; Selten, Van Der Ven, Rutten, & Cantor-Graae, 2013), which posits that long-term exposure to the negative experience of being excluded from the majority group may lead to sensitisation of the mesolimbic dopamine system, and thereby increased risk for psychosis, it is different in the crucial premise of maintaining that individuals don't need to experience feeling defeated for the stressors to increase their

risk of psychosis. This thesis incorporates both structural causes and inequities and a measure of psychosocial disempowerment, whereas the social defeat hypothesis focusses exclusively on the subjective experience of being excluded. Whilst this is entirely speculative and outside my area of expertise, I suggest the neurobiological mechanism is more akin to the concept of 'allostatic load'. The hormones that in response to acute stress protect the body through a 'fight or flight' response and adaptation (allostasis), become overburdened through chronic stress (allostatic load), which causes changes in the brain and may lead to disorder (McEwen, 2012). This would mean that the stressors itself, rather than their cognitive interpretation of it, are crucial.

5.5 Summary and conclusions

In this Chapter, I have suggested that our understanding of causality was based on a mechanistic explanation underpinned by a probabilistic correlation, left room for multiple INUS conditions, and was able to include distant yet salient causes. I suggest the Bradford-Hill criteria are useful to assess likely causality of risk factors. I have argued that ethnicity is not a stable cause for the excess risk of psychosis in ethnic minority groups, as the risk for the same ethnic group varies depending on background conditions such as geographical location. I also argue that whereas pre-migration circumstances and migration itself are undeniably an ingredient of the causal cake, they are unable to account for the increased risk in later-generation migrants who have never themselves migrated.

I subsequently turned to the social sciences to explore the potential of the social circumstances of minorities being a plausible mechanism. I argued that particularly culturally distant minorities were at increased risk of social exclusion or falling outside the group of 'fellow citizens', leading to increased social distance. Such an outsider experience that could not be behaviourally avoided increased psychosocial disempowerment, which in turn was hypothesised to increase risk of psychosis.

This account of why minorities are at increased risk of developing psychosis is broad: it is the social and cultural distance and psychosocial disempowerment that are of causal relevance, not the simple fact of having a different ethnic background. This means that this same explanatory framework isn't necessarily limited to ethnic minorities: in Western societies, they currently occupy a marginalised position. However, so do other minority groups such as Muslims (a religious minority) and non-heterosexual and non-binary people. It is possible that such groups find themselves in a very similar position, if not now, then in the future.

Before I expand on such speculation however, it is prudent to test the plausibility of this explanatory framework in the minority group with the best-established excess psychosis risk: ethnic minorities. I will use data from the case-control arm of the EU-GEI study to do so, and as such am able to look at ethnic minorities across six countries. The empirical details of this are found in the subsequent Chapter (6), where I also explore excess risk in religious minorities. In Chapter 7, I will incorporate the more novel and perhaps

controversial element of genetic distance. This has been eluded to in the introduction, but will be explained in more detail in the background of Chapter 7.

Chapter 6 - Social and cultural distance as an explanation of higher rates of psychotic disorders in ethnic minority groups

6.1 Background

Ethnic minorities have a well-established excess risk of psychosis (Anderson et al., 2015; Bourque et al., 2011; Kirkbride, Hameed, Ioannidis, et al., 2017; Tortelli et al., 2014), the causes of which are not well-understood. In Chapters 1 and 5, I discussed that ethnicity itself is not a stable cause, as the risk for specific ethnic groups varies depending on background conditions (Bhugra et al., 1996; Fearon et al., 2006; Hickling, 1995; Selten et al., 2005; Veling et al., 2006). Whilst pre-migratory factors and migration itself are important, particularly for refugees (Anderson et al., 2015; Hollander et al., 2016), these can't explain excess risk in second and later generations. In the previous Chapter, I argued that falling outside the group of fellow citizens increases social distance and psychosocial disempowerment, and subsequent risk of psychosis. Those at a larger cultural distance are particularly at risk. In this Chapter, I will test whether cultural distance, social distance and psychosocial disempowerment explain the excess psychosis risk seen in some ethnic minority groups. I will also examine if this sociocultural distance model can explain any excess risk in religious minorities.

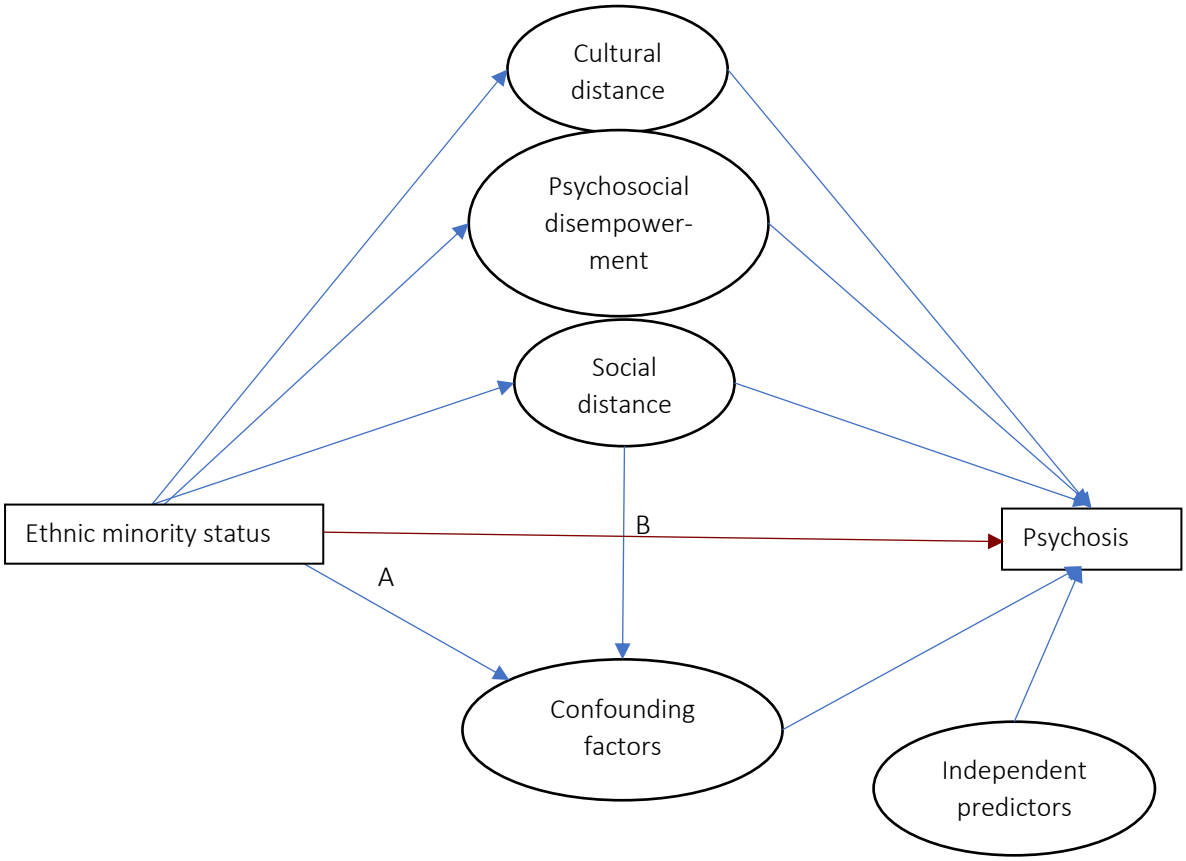
There is existing epidemiological evidence supporting the link between indicators of social distance and psychosis. The hypothesis that social isolation increases psychosis risk is long-standing (Faris & Dunham, 1939; Kohn & Clausen, 1955), and has been summarised in a recent systematic review (Gayer-Anderson & Morgan, 2013). Social isolation and educational attainment were elements of Morgan's concept of cumulative social disadvantage. Both were associated with an increased risk of psychosis and were more common in the Black-Caribbean group (Morgan et al., 2008), increasing their impact on psychosis risk in this group. In a Swedish registry study educational attainment was also shown to be associated with increased risk of being admitted to hospital for a psychotic disorder, after allowing for age, sex, immigrant status and socioeconomic status (hazard ratio (HR) for low education: 1.46, 95%CI: 1.36-1.56)(Leão, Sundquist, Johansson, Johansson, & Sundquist, 2005). Educational attainment possibly moderated psychosis risk in second-generation migrants (Leão et al., 2005). Higher IQ and other makers of cognitive ability have also been conceptualised as evidence of cognitive reserve, which may be protective against psychosis (Khandaker et al., 2011). There are other influences on educational attainment (notably, socio-economic status and parental education (Putnam, 2015)), suggesting it is both a suitable indicator of social distance and determinant of psychosis risk.

The majority of evidence for the importance of psychosocial disempowerment in psychosis risk is derived experimentally, under a social stress paradigm (Akdeniz, Tost, Streit, et al., 2014; Howes et al., 2017; Lederbogen et al., 2011; Mizrahi et al., 2012; Van Winkel, Stefanis, & Myin-Germeys, 2008), but it has some epidemiological support (Berg et al., 2011.; Karlsen, Nazroo, McKenzie, Bhui, & Weich, 2005; Veling

et al., 2007; Veling, Hoek, et al., 2008). For instance, a Dutch incidence study demonstrated a relationship between perceived discrimination and increased incidence at population-level (Veling et al., 2007), but not at individual level in a case-control sample (Veling, Hoek, et al., 2008). In a cross-sectional study investigating psychiatric illness in the community, experiences of interpersonal racism and perceptions of racism in society as whole were associated with an increased risk of psychosis, after controlling for age, sex and socioeconomic status (Karlsen et al., 2005).

The introduction of cultural distance is theoretical and innovative, and as such no studies have sought to operationalise and investigate this construct in the context of psychiatric epidemiology. A justification for using language distance and fluency in the majority language as proxies was given in Section 5.4. As detailed in Section 5.2, psychotic disorders are multi-causal and this PhD seeks to identify one element of the causal cake. Other elements of the cake include socioeconomic status (Marwaha & Johnson, 2004; Werner, Malaspina, & Rabinowitz, 2007), childhood trauma (Bendall et al., 2008; Matheson et al., 2012; Varese et al., 2012), cannabis use (Manrique-Garcia et al., 2012; Moore et al., 2007), younger age, male sex and their interaction (Häfner et al., 1993; Thorup et al., 2007)(Chapter 4), as well as increased paternal age at birth (Sipos et al., 2004; Zammit et al., 2003). These elements might confound the relationship between ethnic minority status and psychosis (line A in Figure 6.1), confound the relationship between social distance and psychosis (line B) or independently predict psychosis risk. In this Chapter, I seek to identify the unique contribution of social distance, cultural distance, and psychosocial disempowerment, and to determine if their combination is the missing ingredient in the causal cake of higher rates of psychosis in ethnic minorities.

Figure 6.1: Full model including covariates and independent predictors



The blue lines are hypothesised to explain the observed association represented by the red line. Line A represents covariates confounding the association between ethnic minority status and psychosis, and line B represents covariates confounding the association between social distance and psychosis.

As mentioned in Chapter 5, this model is not necessarily limited to ethnic minorities and I will also explore whether this model applies to religious minorities (those who follow a non-Christian religion). The yield from the literature on religion and psychosis is limited: a PubMed search using the terms ((psychosis[Title/Abstract] OR schizophrenia [Title/Abstract]) AND (religion [Title/Abstract])), yielded 130 articles. Most of these dealt with religious content of hallucinations and delusions (Gearing et al., 2011), or looked at the role of religion for social support (Sariah, Outwater, & Malima, 2014; Tabak & de Mamani, 2014). A multi-national study looking at spirituality and religion showed that religious involvement was higher in outpatients with schizophrenia or schizoaffective disorder compared with the general population (Mohr et al., 2012). Data from the Adult Psychiatric Morbidity Survey showed that non-Christians reported increased religious discrimination, and that those who perceived religious discrimination had an increased prevalence of common mental disorders (Jordanova, Crawford, McManus, Bebbington, & Brugha, 2015). Both of these studies were cross-sectional, the first study was small (n=276)(Mohr et al., 2012), and the

Adult Psychiatric Morbidity Survey looked at common mental disorders more broadly as opposed to psychosis specifically (Jordanova et al., 2015). This topic has not been well-researched, thus far.

6.2 Hypotheses

The overall aim of this Chapter was to examine if cultural and social distance and psychosocial disempowerment could explain the higher rates of psychotic disorders in some minority groups. In order to investigate this I tested the following hypotheses:

1. Ethnic minority status would be positively associated with increased odds of developing psychosis;
2. Ethnic minority status would be positively associated with increased markers of social distance, cultural distance and increased psychosocial disempowerment;
3. Social distance, cultural distance and psychosocial disempowerment would be associated with increased odds of developing psychosis;
4. The association between ethnic minority status and psychosis would be explained by cultural distance, social distance and psychosocial disempowerment.

I also aimed to explore the risk of developing a psychotic disorder in religious minority groups, and to test if this would also be accounted for by the sociocultural distance hypothesis. In order to investigate this, I tested the following hypotheses:

5. Religious minority status would be positively associated with increased odds of developing psychosis;
6. Religious minority status would be positively associated with increased cultural distance, markers of social distance and psychosocial disempowerment;
7. The relationship between religious minority status and psychosis would not be accounted for by ethnicity, but by the inclusion of markers of social distance, language distance, and discrimination.

6.3 Methods

The study design, sampling and recruitment strategy, case ascertainment and data collection and management of the EU-GEI study were detailed in Chapter 3. This Section contains only additional information relevant to this Chapter.

In addition to the exclusion criteria specified in Chapter 3, there were two additional exclusion criteria for this study: participants with insufficient data to compute cultural distance were excluded, as were cases with insufficient clinical data to arrive at an OPCRIT diagnosis.

6.3.1 Outcome, exposures, covariates

An overview of all variables included in the present study can be found in Table 6.1 below

Table 6.1: Outcomes, exposures and covariates of interest

Outcome	Exposures of interest	Covariates
All psychotic disorders (F20-F33)	Ethnicity	Age
	White majority	
Non-affective disorders (F20-25)	*Black*	Sex
	Mixed	
Affective disorders (F30-33)	*Asian*	Paternal age
	North African	
	Other	Childhood trauma
	White other	
		Lifetime cannabis use
	Indicators of social distance	
	Level of education	Paternal socioeconomic status
	Living arrangements	
	Relationship status	
	Cultural distance	
	Linguistic distance	
	Fluency	
	Psychosocial disempowerment	
	Discrimination	

6.3.1.1 Outcome

The outcomes of interest were all psychotic disorders (ICD10:F20-33), non-affective psychosis (F20-25) and affective disorders (F30-33).

6.3.1.2 Exposures

The exposure variables of interest were broad ethnic group, indicators of social distance (education, lifetime relationship status, living arrangements), cultural distance, psychosocial disempowerment and religious minority status.

Broad ethnic group

Self-ascribed ethnicity was recorded on the MRC Sociodemographic Schedule (see Appendix 3A), and could be derived from four indicators: a six-category broad ethnic group that was identical across all EU-GEI countries (part I, question 3), a country-specific ethnicity variable to permit ascription to pertinent ethnic groups in each setting (question 4), participants' country of birth (question 5), and parental country of birth (questions 7 and 8). Self-reported broad ethnic group was used to define six initial groups: White, Black, Mixed, Asian, North African, and the 'Other' category. Using the additional data I parsed the White group into the majority White group in each setting and people of White minority background (those born outside of that country, or those with two parents born outside of that country). In practice this included mainly White European migrants, as well as those from North America and Australasia.

Social distance

Social distance was approximated by educational attainment, relationship status and living arrangements. Educational attainment was measured using the MRC Sociodemographic Schedule in six levels of education (part II, question 8: school, no qualifications; school qualifications; tertiary; vocational; higher (undergraduate); higher (postgraduate)). Level of education was preferred over years of education as data was more complete (only 18 participants had missing data on level of education, and 448 on years of education). To minimise recall bias, I included binary lifetime relationship status (have you ever had a long-term relationship?) and living situation (since leaving your parents, have you ever lived with others?), as opposed to time-specific measures (both as recorded in the MRC Sociodemographic Schedule, part II questions 5 and 1).

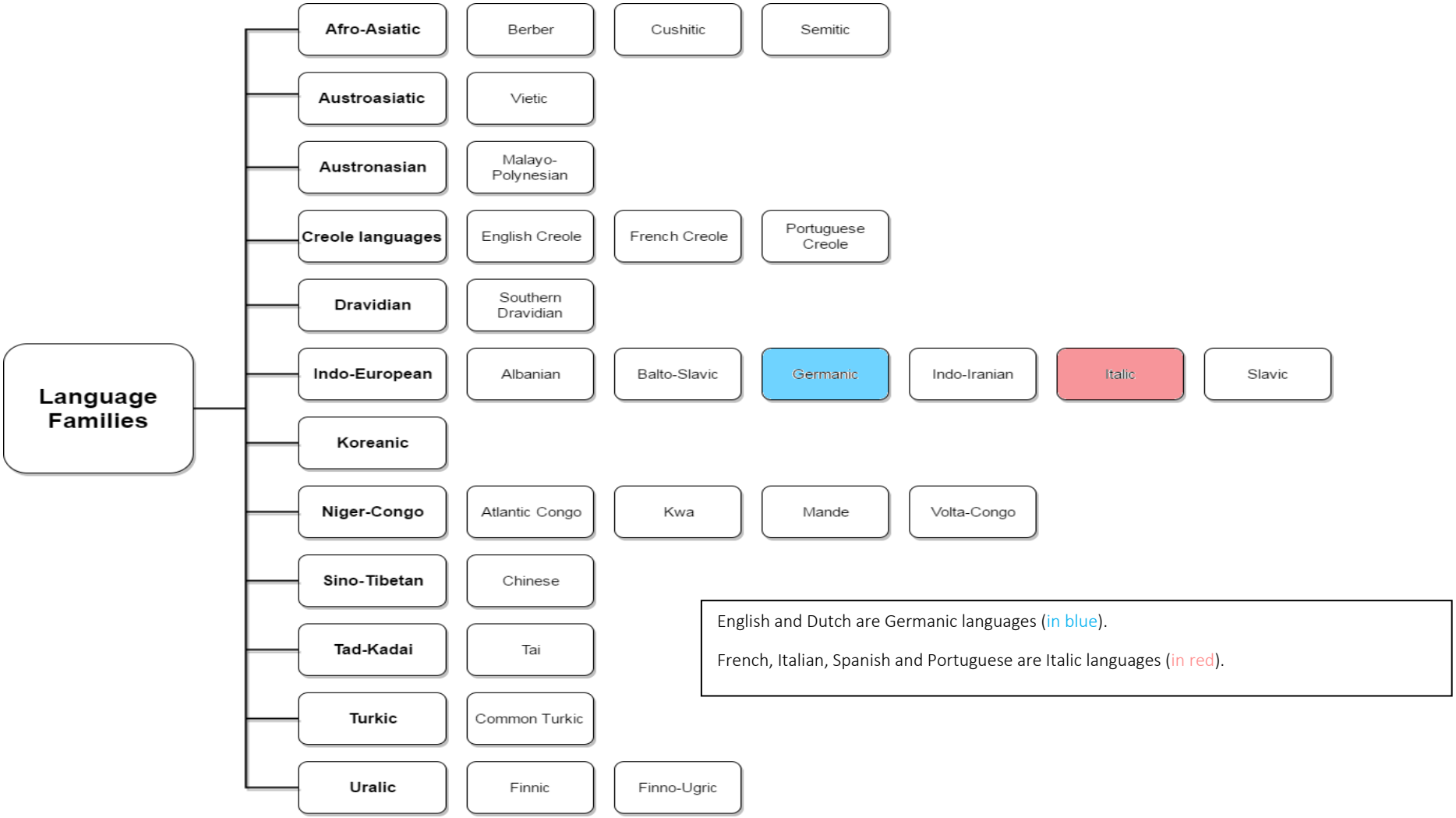
Cultural distance

Cultural distance was approximated using linguistic distance and fluency in the majority language. Linguistic distance was operationalised using information on participants first language as recorded in the MRC Sociodemographic Schedule (part I, question 9), and the majority language of their country of residence. An overview of all language branches represented in the sample can be found in Figure 6.2 below. If a participant's first language was the same as the official language of the country they were living in at the time of the study, they received a score of 0. If it was a language on the same branch in the same language family, they received a score of 1. An example of this would be a person whose native language is Spanish but was living in Italy (both Italic languages in the Indo-European language family). A first language in a different branch within the same family would receive a score of 2, such as a Dutch person living in Italy (a Germanic and Italic language in the Indo-European language family). A language in a different family altogether would receive a score of 3. An example of this is a native speaker of Chinese living in Italy (a Sino-Tibetan and Indo-European language). All majority languages were in the Indo-European family, either in the Germanic branch (English, Dutch), or in the Italic branch (Spanish, Italian, French, Portuguese) (see Figure 6.2 below).

Fluency in the majority language was self-rated on a 10-point scale (part II, question 14), with a higher score indicating a higher fluency. Scores were categorised into low (1-3), moderate (4-6), high (7-9) and very high (10) fluency. To compute cultural distance, those with very high fluency were given a score of 0, a high fluency yielded a score of 1, moderate of 2 and low of 3. This score was then added to participants' linguistic distance to arrive at total cultural distance.

Because of the skewness in cultural distance scores (91.5% of controls and 84.0% of cases received a total cultural distance score of 0), this was operationalised and analysed as a binary variable (0 and 1-6).

Figure 6.2: Language families and branches included in the EU-GEI study.*



* Please note this is not a complete language tree. It only includes languages spoken by EU-GEI participants. Language families and branched are ordered alphabetically.

Psychosocial disempowerment

Psychosocial disempowerment was approximated using a 12-item discrimination questionnaire included in the MRC Sociodemographic Schedule (see pages 259 and 260). Participants were asked if they had ever been unfairly treated [sic] on the following domains: fired; not hired for a job; denied promotion; stopped, questioned or threatened by the police; treated by the court system; discouraged from continuing education; prevented from buying or renting a flat or house; treated by neighbours or family; denied a loan or preferable mortgage rate; received worse service than others; treated when getting medical care; or treated when using public transport. For each of these possible occurrences, participants were asked the number of times they experienced them, and the age at which they first experienced them. They were also asked for the reason why they experienced this discrimination, with the answering options of gender, race/ethnicity, religion, mental illness, sexuality, age, or other. I used the total number of items participants reported ever having felt discriminated against (0-12), regardless of the number of events per item.

Religious minority status

Religious minority status was approximated using participants' self-reported religious affiliation as recorded in the MRC Sociodemographic Schedule (part II, question 12: none, Christian, Jewish, Muslim and other), reclassified into majority (none, Christian) and minority (Jewish, Muslim, other) status. A number of participants indicated in the 'other' Section that they were of a particular Christian denomination or atheist and have been re-categorised accordingly.

6.3.1.3 Covariates

Age, sex, their interaction, paternal age, lifetime cannabis use, childhood trauma and paternal socioeconomic status (definitions below) were included as either confounders or independent predictors. Sex was binary and self-reported via the MRC Sociodemographic Schedule (part I, question 1). Age was continuous, and derived on date of assessment (as listed on the MRC Sociodemographic Schedule, question 2). Paternal age was continuous, reported by participants and retrieved from the MRC Sociodemographic Schedule (question 14).

Childhood trauma was operationalised as the total score on the childhood trauma questionnaire. This questionnaire was originally developed by Bernstein and colleagues (Bernstein et al., 2003), and consisted of 25 items, each scored on a 5-point Likert scale. The total score ranged from 25 to 125 (a higher score indicating higher levels of childhood trauma). Cannabis use was operationalised using the 'Have you ever smoked cannabis?' question included in the Cannabis Experience Questionnaire. I decided that binary lifetime cannabis use was an appropriate predictor for these analyses: it is a good predictor of psychosis risk (Moore et al., 2007), and data were readily and relatively completely available. Paternal socioeconomic status was preferred over participants' socioeconomic status due to the risk of reverse causality associated

with the latter, and was based on fathers' highest occupation as listed on the MRC Sociodemographic Schedule (part I, question 12). This was originally categorised according to the European Socio-economic Classification (Harrison & Rose, 2006). For the purposes of this study, this was categorised into six categories: professional (both higher grade and lower grade), intermediate (including intermediate occupations, small employers and self-employed), lower (lower supervisory, technician, services, sales, clerical and technical occupations), routine, never worked (including long-term unemployed) and not classified (including students).

6.3.2 Missing data

I investigated missing data in cooperation with local researchers in all catchment areas. This included identifying apparent outliers (such as a paternal age of 72), and inconsistencies (such as a negative response to an item on the discrimination questionnaire, but at the same time indicating this occurred for the first time at age 13). A degree of missing data remained; in order to avoid dropping observations with missing data, which would result in a loss of precision (power), and could potentially yield biased results, I used multiple imputation. I imputed date using multiple imputation by fully conditional specification using chained equations (Little & Rubin, 2002; Sterne et al., 2009). Each imputation model (n=25) used all variables in the analyses, as well as auxiliary variables (see below), to increase the likelihood of satisfying the assumption that data was Missing At Random (MAR) (Sullivan et al., 2016). I used the user-written *ice* command as Stata's in-build imputation mechanism is unable to accommodate the presence of multiple categorical variables, which presented a problem in this dataset.

I imputed missing exposures and covariates for participants with complete outcome and language data. The following variables had missing values and were imputed: age, sex, ethnicity, discrimination, level of education, relationship status, living arrangements, cannabis use, childhood trauma, paternal age and paternal socioeconomic status. Case-control status, diagnosis (non-affective, affective, psychosis NOS), Wechsler Adult Intelligence Scale (WAIS) score, neighbourhood trust (measured through the Social Environment Assessment Tool), linguistic distance and fluency were used as auxiliary variables: they were not themselves imputed but used to inform the expected values for the imputed variables.

Analyses were conducted post-imputation, combining estimates across the 25 imputed data sets using Rubin's rule (White, Royston, & Wood, 2011).

6.3.3 Statistical model and analyses

After describing recruitment and missing data, I plotted the distribution of exposures and covariates by case-control status and used appropriate parametric and non-parametric tests to assess difference between cases and controls. I used multinomial regression to test the relationship between ethnicity (using the White majority group as the reference category) and other exposures and covariates, and estimated correlations between remaining exposures and covariates using polychoric (for two binary

variables or a binary and a categorical variable), polyserial (for a categorical and a continuous variable) and Pearson's (for two continuous variables) correlations.

I then used a univariable logistic regression model to assess the associations between the outcomes of interest and exposures and covariates. I subsequently followed the methodology originally developed for analysis of the AESOP data, of progressively including predictors in a multivariable model in order to explain crude excess odds[5]. Order of model inclusion was based on univariable Pseudo-R^2 of the exposures. I built a multivariable regression model for each outcome of interest (all psychotic disorders, non-affective psychoses and affective psychoses) to mutually control for the variables in each model as follows:

- Univariable model with outcome variable and ethnicity
- Model A: multivariable model with outcome variable, ethnicity and *a priori* covariates (age, sex, their interaction, paternal age, childhood trauma, cannabis use, and family history of psychosis)
- Model B: multivariable model with outcome variable, ethnicity, covariates and markers of social distance (education level, relationship status, living arrangements)
- Model C: multivariable model with outcome variable, ethnicity, covariates, markers of social distance and cultural distance.
- Model D: multivariable model with outcome variable, ethnicity, covariates, markers of social distance, cultural distance and self-perceived discrimination.

Multilevel random effects were introduced to allow for the clustering of data at catchment area-level. Descriptive analyses were carried out on the complete case dataset, and regression analyses were based on the datasets with imputed values.

Computationally, it was impossible to assess model fit using likelihood ratio-testing, or to run post-estimation commands to estimate Akaike's Information Criterion (AIC) in multiple-imputed datasets in Stata, so I computed and used Nagelkerke's Pseudo-R^2 instead (Nagelkerke, 1991). This measure of model fit is akin to R^2 in linear regression analyses, designed for logistic regression, and can be interpreted as the proportion of variance explained by the model (more precisely 1- R^2 is the proportion of variance unexplained)(Nagelkerke, 1991).

The association between religious minority status and psychosis and the outcomes of interest was also assessed using models A-D above, with religious minority status taking the place of ethnicity and ethnicity introduced in model A as a confounder. I conducted *post-hoc* analyses to examine whether ethnicity or religion was a stronger predictor following adjustment for covariates and remaining exposures.

[5] This was included in the protocol, although it not in any publications resulting from the AESOP study.

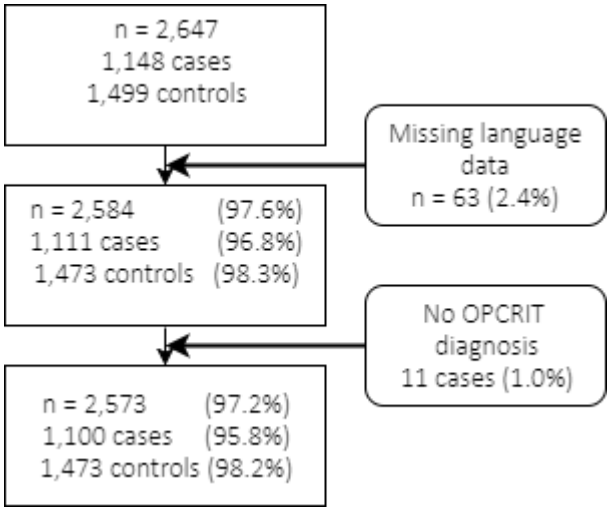
By means of sensitivity analyses, the models (A-D above) were also tested on non-imputed observations (participants) with complete data only. It was anticipated this would more readily 'explain' the increased risk of psychotic disorders in ethnic minorities: an effect of the study becoming underpowered to detect true differences.

6.4 Results

6.4.1 Sample description

As described in Section 3.7, the EU-GEI study recruited a total of 1,148 cases and 1,499 controls. Following exclusion of those with missing language data (37 cases (3.2%) and 26 controls (1.7%)) and insufficient diagnostic data (11 cases; 1.0%), the final sample size was 2,573 (1,100 cases (95.8% of total recruited) and 1,473 controls (98.2% of total recruited)) (Figure 6.3 below). Of the 1,100 cases, 767 (69.7%) had a non-affective disorder, and 305 (27.7%) an affective disorder. Twenty-eight cases (2.6%) had diagnosis of psychosis NOS and were not analysed separately (but were retained for the main analyses).

Figure 6.3: Flowchart of case retention



6.4.2 Missing data

6.4.2.1 Exclusion criteria

Cases were more likely to be missing language data than controls (χ^2: 6.2, p=0.01), as were ethnic minorities (χ^2: 23.9, p<0.01). No difference was found by age (Mann-Whitney U-test: 1.3, p=0.22) and sex (χ^2: 0.0, p=0.90). Those cases without an OPCRIT diagnosis were no different in age (Mann-Whitney U-test: -0.6, p=0.57), sex (χ^2: 1.10, p=0.29) and minority status (χ^2: 3.0, p=0.08) than those with an OPCRIT diagnosis.

6.4.2.2 Final sample for multiple imputation

Of the final sample (n=2,573), one subject was missing data on sex, two subjects (0.1%) were missing data on age and ethnicity respectively, 192 subjects (7.5%) on paternal age, 43 (1.7%) on cannabis use, 129

(5.0%) on childhood trauma, 134 (5.2%) on paternal socioeconomic status, 18 (0.7%) on level of education, 14 (0.5%) on relationship status, 199 (7.7%) on living arrangements, and 348 (13.5%) on self-reported discrimination (Table 6.2 below).

Logistic regression confirmed that missingness was not associated with case-control status for sex, age and ethnicity. However, cases were more likely to have missing data on paternal age (OR: 3.99, 95%CI: 2.87-5.54), cannabis use (OR: 2.29, 95%CI: 1.23-4.27), childhood trauma (2.72, 95%CI: 1.25-10.63), paternal socioeconomic status (OR: 1.65, 95%CI: 1.16-2.33), social distance (OR: 1.74, 95%CI: 1.31-2.31) and self-perceived discrimination (OR: 1.26, 95% CI: 1.01-1.59). Missingness in exposures and covariates was not associated with sex, age or ethnicity broadly. Exceptions were associations between younger age and missing data on paternal age (OR: 0.98, 95%CI: 0.97-0.99), paternal socioeconomic status (OR: 0.98, 95%CI: 0.97-1.00) and living arrangements (OR: 0.96, 95%CI: 0.95-0.97), as well as ethnic minorities having higher odds of missing data on childhood trauma (OR: 1.64, 95%CI: 1.15-2.35) and paternal age (OR: 2.80, 95%CI: 2.07-3.77) and lower odds of missing data on discrimination (OR: 0.61, 95%CI: 0.47-0.78).

Table 6.2: Distribution of covariates and exposures, by case-control status

Variable		Controls *n (%)*	Cases *n (%)*	χ^2; p-value/ MWU[1]; p-value
Age	Median (IQR)	33 (26-47)	28 (22-37)	MWU: 9.9;
	Missing	1 (0.1%)	-	p<0.01
Sex	Men	696 (47.3)	678 (61.6)	
	Women	775 (52.6)	422 (38.4)	χ^2: 51.9;
	Missing	2 (0.1)	-	p<0.01
Paternal age	Median (IQR)	31 (27-35)	31 (27-36)	MWU: -0.7;
	Missing (%)	52 (3.5%)	140 (12.7%)	p=0.48
Childhood trauma	Median (IQR)	35 (31-41)	41 (35-51)	MWU: -15.8;
	Missing (%)	23 (1.6%)	106 (9.6%)	p<0.01
Cannabis use	Yes	691 (46.9)	681 (61.9)	
	No	766 (52.0)	392 (35.6)	χ^2: 64.1;
	Missing	16 (1.1)	27 (2.5)	p<0.01
Ethnicity	White majority	1,055 (71.6)	630 (57.3)	
	Black	119 (8.1)	182 (16.6)	
	Mixed	117 (7.9)	116 (10.6)	
	Asian	33 (2.2)	27 (2.5)	
	North African	23 (1.6)	46 (4.2)	
	Other	20 (1.4)	30 (2.7)	
	White other	104 (7.1)	69 (6.3	χ^2: 86.0;
	Missing	2 (1.4)	-	p<0.01
Paternal socioeconomic status	Professional	394 (26.8)	207 (18.8)	
	Intermediate	323 (21.9)	209 (19.0)	
	Lower	371 (25.9)	296 (26.9)	
	Routine	294 (20.0)	239 (21.7)	
	Never worked	7 (0.5)	25 (2.3)	
	Not classified	23 (1.6)	51 (4.6)	χ^2: 58.1;
	Missing	61 (4.1)	73 (6.6)	p<0.01
Level of education	Postgraduate	207 (14.0)	50 (4.6)	
	Undergraduate	340 (23.1)	126 (11.5)	
	Vocational	233 (15.8)	191 (17.4)	
	Tertiary	427 (29.0)	259 (23.6)	
	School qualifications	190 (12.9)	285 (25.9)	
	School, no qualifications	69 (4.7)	178 (16.2)	χ^2: 256.6;
	Missing	7 (0.5)	11 (1.0)	p<0.01
Relationship status	Yes	1,308 (88.8)	735 (66.8)	
	No	163 (11.1)	353 (32.1)	χ^2: 177.3;
	Missing	2 (0.1)	12 (1.1)	p<0.01
Living arrangements	Yes	1,125 (76.4)	646 (58.7)	
	No	259 (17.6)	344 (31.3)	χ^2:78.3;
	Missing	89 (6.0)	110 (10.0)	p<0.01
Linguistic distance	0 – Majority language	1,349 (91.6)	924 (84.0)	
	1 – Same branch	34 (2.3)	28 (2.6)	
	2 – Same family	38 (2.6)	51 (4.6)	χ^2: 42.4
	3 – Different family	52 (3.5)	97 (8.8)	p<0.01
Fluency	0 – Low	5 (0.3)	14 (2.7)	
	1 – Moderate	28 (1.9)	45 (4.1)	
	2 – High	68 (4.6)	85 (7.7)	χ^2: 31.0;
	3 – Very high	1,372 (93.1)	956 (86.9)	p<0.01
Discrimination	Median (IQR)	0 (0-1)	0 (0-1)	MWU: -4.5;
	Missing	182 (12.4%)	166 (15.1%)	p<0.01

[1] MWU: Mann-Whitney U test used to test for differences in median value between cases and controls.

6.4.3 Distribution of covariates and exposures

6.4.3.1 Distribution of covariates

Controls had a median age of 33 (IQR: 26-47), whereas cases were significantly younger with a median age of 28 (IQR: 22-37) (Mann-Whitney U-test: 9.9, p>0.01, Table 6.2 above). There was no difference in age between diagnostic categories (Kruskal-Wallis χ^2: 2.2 2df, p=0.33) (Table 6.3 below). Cases were significantly more likely to be male than controls (χ^2: 51.9, p<0.01), with 61.6% (n=678) of cases being male, compared with 47.3% (n=696) of controls (Table 6.2). Those with non-affective disorders were more likely to be male than those with affective disorders and psychosis NOS (χ^2: 23.5, p<0.01) (Table 6.3).

There was no difference in paternal age between cases (median: 31, IQR: 27-36) and controls (median 31, IQR: 27-35; Mann-Whitney U-test -0.7, p=0.41), or between diagnostic categories (Kruskal-Wallis χ^2: 4.8 2df, p=0.10). Cases were more likely than controls to have reported lifetime cannabis use (61.9% vs 46.9%, (χ^2: 64.1, p <0.01)); this was higher in those with a non-affective disorder compared with those with an affective disorder or psychosis NOS (χ^2: 8.9, p<0.01) (Table 6.3). Cases reported higher levels of childhood trauma, with a median score of 41 (IQR: 35-51) on the Childhood Trauma Questionnaire, compared with 35 (IQR: 31-41) for controls (Mann-Whitney U-test: -15.8, p<0.01); this did not differ between diagnostic categories (Kruskal-Wallis χ^2: 0.8, 2df, p=0.66) (Table 6.3). Controls' paternal socioeconomic status was also different to cases' (χ^2: 58.1, p>0.01) with controls' fathers more likely to belong to the professional or intermediate category, and less likely to never have worked or not be classifiable (Table 6.2); there were differences by diagnostic category (χ^2: 111.6, p<0.01), but no clear patterns could be observed (Table 6.3).

6.4.3.2 Distribution of exposures

Ethnicity

The ethnic composition of the case-group was significantly different from the control-group (χ^2: 86.0, p<0.01), with cases more likely to be Black, Mixed, North African or 'Other', and less likely to have belonged to the White majority population or to the White non-native group (Table 6.2). Ethnic composition differed by diagnostic category (χ^2: 39.8, p<0.01) with those diagnosed psychosis NOS more likely to belong to the White majority than those with an affective or non-affective disorder, and those with an affective disorder more likely to be from a Mixed ethnic background (Table 6.3).

Social distance

Controls had a higher level of education than cases (χ^2: 256.6, p<0.01) (Table 6.2). Those with psychosis NOS had a higher level of education compared with those with an affective or non-affective disorder (χ^2: 20.8, p=0.02) (Table 6.3). Almost 89% of controls had ever had a relationship (n=1,308), and this percentage was significantly lower in cases at 66.8 (n=735, χ^2: 17.3, p<0.01) (Table 6.2). Those with a non-affective disorder were less likely to have been in a relationship than those with an affective disorder or

Table 6.3: Distribution of exposures and covariates, by diagnostic category

Variable		Non-affective *n* (%)	Affective *n (%)*	Psychosis NOS *n (%)*	χ^2; p-value/ KW[1] χ^2; p-value
Age	Median (IQR)	28 (22-36)	30 (22-41)	29.5 (22.5-39)	KW χ^2: 2.2,
	Missing	0 (0.0)	0 (0.0)	0 (0.0)	2df; p<0.01
Sex	Men	507 (66.1)	153 (50.2)	18 (64.3)	
	Women	260 (33.9)	152 (49.8)	10 (35.7)	χ^2: 23.5;
	Missing	0 (0.0)	0 (0.0)	0 (0.0)	p<0.01
Paternal age	Median (IQR)	31 (36-37)	30.5 (27-36)	33 (30-40)	KW χ^2: 4.8,
	Missing (%)	105 (13.6)	33 (10.8)	2 (7.1)	2df; p=0.08
Childhood trauma	Median (IQR)	42.5 (35-50)	41 (35-51)	37 (34-46)	KW χ^2: 0.8,
	Missing (%)	84 (11.0)	24 (6.9)	1 (3.6)	2df p=0.66
Cannabis use	Yes	496 (64.7)	172 (56.4)	13 (46.4)	
	No	252 (32.9)	128 (42.0)	12 (42.9)	χ^2: 92.2;
	Missing	19 (2.5)	5 (1.6)	3 (10.7)	p<0.01
Ethnicity	White majority	434 (56.6)	174 (57.1)	22 (78.4)	
	Black	134 (17.5)	47 (15.4)	1 (3.6)	
	Mixed	63 (8.2)	53 (17.4)	0 (0.0)	
	Asian	22 (2.9)	5 (1.6)	0 (0.0)	
	North African	39 (5.1)	7 (2.3)	0 (0.0)	
	Other	24 (3.1)	5 (1.6)	1 (3.6)	
	White other	51 (6.7)	14 (4.6)	4 (14.3)	χ^2: 135.3;
	Missing	0 (0.0	0(0.0)	0 (0.0)	p<0.01
Paternal socioeconomic status	Professional	143 (18.6)	59 (19.4)	5 (17.9)	
	Intermediate	141 (18.4)	66 (21.6)	2 (7.1)	
	Lower	215 (28.0)	73 (23.9)	8 (28.6)	
	Routine	153 (20.0)	81 (26.6)	5 (17.9)	
	Never worked	23 (3.0)	2 (0.7)	0 (0.0)	
	Not classified	42 (5.5)	8 (2.6)	1 (3.6)	χ^2: 111.6;
	Missing	50 (6.5)	16 (5.3)	7 (25.0)	p<0.01
Level of education	Postgraduate	35 (4.6)	11 (3.6)	4 (14.3)	
	Undergraduate	88 (11.6)	36 (11.8)	2 (7.1)	
	Vocational	145 (18.9)	41 (13.4)	5 (17.9)	
	Tertiary	164 (21.4)	84 (27.5)	11 (39.3)	
	School qualifications	201 (26.2)	79 (25.9)	5 (17.9)	
	School, no qualifications	124 (16.7)	53 (17.4)	1 (3.6)	χ^2: 281.3;
	Missing	10 (1.3)	1 (0.3)	0 (0.0)	p<0.01
Relationship status	Yes	478 (62.3)	234 (76.7)	23 (82.1)	
	No	248 (36.3)	70 (23.0)	5 (17.9)	χ^2: 221.9;
	Missing	11 (1.4)	1 (0.3)	0 (0.0	p<0.01
Living arrangements	Yes	427 (55.7)	207 (67.9)	12 (42.9)	
	No	262 (34.2)	71 (23.3)	11 (39.3)	χ^2:112.0;
	Missing	78 (10.2)	27 (8.9)	5 (17.9)	p<0.01
Linguistic distance	0 – Majority language	629 (82.0)	268 (87.9)	27 (96.4)	
	1 – Same branch	22 (2.3)	5 (1.6)	1 (3.6)	
	2 – Same family	40 (5.2)	11 (3.6)	0 (0.0)	χ^2: 42.4
	3 – Different family	76 (9.9)	21 (6.7)	0 (0.0)	p<0.01
Fluency	0 – Low	2 (1.6)	2 (0.7)	0 (0.0)	
	1 – Moderate	33 (4.3)	12 (3.9)	0 (0.0)	
	2 – High	70 (9.1)	14 (5.6)	1 (3.6)	χ^2: 31.0;
	3 – Very high	652 (82.0)	277 (90.8)	27 (96.4)	p<0.01
Discrimination	Median (IQR)	0 (0-1)	0 (0-1)	0 (0-1)	KW χ^2: 1.4,
	Missing	121 (15.8)	38 (12.5)	7 (25.0)	2df; p=0.50

[1]KW: Kruskal-Wallis test used to test for differences in median value between diagnostic outcomes.

psychosis NOS (χ^2: 21.5, p<0.01). Sixty-seven percent (n=1,125) of controls has ever lived with others since leaving their parental home, compared with only 59% (n=646) of cases (χ^2: 91.6, p<0.01). Those with an

affective disorder were more likely to have ever lived with someone other than their parents compared with those with a non-affective disorder and psychosis NOS (χ^2: 15.4, p<0.01).

Cultural distance

Overall, controls had a lower linguistic distance than cases. Although the majority (n=2,273; 88.3%) of all participants had a language distance of 0, cases had significantly higher language distance compared with controls (Table 6.2, χ^2: 42.4, p<0.01). Cases were less likely to report a very high fluency compared with controls (χ^2: 31.0, p<0.01). When combining this into the binary cultural distance measure, cases were more likely to have a degree of cultural distance compared with controls (n=176, 16.0% compared with 8.5% (n=125), χ^2: 34.4, p<0.01); this was higher in those with a non-affective disorder (n=138; 18.0%) than in those with an affective disorder (n=37; 12.1%) and those with psychosis NOS (n=1; 3.6%) (χ^2: 8.9, p=0.01).

Psychosocial disempowerment

Despite an identical median (0) and IQR (0-1), there was a significant difference in levels of self-perceived discrimination between cases and controls (Mann-Whitney U-test: -4.5, p <0.01) (Table 6.2), with cases reporting a higher level of discrimination (cases had a higher 90th percentile at 3, compared with 2 in controls). There was no difference in level of discrimination between diagnostic categories (Kruskal-Wallis χ^2: 1.13 2df, p=0.50) (Table 6.3).

6.4.4 Associations between exposures and covariates

6.4.4.1 Associations between ethnicity and remaining exposures and covariates

As per hypothesis 2, I investigated the association between ethnicity and cultural and social distance and psychosocial disempowerment using univariable multinomial logistic regression with ethnicity as the outcome variable on the imputed dataset.

Ethnicity was associated with all covariates to at least some extent (Table 6.4 below). Most ethnic minority groups had higher childhood trauma than the White majority group. The Black, Mixed, North African and 'Other' group tended to have a lower paternal socioeconomic status, whereas the Asian and White other group had a higher paternal socioeconomic status. Younger age and male sex were only associated with some ethnic groups, and there was no evidence ethnic minority groups had a higher lifetime cannabis use (Table 6.4).

Findings for indicators of social distance were mixed. The Black group had higher odds of having a vocational education (OR: 2.75, 95%CI: 1.49-5.08), leaving after obtaining school qualifications (OR: 2.51, 95%CI: 1.35-4.66) or without any qualifications (OR: 5.48, 95%: 2.89-10.40), and the Mixed group had higher odds of having a tertiary education (OR: 3.69, 95%CI: 1.64-8.22) or leaving with (OR: 5.31, 95%CI: 2.36-11.93) or without (OR: 11.50, 95%CI: 5.04-26.26) school qualifications (Table 6.4).The Asian and

White other group had a higher level of education. The Asian group had lower odds of having an undergraduate education (OR: 0.60, 95%CI: 0.28-1.30), a vocational degree (OR: 0.39, 95%CI: 0.16-0.98) or a tertiary education (OR: 0.30, 95%IC: 0.13-0.69), and the White other group was also less likely to leave with only school qualifications (OR: 0.51, 95%CI: 0.29-0.89). The Black (OR: 1.83, 95%CI: 1.31-2.57), Mixed (OR: 1.53, 95%CI: 1.05-2.23) and White other groups (OR: 2.72, 95%CI: 1.61-4.57) were more likely to ever have been in a relationship. The Black (OR: 0.48, 95%CI: 0.35-0.66), North African (OR: 0.47, 95%CI: 0.25-0.87) and 'Other' groups (0.37, 95%CI: 0.18-0.73) were less likely than their White majority counterparts to have ever lived with someone other than their parents.

All ethnic groups had a higher cultural distance, with ORs highest in the North African (OR: 12.10, 95%CI: 7.61-19.23) and the Asian (OR: 9.95, 95%CI: 6.23-15.82) groups. The Black (OR: 1.61, 95%CI: 1.47-1.77), North African (OR: 1.53, 95%CI: 1.30-1.80), Other (OR: 1.69, 95%CI: 1.43-1.99) and White other (OR: 1.32, 95%CI: 1.15-1.51) groups reported higher levels of discrimination.

Table 6.4: Associations between ethnicity and remaining exposures (indicators of social distance, cultural distance and discrimination) and covariates

Ethnic group	Black	Mixed	Asian	North African	Other	White other
Variable	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
Age	**0.97 (0.96-0.98)**	**0.97 (0.96-0.98)**	0.98 (0.96-1.00)	0.99 (0.97-1.01)	0.99 (0.95-1.00)	1.00 (0.98-1.01)
Sex						
Male	1.07 (0.84-1.38)	0.93 (0.70-1.22)	0.95 (0.57-1.58)	**1.89 (1.13-3.16)**	1.33 (0.75-2.36)	0.92 (0.68-1.26)
Female	Reference	Reference	Reference	Reference	Reference	Reference
Paternal age	1.01 (0.99-1.03)	0.99 (0.97-1.01)	1.01 (0.98-1.05)	**1.08 (1.04-1.11)**	1.01 (0.98-1.06)	1.00 (0.97-1.02)
Childhood trauma	**1.05 (1.04-1.06)**	**1.03 (1.02-1.04)**	1.02 (1.00-1.05)	**1.03 (1.01-1.05)**	**1.04 (1.02-1.06)**	**1.02 (1.01-1.03)**
Cannabis use						
Yes	0.94 (0.7-1.20)	**0.60 (0.45-0.79)**	0.76 (0.45-1.27)	0.95 (0.59-1.56)	0.95 (0.53-1.67)	1.33 0.96-1.84)
No	Reference	Reference	Reference	Reference	Reference	Reference
Paternal SES						
Professional	Reference	Reference	Reference	Reference	Reference	Reference
Intermediate	0.84 (0.58-1.22)	**2.45 (1.43-4.23)**	0.89 (0.49-1.63)	**4.40 (1.61-11.98)**	1.46 (0.60-3.59)	0.71 (0.45-1.12)
Lower	**0.63 (0.44-0.91)**	**1.72 (1.00-2.96)**	**0.21 (0.08-0.51)**	**4.32 (1.63-11.45)**	1.45 (0.62-3.36)	0.80 (0.53-1.20)
Routine	1.22 (0.86-1.73)	**6.60 (4.04-10.78)**	**0.36 (0.15-0.84)**	**3.11 (1.08-8.97)**	1.34 (0.52-3.43)	**0.45 (0.26-0.78)**
Never worked	1.02 (0.34-3.06)	2.97 (0.82-10.73)	n/a	**7.50 (1.37-41.13)**	2.19 (0.26-18.22)	0.35 (0.05-2.67)
Not classified	**2.99 (1.65-5.40)**	2.42 (0.81-7.25)	n/a	4.92 (0.96-25.36)	**5.11 (1.50-17.40)**	1.19 (0.48-2.96)
Level of education						
Postgraduate	Reference	Reference	Reference	Reference	Reference	Reference
Undergraduate	1.79 (0.96-3.31)	1.58 (0.66-3.80)	**0.60 (0.28-1.30)**	0.91 (0.26-3.14)	0.40 (0.15-1.04)	**0.48 (0.28-0.81)**
Vocational	**2.75 (1.49-5.08)**	2.25 (0.95-5.36)	**0.39 (0.16-0.98)**	2.17 (0.89-8.22)	0.77 (0.32-1.84)	**0.49 (0.28-0.85)**
Tertiary	1.80 (0.98-3.31)	**3.69 (1.64-8.22)**	**0.30 (0.13-0.69)**	1.62 (0.53-4.92)	**0.26 (0.09-0.71)**	**0.58 (0.36-0.94)**
School qualifications	**2.51 (1.35-4.66)**	**5.31 (2.36-11.93)**	0.52 (0.22-1.19)	2.25 (0.73-6.97)	0.53 (0.21-1.36)	**0.51 (0.29-0.89)**
No qualifications	**5.48 (2.89-10.40)**	**11.50 (5.04-26.26)**	0.35 (0.10-1.27)	2.95 (0.87-10.10)	0.59 (0.17-1.94)	0.64 (0.32-1.26)
Relationship status						
Yes	Reference	Reference	Reference	Reference	Reference	Reference
No	**1.83 (1.31-2.57)**	**1.53 (1.05-2.23**)	1.31 (0.65-2.65)	2.11 (0.98-4.50)	2.07 (0.89-4.81)	**2.72 (1.61-4.57)**
Living arrangements						
Yes	Reference	Reference	Reference	Reference	Reference	Reference
No	**0.48 (0.35-0.66)**	0.69 (0.15-1.01)	0.83 (0.40-1.70)	**0.47 (0.25-0.87)**	**0.37 (0.18-0.73)**	1.16 (0.69-1.93)
Cultural distance						
Yes	**7.34 (4.68-11.50)**	**3.55 (2.18-5.78)**	**9.95 (6.25-15.82)**	**12.10 (7.61-19.25)**	**9.04 (5.66-14.43)**	**8.95 (5.70-14.07)**
No	Reference	Reference	Reference	Reference	Reference	Reference
Discrimination (0-12)	**1.61 (1.47-1.77**	1.12 (0.98-1.28)	1.22 (0.98-1.54)	**1.53 (1.30-1.80)**	**1.69 (1.43-1.99)**	**1.32 (1.15-1.51)**

Odds are relative to the White majority group.

Odds ratios in **bold** are statistically significant (p<0.05).

6.4.4.2 Associations between remaining covariates and exposures.

The full polychoric correlation matrix is found in Table 6.5 below. Whilst there were a large number of statistically significant correlations, the only correlations exceeding ± 0.5 were between older age and relationship status and living arrangements, and between having ever lived with someone other than one's parents and ever having been in a relationship.

Table 6.5: Polychoric correlation matrix of covariates and exposures

	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1.Age	1										
2.Male sex	**-0.08**	1									
3.Paternal age	0.03	0.03	1								
4.Childhood trauma	**-0.04**	-0.05	-0.03	1							
5.Cannabis use	**-0.30**	**0.34**	-0.01	**0.17**	1						
6.Paternal SES	0.03	0.06	**-0.07**	**0.13**	**-0.15**	1					
7. Education level	0.04	**-0.10**	-0.00	**-0.21**	0.03	**-0.38**	1				
8. Relationship	**0.54**	**-0.29**	**-0.04**	-0.04	-0.07	-0.00	**0.16**	1			
9. Living situation	**0.55**	**-0.23**	-0.02	**0.07**	**-0.11**	**-0.07**	**0.22**	**0.61**	1		
10. Cultural distance	-0.06	-0.05	**0.09**	**0.14**	-0.00	**-0.15**	**0.10**	0.06	**0.26**	1	
11. Discrimination	0.03	**0.07**	0.02	**0.28**	**0.18**	0.03	-0.00	**0.09**	**0.11**	**0.19**	1

Correlation coefficients in **bold** are significant (p<0.05)

6.4.4 Regression by diagnostic outcome

In this Section, I established the odds of FEP, non-affective disorders and affective disorders for all covariates and exposures, and built the multivariable model as outlined in Section 6.3.3.

6.4.4.1 All psychotic disorders

Crude associations

As can be seen in Table 6.6 below, there was a crude association between younger age (OR: 0.96, 95%CI: 0.96-0.97) and male sex (OR: 1.78, 95%CI: 1.51-2.09) and increased odds of psychosis. Those who ever smoked cannabis also had increased odds of developing psychosis (OR: 1.94, 95%CI: 1.63-2.30), as did those with increased levels of childhood trauma (OR for a 1-point increase: 1.06, 95%CI: 1.04-1.07). Paternal socioeconomic status was associated with odds of psychosis in a dose-response relationship (OR for intermediate 1.32, 95%CI: 1.02-1.69 and OR for never worked 6.83, 95%CI: 3.85-16.37). No relationship was found between psychosis odds and paternal age (Table 6.6).

Table 6.6: Psychosis risk associated with covariates, by diagnostic category.

	All psychotic disorders		Non-affective disorders		Affective disorders	
Cases	*n = 1,100*		*n = 767*		*n = 305*	
Variable	OR (95%CI)	P-value	OR (95%CI)	P-value	OR (95%CI)	P-value
Age	**0.96 (0.96-0.97)**	<0.01	**0.96 (0.95-0.97)**	**<0.01**	**0.98 (0.96-0.99)**	**<0.01**
Sex						
Male	**1.78 (1.51-2.09)**	<0.01	**2.15 (1.78-2.59)**	**<0.01**	1.13 (0.88-1.45)	0.35
Female	Reference		Reference		Reference	
Paternal age	1.00 (0.99-1.02)	0.49	1.00 (0.98-1.02)	0.69	1.00 (0.98-1.02)	0.77
Childhood trauma	**1.06 (1.04-1.07)**	**<0.01**	**1.05 (1.05-1.06)**	**<0.01**	**1.05 (1.04-1.06)**	**<0.01**
Cannabis use						
Yes	**1.94 (1.63-2.30)**	**<0.01**	**2.05 (1.68-2.49)**	**<0.01**	**1.76 (1.34-2.30)**	**<0.01**
No	Reference		Reference		Reference	
Paternal SES						
Professional	Reference		Reference		Reference	
Intermediate	**1.32 (1.02-1.69)**	**0.03**	**1.34 (1.01-1.79)**	**0.04**	1.31 (0.89-1.94)	0.17
Lower	**1.62 (1.27-2.05)**	**<0.01**	**1.83 (1.39-2.40)**	**<0.01**	1.34 (0.92-1.96)	0.13
Routine	**1.81 (1.38-2.36)**	**<0.01**	**1.95 (1.43-2.66)**	**<0.01**	**1.59 (1.06-2.39)**	**0.06**
Never worked	**6.83 (2.85-16.37)**	**<0.01**	**9.18 (3.73-22.54)**	**<0.01**	2.49 (0.48-12.80)	0.27
Not classified	**4.04 (2.39-6.82)**	**<0.01**	**4.79 (2.75-8.34)**	**<0.01**	**2.65 (1.12-6.27)**	**0.03**

Odds ratios in **bold** are significant (p<0.05)

In terms of indicators of social distance, level of education was associated with psychosis risk in an approximate dose-response relationship (OR for no qualifications: 10.67, 95%CI: 7.04-16.16) (Table 6.7 below), and both ever having been in a relationship (OR: 0.26, 95%CI: 0.21-0.32) and ever having lived with someone (OR: 0.41, 95%CI: 0.34-0.50) were strongly protective. Both cultural distance (OR: 2.05, 95%CI: 1.61-2.62) and self-perceived discrimination (OR for a one event increase: 1.21, 95%CI: 1.12-1.31) were associated with increased psychosis risk.

Ethnic minority status was associated with excess odds of psychosis (OR: 2.09, 95%CI: 1.73-2.52), and this risk was highest in the North African (OR: 3.97, 95%CI: 2.27-6.95) and Black (OR: 2.95, 95%CI: 2.21-3.93) groups. Odds were also attenuated in the 'Other' (OR: 2.28, 95%CI: 1.26-4.15) and Mixed (OR: 2.02, 95%CI: 1.49-2.74) groups, but not in the White other and Asian groups (Table 6.7).

Multivariable adjustments

After adjusting for covariates (Model A: age, sex, their interaction, paternal age, cannabis use, childhood trauma and paternal socioeconomic status), the OR of psychosis for all ethnic minorities diminished to 1.65 (95%CI: 1.34-2.03). ORs were reduced markedly in most ethnic groups (Table 6.7, Figure 6.4), but an excess risk remained in the North-African (OR: 3.05, 95%CI: 1.64-5.67), Black (OR: 2.21, 95%CI: 1.60-3.04) and Mixed groups (OR: 1.49, 95%CI: 1.06-2.08). This model explained 13.8 % of variance.

The further addition of indicators of social distance (Model B: level of education, relationship status, living arrangements) did not change the odds ratio for all ethnic minorities (1.65, 95%CI: 1.36-2.12), but this masked heterogeneity between ethnic groups. For groups with a lower level of education (the Black and Mixed groups), ORs were reduced, to 1.88 (95%CI: 1.34-2.64) and 1.30 (95%CI: 0.91-1.86) respectively (Table 6.7, Figure 6.4). ORs increased for groups with a higher level of education (Asian, OR: 1.78, 95%CI:

0.96-3.32 and White other, OR: 1.29, 95%CI: 0.88-1.90), suggesting a protective effect. This model explained 17.5 % of variance. As can be seen in Table 6.7 and Figure 6.4, the addition of cultural distance (Model C) substantially attenuated the odds of psychosis in all ethnic minority groups (OR: 1.30, 95%CI: 1.01-1.67), although an excess risk remained. This appeared to be driven by the Black (OR: 1.58, 95%CI: 1.11-2.24) group, as no other individual ethnic group retained excess odds. Adding self-perceived discrimination to the model (D) did not improve model fit (R^2 remained 21.1%) and did not further attenuate psychosis risk in the overall ethnic minority group (Table 6.7, Figure 6.4). Discrimination itself was no longer significantly associated with psychosis risk (OR: 1.04, 95%CI: 0.95-1.14), but indicators of social distance (OR for no educational qualifications: 8.72, 95%CI: 5.28-14.41, OR for relationship status: 0.37, 95%CI: 0.29-0.48 and OR for living arrangements 0.70 (95%CI: 0.54-0.89) and cultural distance (OR: 2.16, 95%CI: 1.45-3.21) remained statistically significant (Table 6.7).

Adjustment from crude to best-fitting multivariable model (C) reduced the odds of psychosis in all ethnic minorities from 2.09 (95%CI: 1.73-2.52) to 1.30 (95%CI: 1.01-1.67), and only the Black group retained excess odds (OR: 1.58, 95%CI: 1.07-2.19) compared with four ethic groups with a crude association. Even in the Black group, the OR was attenuated substantially (Figure 6.4).

Figure 6.4: Associations between FEP risk and ethnic minority groups, by statistical model.

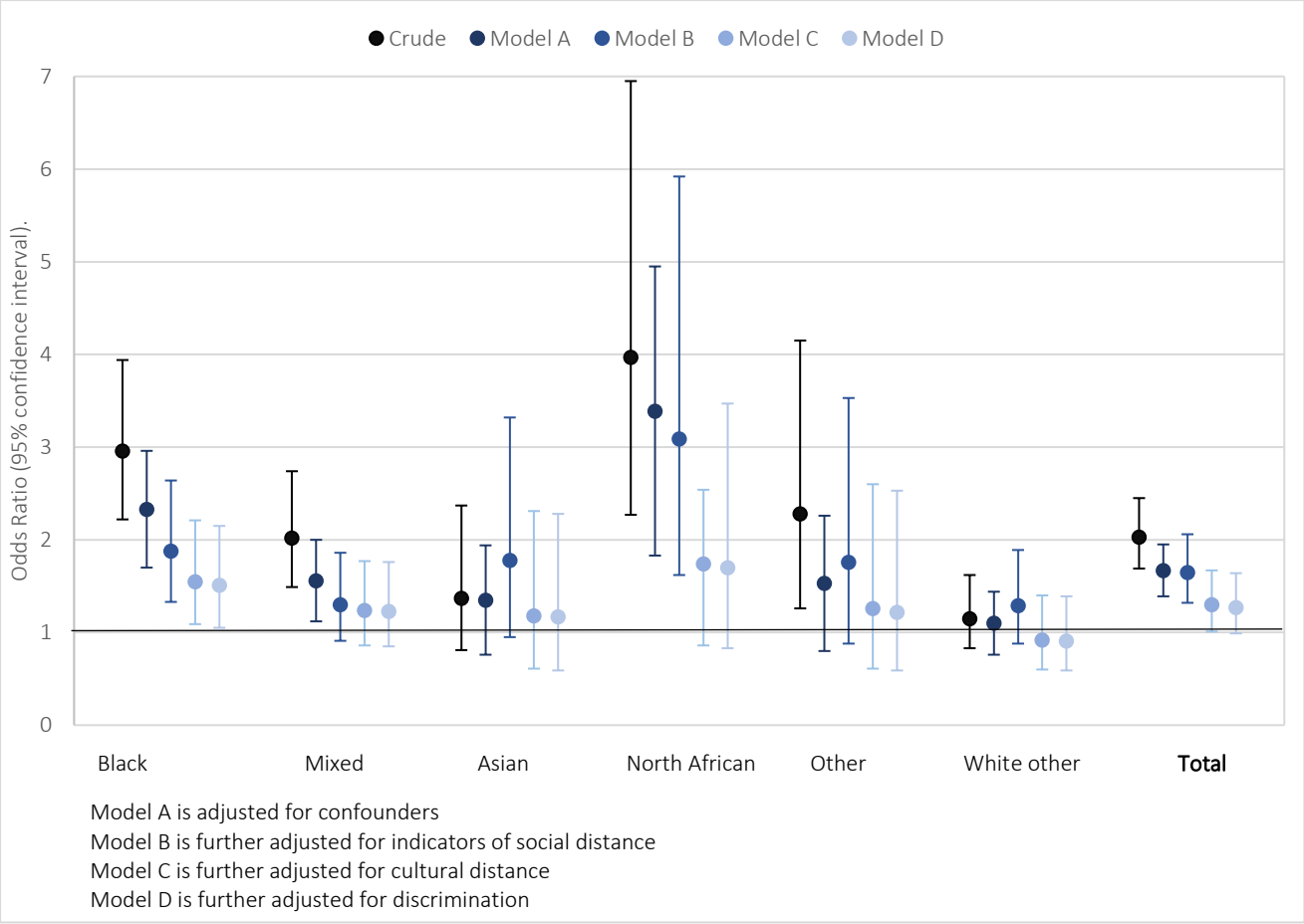


Table 6.7: Associations between FEP risk and ethnic minority groups for each statistical model.

	Crude		Model A[1]		Model B[2]		Model C[3]		Model D[4]	
Variable	OR (95%CI)	P-value	OR (95%CI)	P-value	OR (95%CI)	P-value	OR (95%CI)	P-value	OR (95%CI)	P-value
Ethnicity										
White native	Reference		Reference		Reference		Reference		Reference	
All minorities	**2.09 (1.73-2.52)**	**<0.01**	**1.65 (1.34-2.03)**	**<0.01**	**1.65 (1.32-2.06)**	**<0.01**	**1.30 (1.01-1.67)**	**0.04**	1.27 (0.99-1.64)	0.06
Black	**2.95 (2.21-3.93)**	**<0.01**	**2.21 (1.60-3.04)**	**<0.01**	**1.88 (1.34-2.64)**	**<0.01**	**1.58 (1.11-2.24)**	**0.01**	**1.53 (1.07-2.19)**	**0.02**
Mixed	**2.02 (1.49-2.73)**	**<0.01**	**1.49 (1.06-2.08)**	**0.02**	1.30 (0.91-1.86)	0.16	1.24 (0.86-1.78)	**0.25**	1.23 (0.85-1.76)	0.27
Asian	1.38 (0.81-2.37)	0.24	1.47 (0.82-2.64)	0.20	1.78 (0.96-3.32)	0.07	1.23 (0.63-2.38)	**0.55**	1.22 (1.63-2.36)	0.57
North-African	**3.97 (2.27-6.95)**	**<0.01**	**3.05 (1.64-5.67)**	**<0.01**	**3.09 (1.61-5.92)**	**<0.01**	1.89 (0.94-3.82)	**0.08**	1.85 (0.92-3.75)	0.09
Other	**2.28 (1.26-4.15)**	**<0.01**	1.39 (0.73-2.66)	0.32	1.76 (0.88-3.53)	0.11	1.24 (0.60-2.56)	**0.57**	1.20 (0.57-2.50)	0.63
White other	1.15 (0.83-1.61)	0.40	1.09 (0.76-1.58)	0.62	1.29 (0.88-1.90)	0.20	0.82 (0.52-1.30)	**0.40**	0.82 (0.52-1.30)	0.40
Level of education			n/a							
Postgraduate	Reference				Reference		Reference		Reference	
Undergraduate	**1.53 (1.06-2.22)**	**0.02**			**1.52 (1.00-2.29)**	**0.05**	**1.52 (1.01-2.30)**	**0.05**	**1.53 (1.01-2.31)**	**0.05**
Vocational	**3.38 (2.35-4.87)**	**<0.01**			**2.61 (1.72-3.95)**	**<0.01**	**2.60 (1.71-3.95)**	**<0.01**	**2.59 (1.71-3.93)**	**<0.01**
Tertiary	**2.51 (1.78-3.56)**	**<0.01**			**1.99 (1.33-2.98)**	**<0.01**	**2.00 (1.33-2.98)**	**<0.01**	**1.99 (1.33-2.98)**	**<0.01**
School qualifications	**6.22 (4.34-8.90)**	**<0.01**			**5.17 (3.40-7.85)**	**<0.01**	**5.23 (3.44-7.96)**	**<0.01**	**5.24 (3.44-7.97)**	**<0.01**
No qualifications	**10.67 (7.04-16.16)**	**<0.01**			**8.67 (5.25-14.31)**	**<0.01**	**8.72 (5.28-14.41)**	**<0.01**	**8.76 (5.30-14.47)**	**<0.01**
Relationship status			n/a							
Yes	**0.26 (0.21-0.32)**	**<0.01**			**0.38 (0.29-0.49)**	**<0.01**	**0.37 (0.29-0.48)**	**<0.01**	**0.37 (0.29-0.48)**	**<0.01**
No	Reference				Reference		Reference		Reference	
Living arrangements			n/a							
Yes	**0.41 (0.34-0.50)**	**<0.01**			**0.74 (0.56-0.91)**	**<0.01**	**0.70 (0.54-0.89)**	**<0.01**	**0.69 (0.54-0.89)**	**<0.01**
No	Reference				Reference		Reference		Reference	
Cultural distance			n/a		n/a					
Yes	**2.05 (1.61-2.62)**	**<0.01**					**2.17 (1.46-3.23)**	**<0.01**	**2.16 (1.45-3.21)**	**<0.01**
No	Reference						Reference		Reference	
Discrimination (0-12)	**1.21 (1.12-1.31)**	**<0.01**	n/a		n/a		n/a		1.05 (0.95-1.15)	0.33
Pseudo R^2	2.4%		13.8%		20.6%		21.1%		21.1%	

[1]: Model A is adjusted for covariates (age, sex, their interaction, paternal age, childhood trauma, cannabis use and paternal socioeconomic status)
[2]: Model B is further adjusted for indicators of social distance (level of education, relationship status and living arrangements)
[3]: Model C is further adjusted for cultural distance
[4]: Model D is further adjusted for discrimination

Odds ratios in **bold** are significant (p<0.05)

6.4.3.2 Non-affective and affective psychoses

Crude associations

The associations between covariates and non-affective and affective disorders were comparable with the FEP associations (Table 6.6). The exception to this was a lack of association between male sex and excess odds of affective disorders (OR: 1.13, 95%CI: 0.88-1.45).

Lower education was associated with increased odds of both a non-affective (OR for no qualifications: 16.62, 95%CI: 6.68-16.87) and an affective disorder (OR: 14.42, 95%CI: 7.14-29.2). Ever having been in a relationship was protective for both outcomes (OR for a non-affective disorder: 0.22, 95%CI: 0.17-0.27,OR for an affective disorder: 0.42, 95%CI: 0.49-0.90), as was ever having lived with someone (OR for a non-affective disorder: 0.35, 95%CI: 0.28-0.43, OR for an affective disorder: 0.67, 0.49-0.90). Cultural distance was associated with an OR of 2.36 (95%CI: 1.82-3.07) of developing a non-affective disorder, and with an OR of 1.49 (95%CI; 1.01-2.20) of an affective disorder. Increased levels of discrimination were also associated with both outcomes to a similar extent (Tables 6.8 and 6.9).

The overall ethnic minority group had increased odds of developing a non-affective (OR: 2.12, 95%CI: 1.71-2.63) and an affective (OR: 1.94, 95%CI: 1.47-2.55) disorder. Odds of a non-affective disorder were raised in the North African (OR: 5.16, 95%CI: 2.81-9.49), Black (OR: 3.10, 95%CI: 2.25-4.28), Other (OR: 2.45, 95%CI: 1.29-4.64) and Mixed (OR: 1.76, 95%CI: 1.21-2.53) groups. Odds of affective disorder were raised in the Black (OR: 2.48, 95%CI: 1.63-3.74) and Mixed (OR: 2.45, 95%CI: 1.64-3.67) groups only (Tables 6.8 and 6.9).

Multivariable adjustment

The overall patterns by diagnostic categories were similar to the pattern observed across all psychotic disorders (Tables 6.8 and 6.9, Figures 6.5 and 6.6). Model C, including covariates, indicators of social distance and cultural distance accounted for overall excess odds for both the non-affective (OR: 1.26, 95%CI: 0.93-1.70) the affective disorders (OR: 1.29, 95%CI: 0.91-1.84). Only the Black (OR: 1.61, 95%CI: 1.08-2.41) and North African (OR: 2.34, 95%CI: 1.08-5.10) groups retained excess odds of non-affective disorders (Tables 6.8 and 6.9, Figures 6.5 and 6.6). This was the best-fitting model for both outcomes with an R^2 of 25.5% for non-affective and of 14.7% for affective disorders. As was the case with all psychotic disorders, there was no evidence of an association between discrimination and non-affective psychosis or affective disorders in a multivariable model (Tables 6.8 and 6.9).

Figure 6.5: Associations between risk of non-affective disorder and ethnic minority groups, by statistical model.

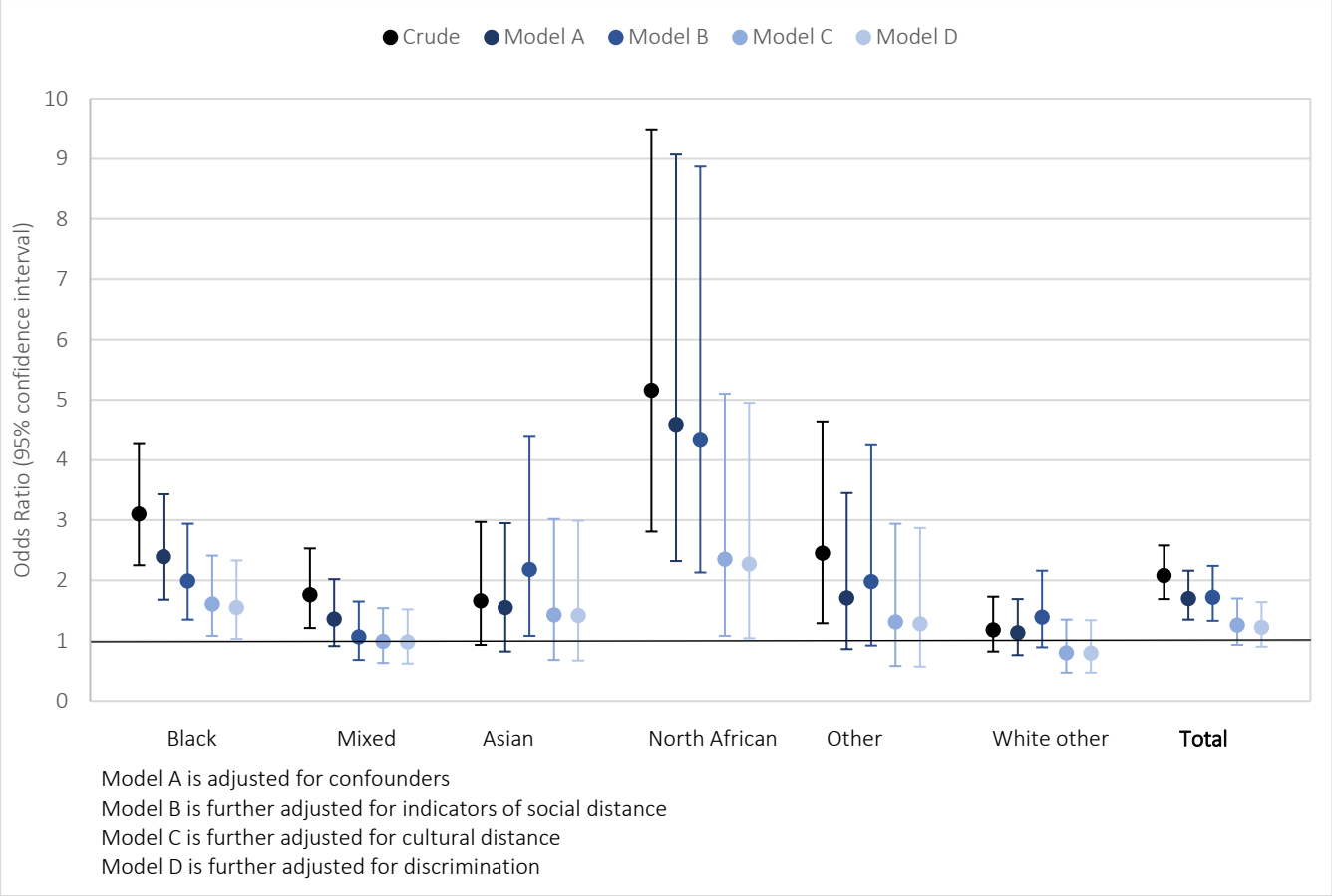


Figure 6.6: Associations between risk of affective disorder and ethnic minority groups, by statistical model.

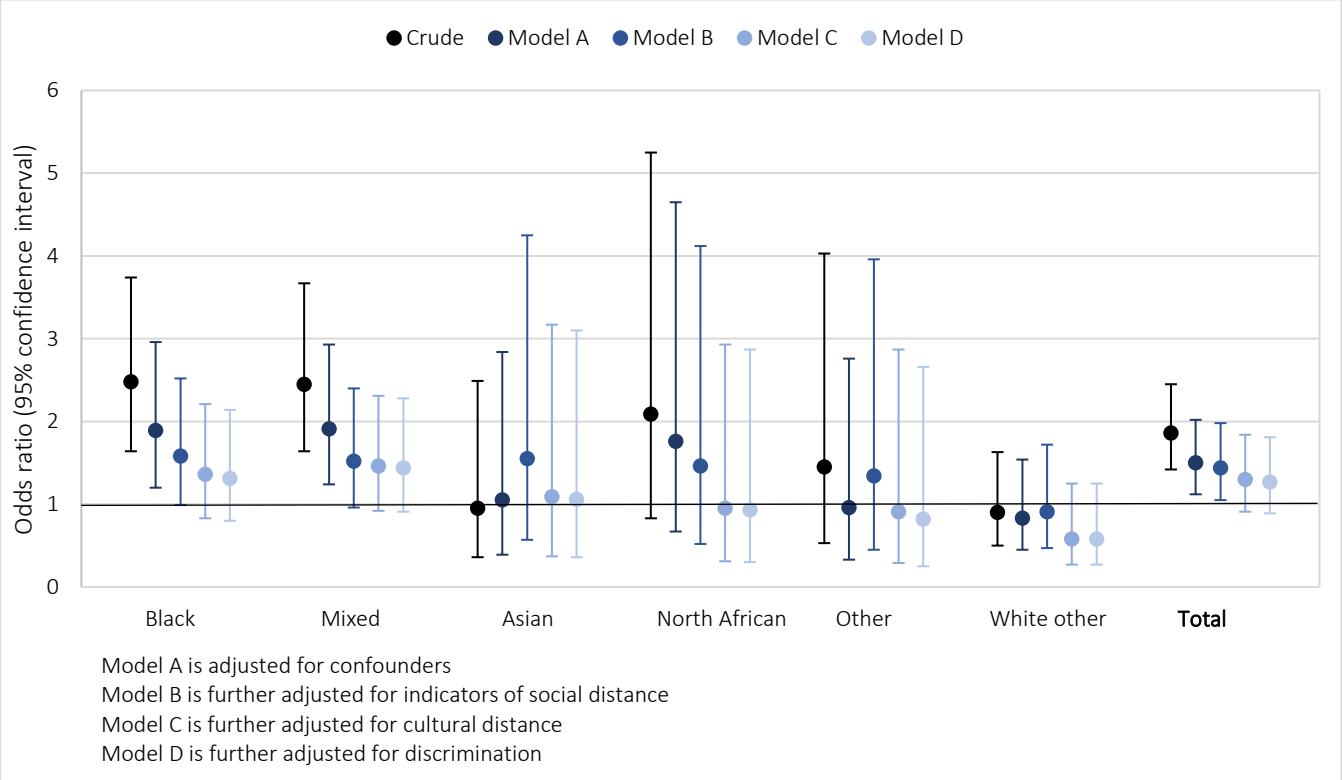


Table 6.8: Associations between risk of non-affective disorder and exposures of interest, for each statistical model (767 cases versus 1,473 controls).

	Crude		Model A[1]		Model B[2]		Model C[3]		Model D[4]	
Variable	OR (95%CI)	P-value	OR (95%CI)	P-value	OR (95%CI)	P-value	OR (95%CI)	P-value	OR (95%CI)	P-value
Ethnicity										
White native	Reference		Reference		Reference		Reference		Reference	
All minorities	**2.12 (1.71-2.63)**	**<0.01**	**1.66 (1.30-2.11)**	**<0.01**	**1.73 (1.33-2.24)**	**<0.01**	1.26 (0.93-1.70)	0.13	1.22 (0.90-1.65)	0.20
Black	**3.10 (2.25-4.28)**	**<0.01**	**2.26 (1.57-3.26)**	**<0.01**	**1.98 (1.35-2.94)**	**<0.01**	**1.61 (1.08-2.41)**	**0.02**	**1.55 (1.03-2.33)**	**0.04**
Mixed	**1.76 (1.22-2.53)**	**<0.01**	1.26 (0.84-1.90)	0.26	1.06 (0.68-1.65)	0.78	0.99 (0.63-1.54)	0.97	0.97 (0.62-1.52)	0.90
Asian	1.66 (0.93-2.97)	0.09	1.74 (0.91-3.33)	0.10	**2.18 (1.08-4.40)**	**0.03**	1.43 (0.67-3.02)	0.35	1.42 (0.67-3.00)	0.36
North-African	**5.16 (2.81-9.49)**	**<0.01**	**4.05 (2.03-8.08)**	**<0.01**	**4.34 (2.13-8.87)**	**<0.01**	**2.34 (1.08-5.10)**	**0.03**	**2.27 (1.04-4.94)**	**0.04**
Other	**2.45 (1.29-4.64)**	**<0.01**	1.51 (0.75-3.05)	0.25	1.98 (0.92-4.26)	0.08	1.31 (0.58-2.94)	0.51	1.28 (0.57-2.88)	0.56
White other	1.19 (0.82-1.73)	0.37	1.13 (0.75-1.70)	0.56	1.39 (0.89-2.15)	0.14	0.80 (0.47-1.35)	0.40	0.79 (0.47-1.34)	0.39
Level of education			n/a							
Postgraduate	Reference				Reference		Reference		Reference	
Undergraduate	1.53 (1.00-2.35)	0.05			1.57 (0.97-2.56)	0.07	1.60 (0.98-2.62)	0.06	1.62 (0.99-2.64)	0.06
Tertiary	**3.67 (2.43-5.55)**	**<0.01**			**2.75 (1.69-4.48)**	**<0.01**	**2.81 (1.72-4.59)**	**<0.01**	**2.80 (1.72-4.58)**	**<0.01**
Vocational	**2.28 (1.52-3.40)**	**<0.01**			**1.76 (1.09-2.84)**	**0.02**	**1.79 (1.11-2.90)**	**0.02**	**1.80 (1.11-2.90)**	**0.02**
School qualifications	**6.26 (4.17-9.42)**	**<0.01**			**5.30 (3.21-8.66)**	**<0.01**	**5.48 (3.34-8.99)**	**<0.01**	**5.52 (3.37-9.06)**	**<0.01**
No qualifications	**10.62 (6.68-16.87)**	**<.0.01**			**8.92 (4.98-15.95)**	**<0.01**	**9.34 (5.20-16.77)**	**<0.01**	**9.42 (5.25-16.90)**	**<0.01**
Relationship status			n/a							
Yes	**0.22 (0.17-0.27)**	**<0.01**			**0.31 (0.23-0.41)**	**<0.01**	**0.30 (0.23-0.41)**	**<0.01**	**0.30 (0.22-0.40)**	**<0.01**
No	Reference				Reference		Reference		Reference	
Living arrangements			n/a							
Yes	**0.35 (0.28-0.43)**	**<0.01**			**0.65 (0.49-0.86)**	**<0.01**	**0.64 (0.48-0.84)**	**<0.01**	**0.64 (0.78-0.84)**	**<0.01**
No	Reference				Reference		Reference		Reference	
Cultural distance			n/a		n/a					
Yes	**2.36 (1.82-3.07)**	**<0.01**					**2.55 (1.63-3.98)**	**<0.01**	**2.55 (1.63-3.98)**	**<0.01**
No	Reference						Reference		Reference	
Discrimination (0-12)	**1.21 (1.11-1.31)**	**<0.01**	n/a		n/a		n/a		1.07 (0.96-1.18)	0.37
Pseudo R^2	2.9%		16.6%		24.6%		25.3%		25.5%	

[1]: Model A is adjusted for covariates (age, sex, their interaction, paternal age, childhood trauma, cannabis use and paternal socioeconomic status)
[2]: Model B is further adjusted for indicators of social distance (level of education, relationship status, living arrangements)
[3]: Model C is further adjusted for cultural distance
[4]: Model D is further adjusted for discrimination

Odds ratios in **bold** are significant (p<0.05)

Table 6.9: Associations between risk of affective disorder and exposures of interest, for each statistical model (305 cases versus 1,473 controls).

	Crude		Model A[1]		Model B[2]		Model C[3]		Model D[4]	
Variable	OR (95%CI)	P-value	OR (95%CI)	P-value	OR (95%CI)	P-value	OR (95%CI)	P-value	OR (95%CI)	P-value
Ethnicity										
White native	Reference		Reference		Reference		Reference		Reference	
All minorities	**1.94 (1.47-2.55)**	**<0.01**	**1.54 (1.14-2.08)**	**<0.01**	**1.44 (1.05-1.98)**	**<0.01**	1.29 (0.91-1.84)	0.15	1.27 (0.89-1.81)	0.19
Black	**2.48 (1.65-3.74)**	**<0.01**	**1.83 (1.17-2.88)**	**<0.01**	1.58 (0.99-2.52)	0.06	1.36 (0.83-2.21)	0.22	1.31 (0.80-2.14)	0.28
Mixed	**2.45 (1.64-3.67)**	**<0.01**	**1.86 (1.21-2.87)**	**<0.01**	1.52 (0.96-2.40)	0.07	1.46 (0.92-2.30)	0.11	1.44 (0.91-2.28)	0.12
Asian	0.95 (0.36-2.49)	0.91	1.09 (0.41-2.94)	0.86	1.55 (0.57-4.25)	0.39	1.09 (0.37-3.16)	0.88	1.06 (0.36-3.10)	0.92
North-African	2.08 (0.83-5.24)	0.12	1.68 (0.63-4.65)	0.30	1.47 (0.52-4.13)	0.47	0.96 (0.31-2.93)	0.94	0.93 (0.30-2.87)	0.91
Other	1.46 (0.53-4.03)	0.47	0.95 (0.33-2.76)	0.93	1.34 (0.45-3.95)	0.60	0.91 (0.29-2.88)	0.87	0.82 (0.25-2.66)	0.74
White other	0.91 (0.50-1.63)	0.73	0.27 (0.46-1.58)	0.61	0.91 (0.47-1.73)	0.76	0.58 (0.27-1.25)	0.17	0.58 (0.27-1.25)	0.17
Level of education			n/a							
Postgraduate	Reference				Reference		Reference		Reference	
Undergraduate	1.99 (0.99-4.00)	0.05			1.83 (0.89-3.74)	0.10	1.84 (0.90-3.78)	0.09	1.88 (0.92-3.87)	0.08
Tertiary	**3.32 (1.66-6.62)**	**<0.01**			**2.92 (1.4-6.03)**	**<0.01**	**2.86 (1.38-5.93)**	**<0.01**	**2.89 (1.40-6.00)**	**<0.01**
Vocational	**3.71 (1.94-7.10)**	**<0.01**			**3.26 (1.64-6.48)**	**<0.01**	**3.27 (1.64-6.50)**	**<0.01**	**3.31 (1.66-6.58)**	**<0.01**
School qualifications	**7.84 (4.05-15.19)**	**<0.01**			**6.87 (3.39-13.93)**	**<0.01**	**6.85 (3.38-13.89)**	**<0.01**	**6.64 (3.42-14.08)**	**<0.01**
No qualifications	**14.44 (7.14-29.21)**	**<0.01**			**10.90 (5.00-23.77)**	**<0.01**	**10.68 (4.89-23.31)**	**<0.01**	**10.90 (4.99-23.80)**	**<0.01**
Relationship status			n/a							
Yes	**0.42 (0.31-0.57)**	**<0.01**			**0.56 (0.38-0.83)**	**<0.01**	**0.56 (0.38-0.83)**	**<0.01**	**0.56 (0.38-0.83)**	**<0.01**
No	Reference				Reference		Reference		Reference	
Living arrangements			n/a							
Yes	**0.67 (0.49-0.90)**	**<0.01**			0.92 (0.62-1.35)	0.67	0.89 (0.60-1.31)	0.55	0.89 (0.60-1.31)	0.55
No	Reference				Reference		Reference		Reference	
Cultural distance			n/a		n/a					
Yes	**1.49 (1.01-2.20)**	**0.05**					**2.06 (1.13-3.77)**	**0.02**	**2.06 (1.12-3.76)**	**<0.02**
No	Reference						Reference		Reference	
Discrimination (0-12)	**1.25 (1.11-1.40)**	**<0.01**	n/a		n/a		n/a		1.08 (0.94-1.24)	0.26
Pseudo R^2	2.7%		9.2%		14.4%		14.7%		14.7%	

[1]: Model A is adjusted for covariates (age, sex, their interaction, paternal age, childhood trauma, cannabis use and paternal socioeconomic status)
[2]: Model B is further adjusted for indicators of social distance (level of education, relationship status, living arrangements)
[3]: Model C is further adjusted for cultural distance
[4]: Model D is further adjusted for discrimination

Odds ratios in **bold** are significant (p<0.05)

6.4.5 Religious minority status

The religious majority (no religion or Christian) made up 83.4% of the sample (n=2,168) and the religious minority (Jewish, Muslim or any other religion) 14.9% (n=383, Table 6.10). Twenty-two people (0.9%) had no data on their religion, and this data was imputed. Cases were more likely to follow any religion (χ^2: 43.2, 5df; p<0.01) and were more likely to belong to a religious minority (χ^2: 27.1, 2df; p<0.01) than controls.

Table 6.10: Distribution of religion.

Religious group	Cases N (%)	Controls N (%)
Majority	882 (80.2)	1,287 (87.4)
Christian	555 (50.5)	739 (50.2)
No religion	327 (29.7)	549 (37.3)
Minority	203 (18.5)	179 (12.2)
Muslim	155 (14.1)	122 (8.3)
Jewish	9 (0.8)	23 (1.6)
Other	39 (3.6)	33 (2.2)
Missing data	15 (1.36)	7 (0.5)

Participants from a religious minority had lower educational attainment than their religious majority counterparts: they were more likely to have a tertiary education (OR: 1.91, 95%CI: 1.21-3.03), only have school qualifications (OR: 1.87, 95%CI: 1.16-3.01) or leaving school without any qualifications (OR: 3.22, 95%CI: 1.95-5.32). They were also less likely to have ever been in a relationship (OR: 0.70, 95%CI: 0.55-0.91). They had increased cultural distance (OR: 6.92, 95%CI: 4.93-9.73) compared with those in the religious majority, and reported higher levels of discrimination (OR: 1.17, 95%CI: 1.06-1.29) (Table 6.11 below).

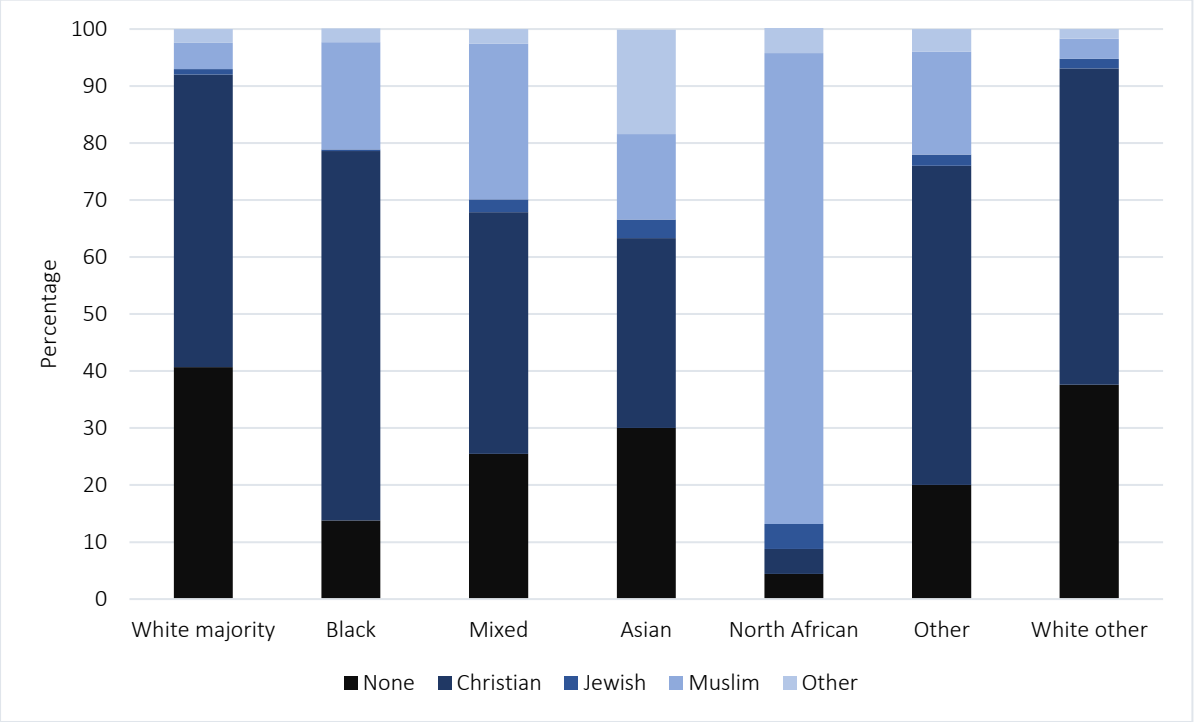
Table 6.11: Associations between religious minority status and remaining exposures

	OR (95%CI)[1]	P-value
Level of education		
Postgraduate	Reference	
Undergraduate	1.08 (0.65-1.80)	0.77
Vocational	1.12 (0.67-1.87)	0.68
Tertiary	**1.91 (1.21-3.03)**	**<0.01**
School qualifications	**1.87 (1.16-3.01)**	**<0.01**
No qualifications	**3.22 (1.95-5.32)**	**<0.01**
Relationship status		
Yes	**0.70 (0.55-0.91)**	**<0.01**
No	Reference	
Living arrangements		
Yes	1.04 (0.80-1.36)	0.78
No	Reference	
Cultural distance		
Yes	**3.54 (2.70-4.64**)	**<0.01**
No	Reference	
Discrimination (0-12)	**1.17 (1.06-1.29)**	**<0.01**

[1] Odds are relative to the religious majority group
Odds ratios in **bold** are significant (p<0.05)

Religious minorities were also 25% more likely (OR: 1.26, 95%CI: 1.20-1.33) to be of any ethnic minority, so it is possible that, to some extent, I would examine the same group of people twice if I treated religious minority status as a unique characteristic. When plotting religion against ethnic group (Figure 6.8 below), this appeared to be the case to some extent. The North-African group, which had the highest risk of psychosis, is also predominantly a Muslim group (n=57; 82.3%). The Black group on the other hand appeared to be predominantly Christian (n=192; 63.7%, Figure 6.7 below). Because of this different religious make-up of the various ethnic minority groups (all were more religious than the White majority group: χ^2: 671.6, p<0.01), I included ethnicity as a confounder (first alone, then from Model A onwards).

Figure 6.7: Religion by ethnic group



Those who belong to a religious minority had increased crude odds of developing psychosis (OR: 1.84, 95%CI: 1.45-2.33) (Table 6.12, Figure 6.9). After adjusting for ethnicity, this was attenuated to 1.44 (95%CI: 1.10-1.87), and after adjusting for further covariates (Model A) this was attenuated, but remained increased (OR: 1.38, 95%CI: 1.03-1.84). This model explained 13.8% of the variance (Table 6.12). Further adjustment for indicators of social distance made little difference to the point estimate (OR: 1.35, 95%CI: 1.00-1.86). This model explained 20.2% of the variance (Table 6.12). Adding cultural distance (Model C) reduced the point estimate to 1.30 (95%CI: 0.95-1.78), and meant that religious minority status was no longer associated with excess odds of a psychotic disorder. This model explained 21.1% of the variance (Table 6.12). Further adjustment for discrimination did not improve model fit (R^2: 21.2%) nor have an effect on the point estimate (OR: 1.30, 95%CI: 0.95-1.77) (Table 6.12).

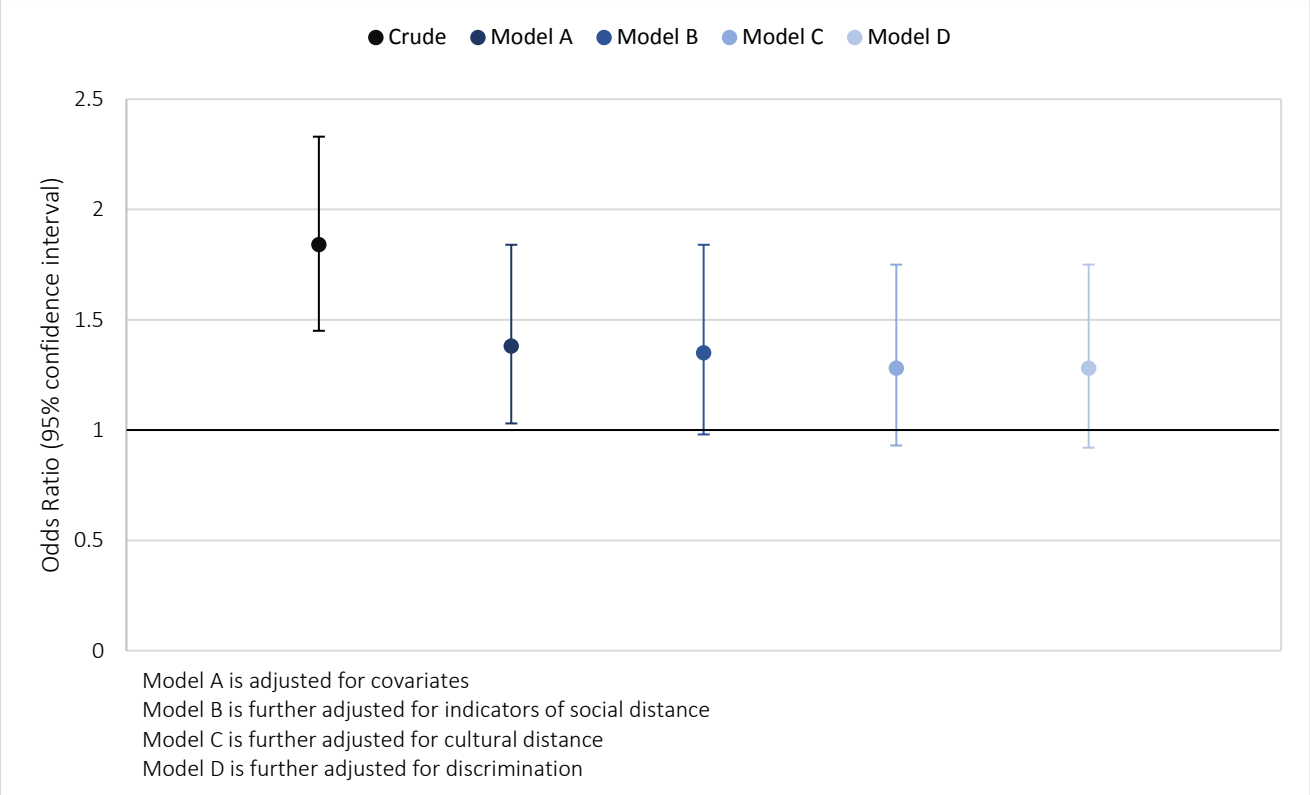
Table 6.12: Associations between FEP risk and religious minority status and remaining exposures, by statistical model.

	Crude		Model A[1]		Model B[2]		Model C[3]		Model D[4]	
Variable	OR (95%CI)	P-value	OR (95%CI)	P-value	OR (95%CI)	P-value	OR (95%CI)	P-value	OR (95%CI)	P-value
Religion										
Majority	Reference		Reference		Reference		Reference		Reference	
Minority	**1.84 (1.45-2.33)**	**0.01**	**1.38 (1.03-1.84)**	**0.03**	**1.357(1.00-1.86)**	**0.05**	1.30 (0.95-1.78)	0.10	1.30 (0.95-1.77)	0.10
Level of education			n/a							
Postgraduate	Reference				Reference		Reference		Reference	
Undergraduate	**1.53 (1.06-2.22)**	**0.02**			1.51 (1.00-2.28)	0.05	**1.53 (1.01-2.31)**	**0.04**	**1.53 (1.01-2.32)**	**0.04**
Vocational	**3.38 (2.35-4.87)**	**<0.01**			**2.53 (1.67-3.83)**	**<0.01**	**2.60 (1.17-3.94)**	**<0.01**	**2.59 (1.71-3.93)**	**<0.01**
Tertiary	**2.51 (1.78-3.56)**	**<0.01**			**1.93 (1.29-2.87)**	**<0.01**	**1.98 (1.32-2.96)**	**<0.01**	**1.99 (1.32-2.96)**	**<0.01**
School qualifications	**6.22 (4.34-8.90)**	**<0.01**			**4.91 (3.24-7.43)**	**<0.01**	**5.20 (3.41-7.92)**	**<0.01**	**5.21 (3.42-7.94)**	**<0.01**
No qualifications	**10.67 (7.04-16.16)**	**<0.01**			**8.21 (5.00-13.50)**	**<0.01**	**8.72 (2.23-5.27)**	**<0.01**	**8.75 (5.30-14.45**)	**<0.01**
Relationship status			n/a							
Yes	**0.26 (0.21-0.32)**	**<0.01**			**0.38 (0.29-0.50)**	**<0.01**	**0.37 (0.29-0.48)**	**<0.01**	**0.37 (0.28-0.48)**	**<0.01**
No	Reference				Reference		Reference		Reference	
Living arrangements			n/a							
Yes	**0.41 (0.34-0.50)**	**<0.01**			**0.71 (0.55-0.91)**	**<0.01**	**0.69 (0.54-0.89)**	**<0.01**	**0.69 (0.54-0.89)**	**<0.01**
No	Reference				Reference		Reference		Reference	
Cultural distance			n/a		n/a					
Yes	**2.05 (1.61-2.62)**	**<0.01**					**2.11 (1.42-3.14)**	**<0.01**	**2.10 (1.41-3.13)**	**<0.01**
No	Reference						Reference			
Discrimination (0-12)	**1.21 (1.12-1.31)**	**<0.01**	n/a		n/a		n/a		1.05 (0.95-1.15)	0.41
Pseudo R^2	0.6%		13.8%		20.2%		21.1%		21.2%	

[1]: Model A is adjusted for covariates (age, sex, their interaction, ethnicity, paternal age, childhood trauma, cannabis use and paternal socioeconomic status)
[2]: Model B is further adjusted for indicators of social distance (level of education, relationship status, living arrangements)
[3]: Model C is further adjusted for cultural distance
[4]: Model D is further adjusted for discrimination

Odds ratios in **bold** are significant (p<0.05)

Figure 6.8: Associations between FEP risk and religious minority status, by statistical model.



6.4.5.1 Post-hoc analysis

Because of the heterogeneity in religious make-up of the various ethnic groups, I conducted *post-hoc* analyses to explore whether ethnicity or religion was a stronger predictor of psychosis risk, following adjustment for covariates and remaining exposures (Table 6.13 below). Christians (OR: 1.38, 95%CI: 1.14-1.67), Muslims (OR: 2.55, 95%CI: 1.88-3.46) and those following any other religion (OR: 2.10, 95%CI: 1.29-3.44) had crude excess odds of psychosis compared with the non-religious group, but no difference was found for the Jewish group (OR: 0.86, 95%CI: 0.38-1.91), possibly due to the low number of Jewish participants (n=31). Mutually adjusting for ethnicity and religion only explained excess risk in the Christian group, but all ethnic minorities with an excess crude risk retained excess risk (Table 6.13). The *post hoc* model explained excess odds in all ethnic groups, but Christians (OR: 1.45, 95%CI: 1.17-1.85), Muslims (OR: 1.62, 95%CI: 1.09-2.40) and those following any other religion (OR: 1.89, 95%CI: 1.05-3.41) retained increased odds of any psychotic disorder, albeit substantially attenuated for Muslims and followers of any other religion. The model explained 21.5% of variance (Table 6.13).

Table 6.13: Associations between FEP risk and exposures following multivariable adjustment.

	Crude		Mutually adjusted[1]		Fully adjusted[2]	
Variable	OR (95%CI)	p-value	Or (95%CI)	p-value	OR (95%CI)	p-value
Religion						
None	Reference		Reference		Reference	
Christian	**1.38 (1.14-1.67)**	**<0.01**	1.18 (0.96-1.44)	0.10	**1.45 (1.15-1.84)**	**<0.01**
Jewish	0.86 (0.38-1.91)	0.71	0.74 (0.33-1.66)	0.46	1.33 (0.50-3.50)	0.57
Muslim	**2.55 (1.88-3.46)**	**<0.01**	**1.70 (1.20-2.40)**	**<0.01**	**1.62 (1.08-2.43)**	**0.02**
Other	**2.10 (1.29-3.44)**	**<0.01**	**1.88 (1.14-3.12)**	**0.01**	**1.94 (1.07-3.53)**	**0.04**
Ethnicity						
White native	Reference		Reference		Reference	
Black	**2.95 (2.21-3.93)**	**<0.01**	**2.66 (1.68-3.58)**	**<0.01**	1.35 (0.94-1.94)	0.11
Mixed	**2.02 (1.49-2.73)**	**<0.01**	**1.90 (1.40-2.59)**	**<0.01**	1.17 (0.81-1.69)	0.40
Asian	1.38 (0.81-2.37)	0.24	1.20 (0.69-2.08)	0.52	1.08 (0.55-2.12)	0.83
North African	**3.97 (2.27-6.95)**	**<0.01**	**2.77 (1.49-5.17)**	**<0.01**	1.46 (0.68-3.12)	0.32
Other	**2.28 (1.26-4.15)**	**<0.01**	**2.05 (1.12-3.75**)	**0.02**	1.08 (0.52-2.27)	0.83
White other	1.15 (0.83-1.61)	0.40	1.13 (0.81-1.59)	0.47	0.80 (0.51-1.27)	0.36
Level of education						
Postgraduate	n/a		n/a		Reference	
Undergraduate					1.50 (0.99-2.27)	0.06
Vocational					**2.55 (1.68-3.88)**	**<0.01**
Tertiary					**1.94 (1.29-2.91)**	**<0.01**
School qualifications					**5.07 (3.33-7.73)**	**<0.01**
No qualifications					**8.58 (5.18-14.19)**	**<0.01**
Relationship status						
Yes	n/a		n/a		**0.37 (0.28-0.48)**	**<0.01**
No					Reference	
Living arrangements						
Yes	n/a		n/a		**0.70 (0.54-0.89)**	**<0.01**
No					Reference	
Cultural distance						
Yes	n/a		n/a		**2.08 (1.40-3.11)**	**<0.01**
No					Reference	
Discrimination (0-12)	n/a		n/a		1.04 (0.95-1.14)	0.40
Pseudo R^2	n/a		2.7%		21.5%	

[1] Only mutually adjusted, no further covariates or exposures included in this model

[2]Further adjusted for covariates: age, sex, their interaction, paternal age, childhood trauma, cannabis use and paternal socioeconomic status.

Odds ratios in **bold** are significant (p<0.05)

6.4.6 Sensitivity analyses

In order to be able to assess bias in the multiply imputed dataset and associated models, I carried out sensitivity analyses. I repeated the analyses for the ethnic minority groups on the complete case dataset for the main outcome of interest (all psychotic disorders). As can be seen in Table 6.14 below, the broad pattern remained similar, with each additional modelling step reducing point estimates for ethnic minorities overall. However, where in the imputed best-fitting model (C), the odds for ethnic minorities overall and for Black group remained significantly increased, in the sensitivity analyses no increased odds were established, which might have been due to loss of precision. In the crude model, there was limited loss of observations due to missing data (n=2; 0.1%). However, with each additional adjustment, further observations were lost. In the final model only 67.8% (n=1,744) of observations remained.

Table 6.14: Sensitivity analyses of the associations between FEP risk and ethnic group, by statistical model (complete cases only)

	Crude		Model A[1]		Model B[2]		Model C[3]		Model D[4]	
N (%)	*2,571 (99.9)*		*2,161 (84.0)*		*1,999 (77.7)*		*1,999 (77.7)*		*1,744 (67.8)*	
	OR (95%CI)	p-value	OR (95%CI)	p-value	OR (95%CI)	p-value	OR (95%CI)	p-value	OR (95%CI)	p-value
Ethnicity										
White majority	Reference		Reference		Reference		Reference		Reference	
All minorities	**1.92 (1.63-2.26)**	**<0.01**	**1.29 (1.05-1.59)**	**0.02**	**1.31 (1.04-1.65)**	**0.02**	1.01 (0.78-1.32)	0.90	1.10 (0.83-1.46)	0.49
Black	**2.56 (1.99-3.29)**	**<0.01**	**1.46 (1.06-2.02)**	**0.02**	1.33 (0.94-1.89)	0.10	1.09 (0.75-1.57)	0.66	1.19 (0.81-1.77)	0.37
Mixed	**1.66 (1.26-2.19)**	**<0.01**	1.12 (0.80-1.56)	0.51	0.95 (0.65-1.37)	0.77	0.91 (0.63-1.32)	0.62	1.02 (0.69-1.52)	0.88
Asian	1.37 (0.82-2.30)	0.23	1.53 (0.84-2.77)	0.16	1.93 (1.00-3.73)	0.05	1.23 (0.30-2.53)	0.57	1.05 (0.48-2.27)	0.90
North-African	**3.35 (2.01-5.58)**	**<0.01**	**2.45 (1.33-4.53)**	**<0.01**	**2.37 (1.18-4.78)**	**0.02**	1.30 (0.60-2.05)	0.51	1.28 (0.58-2.85)	0.54
Other	**2.51 (1.41-4.46)**	**<0.01**	1.29 (0.63-2.65)	0.48	1.93 (0.83-4.49)	0.13	1.42 (0.58-3.44)	0.44	1.45 (0.58-3.67)	0.42
White other	1.11 (0.81-1.53)	0.52	0.87 (0.58-1.30)	0.50	1.07 (0.69-1.65)	0.76	0.64 (0.38-1.08)	0.09	0.68 (0.39-1.19)	0.18

[1]: Model A is adjusted for covariates (age, sex, their interaction, paternal age, childhood trauma, cannabis use and paternal socioeconomic status)
[2]: Model B is further adjusted for indicators of social distance (level of education, relationship status, living arrangements)
[3]: Model C is further adjusted for cultural distance
[4]: Model D is further adjusted for discrimination

Odds ratios in **bold** are significant (p<0.05)

6.5 Discussion

6.5.1 Summary of main findings

The results confirmed the majority of my hypotheses while for some I did not have sufficient evidence to reject the null-hypothesis. Social and cultural distance explained excess psychosis risk in ethnic minority groups to a large extent (hypothesis 4, Table 6.15). These appeared to be important predictors, but self-perceived discrimination was not, and the model including covariates, social distance and cultural distance was the best-fitting model. Whilst excess odds of all psychotic disorders remained, these were considerably attenuated and only one instead of four ethnic groups still had excess risk of developing any psychotic disorder. Outcomes were broadly similar across diagnostic categories: the final model explained excess odds for all ethnic minorities for both non-affective and affective disorders, and only the Black and North African groups still had excess odds of non-affective disorders. This model initially also appeared to explain excess odds of disorder in religious minorities (hypothesis 7, Table 6.15), however *post hoc* analyses revealed that followers of any religion (including Christianity) were still at increased risk so there appeared to be a degree of confounding by indication in the original reference group. This model attenuated excess risk in the Muslim and other religious groups, but increased risk in the Christian group.

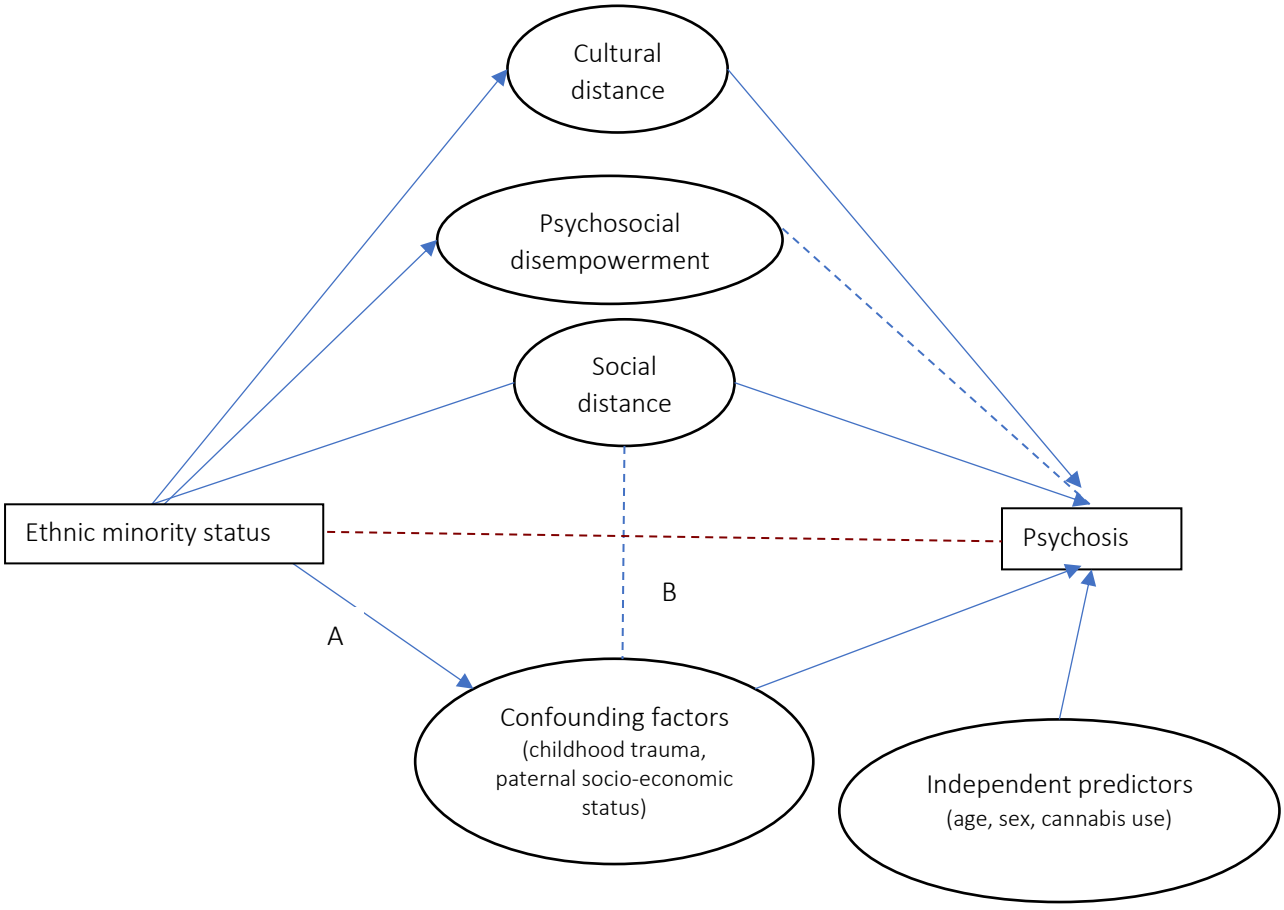
Table 6.15: Reappraisal of hypotheses

Hypothesis	Outcome
1. Ethnic minorities have increased odds of psychosis	Confirmed
2. Ethnic minorities have increased odds of:	
Markers of social distance	Confirmed
Cultural distance	Confirmed
Self-perceived discrimination	Confirmed
3. Increased odds of psychosis are associated with:	
Increased cultural distance	Confirmed
Increased markers of social distance	Confirmed
Increased discrimination	Confirmed
4. The association between ethnic minority status and psychosis is accounted for by:	
Markers of social distance	Partially confirmed
Cultural distance	Partially confirmed
Discrimination	Rejected
5. Religious minorities have increased odds of psychosis	Confirmed
6. Religious minorities have increased odds of:	
Markers of social distance	Partially confirmed
Cultural distance	Confirmed
Self-perceived discrimination	Rejected
7. The association between religious minority status and psychosis is accounted for by:	
Markers of social distance	Partially confirmed
Cultural distance	Partially confirmed
Discrimination	Rejected

Overall, the empirical results from this Chapter have left the conceptual model (Figure 6.1) largely intact, see Figure 6.9 below. There was a strong association between ethnic minority status and social distance: ethnic groups with a higher social distance had their psychosis risk attenuated, and those with a lower social distance saw an increase in risk once social distance was accounted for. The

polychoric correlations showed very little confounding between social distance and the covariates (line B). There was also no clear link between psychosocial disempowerment and psychosis after accounting for other exposures in the model. The multinomial regression revealed that age, sex and cannabis use are best understood as independent predictors of psychosis: there was no clear relationship between these covariates and ethnic minority status. On the other hand, most ethnic minority groups experienced higher levels of childhood trauma and lower paternal socio-economic status, so these were best conceptualised as confounders.

Figure 6.9: Updated conceptual model of the relationship between ethnic minority status, cultural distance, social distance, psychosocial disempowerment, and psychosis.



The blue lines are hypothesised to explain the observed association represented by the red line. Line A represents covariates confounding the association between ethnic minority status and psychosis, and line B represents covariates confounding the association between social distance and psychosis. Dotted lines represent no clear association found in Model C (polychoric correlations).

6.5.2 Strengths and limitations

The findings of this Chapter should be interpreted in light of the methodological strengths and limitations inherent to this work. These analyses were cross-sectional with many measures collected retrospectively, and as such can only establish a probabilistic association and causal claims are not warranted.

6.5.2.1 Chance

Whilst chance could not be ruled out as an explanation for the results obtained in this Chapter, the study benefits from a large sample size (n=2,573) allowing precise estimates of effect sizes. Use of multiple imputation avoided loss of precision due to missing data, use of multilevel modelling allowed correct (wider) estimation of standard errors in nested data, and I used an *a priori* specified modelling approach (with the exception of *post hoc* analyses). Results were broadly similar across disease categories and across main analyses and sensitivity analyses, increasing confidence that findings were not due to chance.

6.5.2.2 Bias

There are, however, a number of biases to consider. The control group might not be entirely representative of the general population. As discussed in Section 3.8, controls were broadly representative of the population at-risk in terms of age, sex and minority status in most settings. However, I was unable to test this for, for instance, socio-economic status. Based on Section 3.8, I believe selection bias is unlikely to be able to fully explain results. The lack of weighting to account for any differences between true population and sample values is a limitation, albeit a minor one. As indicated in Section 3.8, I do not anticipate that including sampling weights changes main conclusions.

The same instruments were used across settings and data-entry was standardised, minimising differential measurement bias, although it could not be determined to what extent these instruments had the same meaning within each catchment area. For instance, tertiary educational attainment might have very different social standing in Brazil compared with the Netherlands (in the Netherlands 36.0% of 25-64 year olds has obtained tertiary education, whereas in Brazil this is 14.8%, presumably giving rise to a higher comparative advantage of obtaining tertiary education in Brazil (OECD, 2017)). A degree of ecological bias remained too: the findings from the sample as a whole might or might not be applicable to any specific catchment area. Relying largely on retrospectively collected measures meant that recall bias was in important consideration: as cognitive deficits are common in psychosis (see Section 1.1), it is possible that cases' recall of events is not as accurate as controls'. However, this was minimised by, where available, using lifetime rather than time-specific exposures.

The further selection criteria of completeness of language data and sufficient clinical information for an OPCRIT diagnosis introduced a small degree of selection bias, and loss of statistical power. The decision on

language data was made *a priori* on the basis that cultural distance was the main novel exposure of interest. Exclusion of those without an OPCRIT diagnosis was done as a quality control mechanism. Nonetheless, this meant that controls and those of the White majority were more likely to be included in the present study. The fact that cases were more likely to be excluded, means that estimates are slightly conservatively biased. The effect of over-exclusion of ethnic minorities depends on their cultural distance: if this is larger than those included, estimated will be biased conservatively and if it is smaller, estimates will be biased away from the null-hypothesis. However, only 74 (2.8%) of participants were excluded at this stage, so it is unlikely this changed main conclusions. Missing data in the final sample was not randomly distributed: cases were more likely to have missing data on paternal age, cannabis use, childhood trauma, paternal socioeconomic status, social distance and discrimination. Use of multiple imputation ensured these participants could be retained and lack of precision could be avoided.

Risk of bias due to missing data depends on the underlying reasons for this missing data. As cases were more likely to have missing data than controls, data was not missing completely at random (MCAR), meaning that complete case analysis might give biased results. When data is missing not at random (MNAR), systematic differences between observed and missing data remain even after observed date are taken into account, and when data is missing at random (MAR), systematic differences between observed and missing data can be explained by the observed data. It is impossible to identify the missingness mechanism (MAR or MNAR) on the basis of the observed data alone (Sterne et al., 2009). A few common pitfalls in reporting and assessing bias have been avoided: I included the outcome variable and a wide range of auxiliary variables in the imputation method, and I used a method (*ice*) which was able to impute multiple categorical variables. I have also reported missingness for all variables by different diagnostic categories (Tables 6.2 and 6.3) and have assessed if missing data was associated with case-control status, age, sex and ethnic minority status. Whilst there were some differences in missingness by age and ethnic minority status (and case-control status), the similarity between the imputed analyses and sensitivity analyses on the bases of complete cases suggest that it is plausible that data is MAR. If this assumption holds, estimates from imputed data are unbiased (Sterne et al., 2009).

6.5.2.3 Confounding and reverse causality

There are a number of further limitations resulting from this work relating to confounding. Whilst I attempted to adjust for as many known confounders as possible, data on family history of psychosis was not yet available and could not be included. A degree of unobserved confounding was therefore a likely limitation of this work. The possibility of unknown confounders (which, by their very nature, are impossible to predict) was also not eliminated. There was a known relationship between pre-morbid educational attainment and IQ and psychotic disorders (Khandaker et al., 2011), and such neurodevelopmental origins of disease have invariably introduced a degree of reverse causality. It was possible this extended to other

measures of social distance (relationship status and living arrangements), despite best efforts to minimise this by using lifetime exposures.

The constructs used in this Chapter were narrower than the concepts they attempted to measure as proposed in Chapter 5, and it was impossible to assess the extent to which they were valid measures of these broader constructs, and the degree of residual confounding that was the result of this. There is a broader explanatory gap between sociological literature and epidemiological practice, and the implications of this for further research are briefly discussed in Chapter 8. One particular concern is discrimination as a measure of psychosocial disempowerment. The latter is a broad paradigm of chronic stress (Fisher & Baum, 2010; Link & Phelan, 1995; Marmot, 2015; Wheaton & Montazer, 2017) (see Section 5.4.3.2), whereas discrimination as operationalised here is a much narrower measure of event stress. This discrepancy may have contributed to the null finding in the multivariable model.

Finally, confounding and reverse causality are possible alternative explanations for the association between religion and psychotic disorders. Some suggest that both religiosity and the formation of delusions and hallucinations depend on similar cognitive strategies (Dein & Littlewood, 2011). There is some evidence for a bias towards external attribution styles in psychosis (Garety et al., 2001; Van Os, Linscott, Myin-Germeys, Delespaul, & Krabbendam, 2009), and this might also lend itself well to religious explanations of daily-life phenomena, confounding the association between religion and psychosis. Being religious is also not a fixed identity, and I had no data on length of religious affiliation, so it is possible that those who were in the prodromal stage of illness were more likely to stay or become religious. Engaging in spiritual activities has also been reported as a helpful coping strategy employed by individuals suffering from a psychotic disorder (Cohen et al., 2017), increasing the probability of reverse causality as an explanation.

6.5.3 Interpretation of main findings

This Chapter has two main findings: the sociocultural distance model, for the first time, explains excess psychosis risk in ethnic minorities, and this model could not be directly extended to religious minorities. It appeared that structural inequalities as captured by social and cultural distance were the main drivers of excess psychosis risk in ethnic minorities, whereas in the multivariable model, discrimination was no longer significantly associated with increased odds of developing psychosis. In Chapter 5, I hypothesised that the more structural drivers of social distance and cultural distance (the stressors), would correspond with increased levels of psychosocial disempowerment (the stress), and that each of these would be associated with increased risk of psychosis. These present results could be interpreted in a number of ways: structural drivers or stressors are more predictive of psychosis risk than subjective stress, structural drivers are more chronic and thereby easier to measure than acute and time-limited stress, the measure of psychosocial disempowerment was inadequate, or the modelling approach was insufficient to disentangle these

concepts fully. It appears that a partial explanation is provided by the measure of discrimination used not being appropriate approximation of the concept of psychosocial disempowerment. However, I mainly interpret these findings to indicate that the crude association is explained fully by the structural inequities and disadvantages captured by the sociocultural distance model. The 'otherness' and structural strain (to return to sociology) captured by this model appear to be predictive of psychosis risk in ethnic minorities, albeit via a different mechanism than direct self-perceived discrimination.

The results of the post hoc analyses examining religion and ethnicity suggest that the original theory proposed to explain excess risk in ethnic minorities might not directly extend to other minorities, and suggest that religious affiliation might be an important independent risk factor for psychotic disorders. The finding that, after accounting for ethnicity, the model attenuated risk in the Muslim and other group is in line with recent European Union reports on discrimination and marginalisation of Muslims (European Union Agency for Fundamental Rights, 2017). For Muslims, crude excess risk was attenuated substantially following full adjustment, suggesting an aetiological role for their 'outsider status' as approximated by their ethnicity, and cultural and social distance. However, the findings on religion were not uniform. Christians also faced crude excess risk and this increased following full adjustment. This is less straightforward to interpret: it suggests perhaps a protective effect of lower cultural and social distance, but also illustrates that even those from a non-persecuted religion have increased odds of developing psychosis.

A full contextualisation of findings in light of Chapter 5 can be found in Chapter 8.

6.5.4 Comparison with existing literature

The increased psychosis risk in ethnic minorities was in line with existing literature (Anderson et al., 2015; Bourque et al., 2011; Kirkbride, Hameed, Ioannidis, et al., 2017; Tortelli et al., 2014), particularly the high risk in the Black and Mixed groups (Bourque et al., 2011; Kirkbride, Hameed, Ioannidis, et al., 2017). The prior evidence for the observed higher risk in the North African group was more mixed, with a higher incidence of schizophrenic disorders reported for the Moroccan minority in the Netherlands (Veling et al., 2006), but no excess FEP admissions in the North African group in Paris (Tortelli et al., 2014). I did not demonstrate an excess risk in the Asian group, but recent data from East Anglia suggest there might be marked heterogeneity within this group, with Pakistani (IRR: 2.93, 95%CI: 1.80-4.74) and Bangladeshi (IRR: 2.34, 95%CI: 1.05-5.23) groups at increased risk, but not the Indian group (Kirkbride, Hameed, Ioannidis, et al., 2017). This has previously been observed in East London (Coid et al., 2008; Kirkbride et al., 2008). In the EU-GEI study, this group further includes the Southeast and East Asian minorities, which were shown not to have increased incidence in Ontario, Canada (Anderson et al., 2015). No evidence of excess risk was found in the White other group, which was partially in line with existing literature: a systematic review did find an excess risk in White ethnic minorities (Bourque et al., 2011), but more recent studies were unable to replicate this (Anderson et al., 2015; Kirkbride, Hameed, Ioannidis, et al., 2017).

Findings on social distance are broadly in line with both the study on cumulative social disadvantage (Morgan et al., 2008) and a wider literature. Social isolation (Hare, 1956; Kohn & Clausen, 1955; Morgan et al., 2008) or lack of social support (Gayer-Anderson & Morgan, 2013) are known risk-factors, and the association with low educational attainment (Morgan et al., 2008) and low premorbid IQ (Khandaker et al., 2011) is also well-established. Findings on cultural distance were novel, although based on a wider literature indicating an important role for language in the forming of a social identity (Candelo et al., 2016; Koczan, 2016) and decreased vulnerability to disease (Akiyama, 1996). The null-finding on discrimination in the multivariable model sat within an equipoised literature: there was an association between discrimination and risk of psychosis on a population-level (Veling et al., 2007), but in the same study setting this was not found at an individual level (Veling, Hoek, et al., 2008). An English study did find an association, but this was not based on a clinical population nor used clinical diagnoses (Psychosis Screening Questionnaire), and specifically examined personal instances of verbal and physical abuse (which was narrower than the discrimination measure used here), and perceived societal discrimination (which was broader than the discrimination measure used here)(Karlsen et al., 2005).

Findings on covariates were broadly in line with the existing literature. Childhood trauma was a known risk factor (Bendall et al., 2008; Matheson et al., 2012; Varese et al., 2012), as were low socioeconomic status (Marwaha & Johnson, 2004; Reininghaus et al., 2008; Werner et al., 2007), cannabis use (Manrique-Garcia et al., 2012; Moore et al., 2007), younger age, male sex, and their interaction (Häfner et al., 1993; Thorup et al., 2007). However, our null-finding on paternal age was unexpected, considering strong previous evidence for an association between increased paternal age and risk of disorder (Sipos et al., 2004; Zammit et al., 2003).

Whilst the present model included a similar range of indicators and found similar results on these indicators to Morgan's cumulative social disadvantage study (Morgan et al., 2008), the present study casted a wider net. The findings from both studies (that indicators of social disadvantage were more common in the Black Caribbean group, and that social and cultural distance largely explained excess risk in minority groups) are complementary. The present study is consistent with a socio-developmental model of psychosis in minorities (Morgan, Charalambides, Hutchinson, & Murray, 2010), although due to its cross-sectional nature and focus on the social environment and not on developmental markers, could not provide conclusive evidence for it. This study provided tentative evidence against the social defeat hypothesis (Selten et al., 2013), which crucially hinged on subjective interpretation of experiencing a situation as defeating. The only subjective indicator used (self-perceived discrimination) was not found to be associated with excess psychosis risk in a multivariable model (although there might be alternative explanations), indicating that structural drivers might be more important.

6.5.5 Conclusion

This study demonstrated that, following adjustment for covariates, social and cultural distance explained a large proportion of excess psychosis risk in ethnic minority groups. Whilst minorities retained excess risk for all psychotic disorders, odds ratios were significantly attenuated and only remained significantly increased in the Black group. Findings were similar for affective and non-affective disorders, with the best-fitting model explaining excess odds across all ethnic minority groups, but a significant effect remained for the Black and North African group (non-affective disorders only). Further inclusion of discrimination did not improve model fit, nor affect point estimates. The model also explained excess risk in religious minorities, although *post hoc* analyses revealed that followers of all religion (including Christianity) had excess risk of psychosis, and this was not explained by the current analytical model. Findings from this Chapter suggest that indicators of structural disadvantage or inequality are crucial to understanding excess risk of psychotic disorders in ethnic minorities.

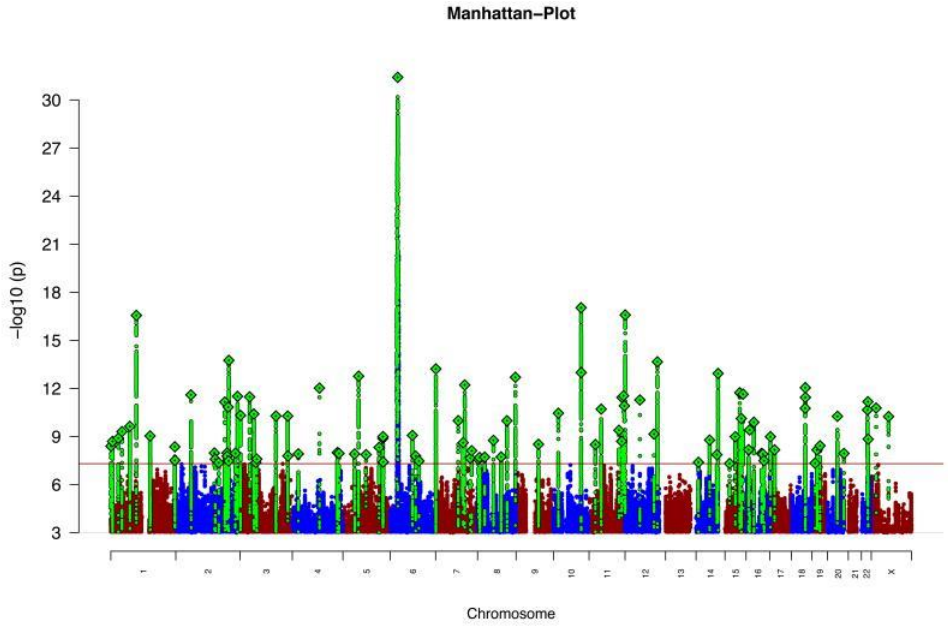
Chapter 7 - An exploration of the role of genetic distance in explaining psychosis risk.

7.1 Background

In the previous Chapter I demonstrated that the excess of psychotic disorders in ethnic minority groups was partially accounted for by confounders and independent predictors, indicators of social distance and cultural distance. As mentioned in Section 1.2, psychotic disorders are also known to have a genetic determinant (Schizophrenia Working Group of the Psychiatric Genomic Consortium, 2015), which is non-Mendelian and involves a large number of risk genes each conveying a small effect (see below). It is unclear whether the increased risk in ethnic minorities can be explained genetically (Vassos et al., 2017). Those of African origin appear to have increased genetic diversity compared with White Europeans, and this might increase their risk of disorder (Tishkoff et al., 2009). On the other hand, many risk genes for psychosis predate African-European divergence, so risk would be expected to be equivalent across population groups (de Candia et al., 2013; Ioannidis, Ntzani, & Trikalinos, 2004). Furthermore, as discussed in Section 1.4, risk of disorder in any particular ethnic group appears to be, at least partially, context-dependent. The EU-GEI study was conceived at a time when it had become clear that the 'schizophrenia gene' did not exist, and that psychotic disorders were polygenic in nature. At the time, it was hypothesised that genetic variation operated in tandem with socio-environmental risk factors, increasing the risk of psychosis for those who have both a genetically increased risk and are exposed to socio-environmental risk factors (Kirkbride, 2009) such as individuals from an ethnic minority background. A gene-environment interaction was also considered possible, where ethnic minorities and other at-risk groups have an increased environmental burden and consequently need a lower genetic threshold for the development of disorder (Cooper, 2005).

Since the conception of the EU-GEI study understanding of psychiatric genetics has moved on. As was mentioned in Chapter 1, the most recent systematic analysis of available genome-wide association studies (GWASs; observational studies of a genome-wide set of genetic variants designed to investigate if any variant is associated with a trait) of schizophrenia identified 108 independently associated loci (places on the genome (Schizophrenia Working Group of the Psychiatric Genomic Consortium, 2015) (Figure 7.1 below)). Each of these explained a small part of the genetic variance, with odds ratios commonly around 1.1, and never exceeding 1.2 (Schizophrenia Working Group of the Psychiatric Genomic Consortium, 2015).

Figure 7.1: Manhattan plot of loci associated with risk of schizophrenia (Schizophrenia Working Group of the Psychiatric Genomic Consortium, 2015).



Every vertical peak surpassing the horizontal line indicates a statistically significant association ($p<5*10^{-8}$).

A common way of combining these small effect sizes (and improving predictive ability) is via so-called polygenic risk scores (PRSs; a weighted sum of genetic variance of disease risk), where the cumulative variance is summarised in a discovery sample, and subsequently used an independent variable in testing samples (Dudbridge, 2013; Janssens et al., 2011). The PRS for schizophrenia was derived from the aforementioned systematic analysis of GWASs (Schizophrenia Working Group of the Psychiatric Genomic Consortium, 2015). This PRS contains genes for known neurobiological mechanisms involved in psychosis, such as dopamine-2 receptors (Howes et al., 2017) as well as genes involved in glutamatergic neurotransmission and synaptic plasticity (Howes & Kapur, 2014). It also maps onto genes involved in immune function, which is implicated in psychosis (Khandaker, Pearson, et al., 2014), and authors demonstrated overlap with genes implicated in the onset of autism-spectrum disorder and intellectual disability, increasing the evidence base there might be common neurodevelopmental origins to these disorders (Sullivan, Daly, & O'Donovan, 2012). Finally the authors tested the validity of the PRS by dividing it up into deciles, and demonstrating a dose-response improvement in predictive ability (Schizophrenia Working Group of the Psychiatric Genomic Consortium, 2015). This dose-response relationship was less clear for an incidence sample compared with samples of chronically ill participants, suggesting that the PRS might be particularly predictive of chronicity. Furthermore, the PRS only explains a small part of genetic predisposition to schizophrenia, with the remainder most likely explained by as of yet undiscovered genetic loci, rare variants and gene-gene and gene-environment interactions (Harrison, 2015).

In the context of ethnicity, the PRS has some notable further limitations, including its low predictability in a non-White sample (Vassos et al., 2017). This is due to the fact that most of the discovery samples used and synthesised to obtain the PRS predominantly include White individuals (Schizophrenia Working Group of the Psychiatric Genomic Consortium, 2015). This 'population stratification' effect is not unique to schizophrenia genetics; allele frequency differences between cases and controls due to systematic ancestry differences are known to have the potential to cause spurious correlations in disease studies (Price et al., 2006). It is possible, to an extent, to account for these systematic ancestry differences using results from a principal component analyses (PCAs) modelling genetic ancestry by including these principal components as covariates. The PRS and PCA scores are based on the same GWAS, but used to indicate two different exposures: the PRS is an indicator of genetic risk for disorder whilst the PCA scores are an indicator of genetic differences in ancestry. Data from Southeast London suggested that even after controlling for genetic ancestry using these PCAs, the discriminatory power of the PRS was still lower in the African group (R^2: 1.1%) than in the White sample (R^2: 9.4%) (Vassos et al., 2017). This presents a problem when examining the causes of higher rates of psychosis in ethnic minorities: it is known that genetics play an important role in psychosis, but using the PRS introduces a differential bias as it has different power to discriminate between cases and controls between different ethnic groups.

Nonetheless, as I am interested in how the social circumstances of ethnic minorities influenced their risk of developing psychosis, it is possible to use the results from these genetic ancestry PCAs for preliminary analyses, and not the PRS as such. I am able to test if those at an increased genetic distance from the white majority population (with corresponding higher PCA scores) are at increased risk of psychosis. I am also able to test if sociocultural distance model developed in the previous two Chapters could explain any excess risk, using the same methodology as in Chapter 6. I hypothesise that genetic distance isn't as good a predictor of case-status as self-ascribed ethnicity, and that there is no association between genetic distance and psychosis risk within ethnic groups. This is because I think it is the social reality of being an ethnic minority, and not a genetic marker of ancestry, that determines increased psychosis risk. In this Chapter, I conduct some preliminary analyses to explore this.

7.2 Hypotheses

This Chapter sought to investigate the following hypotheses:

1. A higher genetic distance (higher PCA score) would be associated with increased odds of developing psychosis across the sample.
2. This increased risk of psychosis associated with genetic distance would be explained by the inclusion of covariates, indicators of social distance, cultural distance and discrimination.
3. Genetic distance would be a poorer predictor of psychosis risk than self-ascribed ethnicity.
4. There would be no association between genetic distance and psychosis risk within ethnic groups.

7.3 Methods

This Section details only the methodology specific to this Chapter. The overall EU-GEI study design, recruitment strategy and methodology are found in Chapter 3, and a description of the exposures and covariates of interest, as well as details of multiple imputation, were given in Section 6.3.

I am using genetic distance (PCA scores on genetic ancestry) only as an exposure, and not the PRS, due to the low predictive ability of the PRS in non-White samples. Therefore, the methodology described here is specific to genetic distance.

7.3.1 Data collection and management

Individuals who had consented to participate in the full EU-GEI study were invited to give two 9ml blood samples for DNA extraction. Participants unwilling or unable to give blood were offered an alternative sampling technique (Oragene, from saliva) which was considered a reliable method of extracting DNA (Abraham et al., 2012; Rylander-Rudqvist, Håkansson, Tybring, & Wolk, 2006). Blood or saliva samples were taken at approved clinical research facilities within each catchment area by an experienced research nurse or trained researcher, were fully anonymised and sent to Cardiff University for genotyping (European Network of National Schizophrenia Networks Studying Gene-Environment Interactions, 2013).

7.3.2 Sample selection

The sample in this Chapter was restricted to case-control participants from the two English sites (Southeast London and Cambridgeshire and Peterborough). I chose to do this because of the variance in genetic ancestry in the White majority population between countries (see Section 7.3.3 below). In addition to participants with incomplete language data (see Chapter 6), participants with incomplete genetic data were excluded from the present analyses.

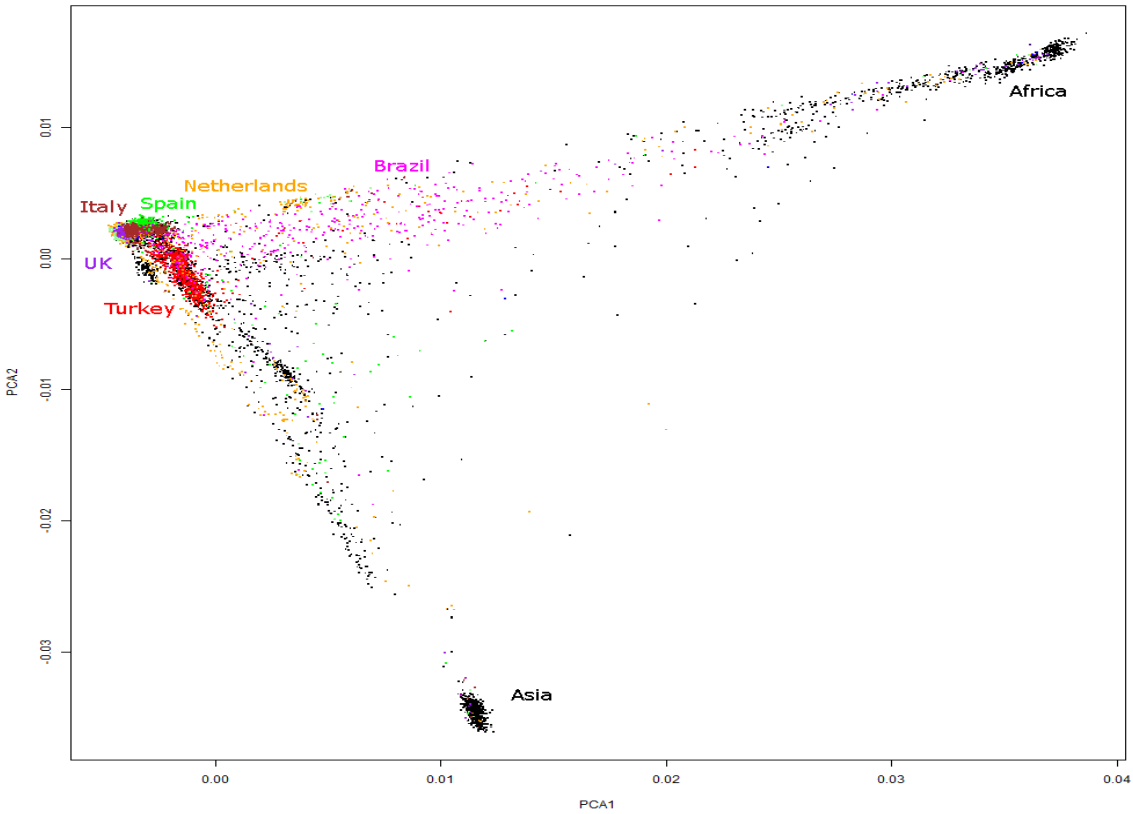
7.3.3 Definition of genetic distance

The creation of genetic distance was done in two stages: genotyping the DNA and computing the PCA using this genotyped data. This took place in Cardiff, and I have been sent the PCA results. Genotyping was performed using custom Illumina HumanCoreExome-24 BeadChip genotyping arrays that contained probes for 570,038 variants (281,725 common variants, 265,919 rare exonic variants and 22,394 custom variants) (Illumina, 2017). These genotypes were subsequently analyses using PLINK, an open source whole genome association analysis toolset (Purcell et al., 2007). Using this toolset, data from the largest dataset characterising the functional and geographic spectrum of human genetic variation, the 1000 genomes project (1000 Genomes Project Consortium, 2015), was used as the reference sample to compute principal components for genetic ancestry. The results from this reference sample were then used to score the EU-GEI dataset.

Increased genetic distance (away from the White majority) in this study refers to increased genetic African ancestry, operationalised as the score on PCA1. As with any principal component analysis, PCA1 explains the largest amount of variance in the sample. The exact meaning of PCA1 therefore differs depending on the reference sample used, but in this sample it is African ancestry, more specifically Yoruban (Nigerian) ancestry (1000 Genomes Project Consortium, 2012). PCA2 refers to Asian (Chinese) ancestry in the graphs presented in this Chapter.

Exact genetic ancestry differed between countries, even in the ethnic majority. This is illustrated in Figure 7.2 below, which has been produced by Dr Alex Richards at Cardiff University and maps the genetic ancestry in the ethnic majority in various countries included in the EU-GEI project, as well as in an African and Asian (Chinese) reference sample from the 1000 genomes project (1000 Genomes Project Consortium, 2015).

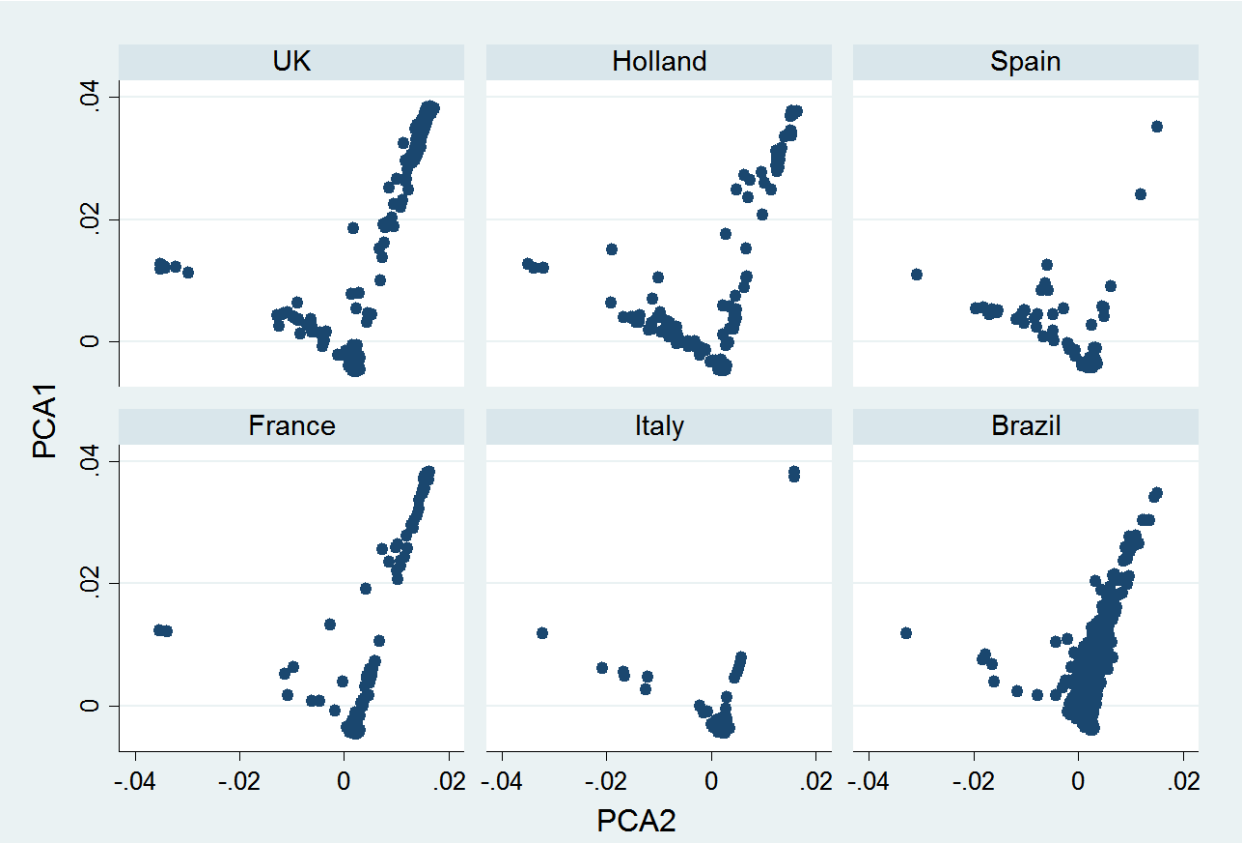
Figure 7.2: African versus Asian ancestry in various EU-GEI countries



This graph represents the countries included in work packages 2 and 6 (GxE vulnerability and severity), as well as the European members of the 1000 genomes reference sample and the Asian and African reference samples from this dataset (1000 Genomes Project Consortium, 2012). Individuals from different EU-GEI countries have been given different colours (the Netherlands and the UK, for instance). PCA1 (X-axis) refers to African Ancestry, and PCA2 (Y-axis) refers to Asian ancestry.

Because of this genetic ancestral heterogeneity across the sample (see also Figure 7.3 below), it was inappropriate to treat the EU-GEI study as one sample, even when taking a multilevel modelling approach to account for the clustering of data at country-level. Therefore, I decided to isolate the EU-GEI sample from Cambridgeshire and Peterborough and Southeast London (the English catchment areas) for the exploratory analyses presented in this Chapter. Despite the fact that the ethnic composition of the population in these catchment areas is very different, the White majority in Southeast London had a similar median genetic distance (-0.64, IQR: -0.64 - -0.63) compared with Cambridgeshire and Peterborough (-0.64, IQR: -0.65 - -0.63; Mann-Whitney U-test: 1.67, p=0.09).

Figure 7.3: Scatterplot of African versus Asian ancestry, by EU-GEI country.



7.3.4 Outcome, exposures and covariates

The outcome of interest was restricted to the all psychotic disorders category (ICD10: F20-F33). Due to the limited sample size it was not feasible to examine data for non-affective and affective disorders separately.

All covariates and most of the exposures of interest were identical to those used in Chapter 6 (Section 6.3), apart from ethnicity and the addition of genetic distance. An overview of the outcome, exposures and covariates is found in Table 7.1 below.

Ethnicity was recoded to three categories: White British, Black and 'other'. Those whose ethnicity was recorded as White native were retained in the White British group, those whose ethnicity was recorded as Black were retained in the Black ethnic group. All other groups were amalgamated (Mixed, Asian, Other,

White other) into the 'Other' group due to small numbers, with the exception of the North African group. This group was divided: those whose more detailed ethnicity variable and/or country of birth identified them as Black African were added to the overall Black category whereas those who were considered of Arab ethnic origin were included in the other category.

Genetic distance was operationalised as PCA1 scores (see above). The values for PCA1 were z-standardised to have a mean of zero and a standard deviation of one, in order to aid interpretability, with a higher score indicating an increased genetic distance.

Table 7.1: Outcomes, exposures and covariates

Outcome	Exposures	Covariates
All psychotic disorders (F20-F33)	Ethnicity	Age
	White majority	
	Black	Sex
	Other	
		Paternal age
	Genetic distance	
		Childhood trauma
	Indicators of social distance	
	Level of education	Lifetime cannabis use
	Living arrangements	
	Relationship status	Paternal socioeconomic status
	Cultural distance	
	Linguistic distance	
	Fluency	
	Psychosocial disempowerment	
	Discrimination	

7.3.5 Missing data

From the sample used in Chapter 6, only the participants from the two English catchment areas who were genotyped were retained. The same method of multiple imputation was used as described in Section 6.3. Where missing, I imputed data on the following variables: level of education, living arrangements, relationship status, discrimination, age, sex, ethnicity, paternal age, childhood trauma and lifetime cannabis use. Using a user-written command in Stata (*ice*), I imputed 25 datasets using chained equations. I used all variables in the analyses as well as a number of auxiliary variables.

7.3.6 Statistical analyses

After describing the distribution of exposures and covariates by case-control status, I plotted genetic distance by case-control status and by broad ethnic group (using Mann-Whitney U test and Kruskal-Wallis test to assess differences in median genetic distance). Using a simple linear regression across the imputed datasets, I tested whether there was an association between genetic distance and the covariates and remaining exposures. Subsequently, using logistic regression, I built a multivariable model across the imputed datasets to examine the effect of the addition of covariates, social distance, cultural distance and discrimination in the same way as in Chapter 6:

- A crude association between psychotic disorders and genetic distance;
- Model A further included covariates (Table 7.1);
- Model B further included indicators of social distance;
- Model C further included cultural distance, and
- Model D further included self-perceived discrimination

As explained in Section 6.3, it was impossible to rely on likelihood ratio-tests or compute the Akaike Information Criterion (or similar measures) in multiply imputed datasets and Nagelkerke's pseudo-R^2 was used to assess model-fit (Nagelkerke, 1991).

Using the complete-case dataset, a logistic regression model was built with only genetic distance and ethnicity as exposures, to assess their relative importance vis-à-vis each other (hypothesis 3).

The association between genetic distance and psychosis was tested within individual ethnic groups using univariable logistic regression on the complete-case data set.

The main analyses (univariable logistic regression and multivariable model building) were repeated on the complete case sample by means of sensitivity analyses, in the same way as in Chapter 6.

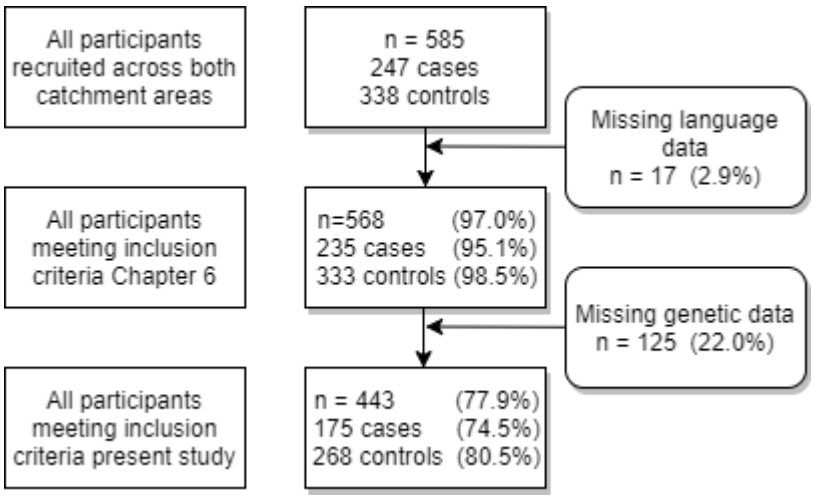
7.4 Results

7.4.1 Sample description

7.4.1.1 Case retention and missing data

A total of 443 participants (175 cases and 268 controls) were included in the sample for this Chapter. Five hundred and eighty-five individuals participated in the EU-GEI study in England, but 14 (2.4%) had been excluded previously for missing language data. A further 60 cases (25.5%) and 65 controls (19.5%) were not genotyped, and were also excluded from these analyses (Figure 7.4 below). This could occur if a participant was unwilling or unable to give DNA via either blood or saliva, or if the DNA obtained was not of sufficient quality. Data was insufficiently detailed to differentiate between these two possibilities.

Figure 7.4: Flow chart of case retention



Cases were not more likely to have missing genetic data (χ^2: 2.9, p=0.09), and whilst those with missing genetic data were not different in sex, they were younger than those with genetic data: median age 28 (IQR: 24-34) and 31 (IQR: 24-42) respectively (Mann-Whitney U-test: 2.7, p<0.01). They were also less likely to be White British with only 32.8% falling into this ethnic category, compared to 51.5% of those with genetic data (χ^2: 14.6, p<0.01).

There were missing data in the remaining covariates and exposures (Table 7.2 below). Cases were more likely to have missing data on paternal age (χ^2: 26.6, p<0.01), childhood trauma (χ^2: 39.0, p<0.01) and discrimination (χ^2: 9., p<0.01) but case-control status was not associated with missingness in other exposures.

Table 7.2: Missing data by case-control status.

Variable	Controls (%)	Cases (%)	X^2; p-value
Paternal age	13 (4.9)	36 (20.5)	26.6; p<0.01
Childhood trauma	2 (0.7)	28 (16.0)	39.0; p<0.01
Cannabis use	4 (1.5)	3 (1.7)	0.0; p=0.86
Paternal socioeconomic status	1 (0.4)	1 (0.6)	0.1; p=0.76
Level of education	1 (0.4)	2 (1.1)	0.9; p=0.33
Relationship status	0 (0.0)	2 (1.1)	3.1; p=0.08
Living arrangements	8 (3.0)	5 (2.9)	0.0; p=0.94
Discrimination	8 (3.0)	17 (9.7)	9.0; p<0.01

NB: this table only includes variables with a degree of missing data.

Those with missing paternal age had a higher genetic distance (Mann-Whitney U-test: -2.81, p<0.01), but genetic distance was not associated with missingness in remaining exposures (Table 7.3 below)

Table 7.3: Associations between genetic distance and missing data

Variable	Complete[1] median (IQR)	Missing[2] median (IQR)	MWU[3]; p-value
Paternal age	-0.63 (-.064 - -0.04)	0.57 (-0.64 – 1.78)	-2.81; p<0.01
Childhood trauma	-0.62 (-0.64 – 0.38)	-0.62 (-0.64 – 0.88)	-0.53; p=0.60
Cannabis use	-0.62 (-0.64 – 0.45)	-0.63 (-0.63 – 0.62)	-0.51; p=0.62
Paternal socioeconomic status	-0.62 (-0.64 – 0.48)	-0.63 (-0.64 - -0.62)	0.65; p=0.51
Level of education	-0.62 (-0.64 – 0.40)	1.25 (-0.63 – 1.50)	-1.15; p=0.25
Relationship status	-0.62 (-0.64 – 0.41)	0.52 (-0.63 – 1.66)	-0.80; p=0.42
Living arrangements	-0.62 (-0.64 – 0.57)	-0.62 (-0.64 - -0.14)	-0.35; p=0.73
Discrimination	-0.62 (-0.64 – 0.39)	-0.63 (-0.64 – 1.78)	-0.94; p=0.35

[1] Refers to participants with complete data on the variable of interest.
[2] Refers to participants with missing data on the variable of interest.
[3] Mann-Whitney U-test

NB: this table only includes variables with a degree of missing data.

As described in Section 3.8, controls for both Southeast London and Peterborough were broadly representative of the population at-risk in terms of sex, although for this study controls from any ethnic minority were oversampled ($\chi^{2:}$ 35.1, p<0.01). Controls from younger age groups were oversampled compared with the population at-risk (χ^2: 54.0, p<0.01).

7.4.1.2 Sample characteristics.

Cases were younger (median age: 27, IQR: 23-34) than controls (34, IQR: 27-48) (Mann-Whitney U-test: 7.17, p<0.01), but were not statistically more likely to be men (χ^2: 3.94, p=0.06), although this might be due to low power. There was no difference in paternal age between cases and controls (Mann-Whitney U-test: 1.51, p=0.13) (Table 7.4). Cases had higher levels of childhood trauma (Mann-Whitney U-test: -8.4, p<0.01), were more likely to have smoked cannabis (χ^2: 11.6, p<0.01), and their fathers had a lower socioeconomic status (χ^2: 27.8, p<0.01) compared with controls (Table 7.4). Cases had a lower level of education compared with controls (24.2% (n=33) of cases obtained a university degree, compared with 56.0% (n= 150) of controls; χ^2: 69.2, p<0.01). Cases were also less likely to have ever been in a relationship (χ^2:35.4, p<0.01, Table 7.4) or to have ever lived with someone other than their parents (χ^2:6.1, p=0.05). They were more likely to have a degree of cultural distance (χ^2:4.9, p=0.03). Cases also reported a higher level of discrimination with a median of 1 event (IQR:0-3) compared with 0 (IQR: 0-1) in controls (Mann-Whitney U-test: -4.9, p<0.01).

Table 7.4: Distribution of exposures and covariates by case-control status.

	Controls (%) /Median (IQR)	Cases (%) /Median (IQR)	χ^2 / MWU[1] test
Age			MWU: 7.17
Median (IQR)	34 (27-48)	27 (23-34)	p<0.01
Sex			
Men	135 (50.4)	102 (59.4)	χ^2: 3.94
Women	122 (49.6)	71 (40.6)	P=0.06
Paternal age			MWU: 1.51
Median (IQR)	31 (28-35)	30 (26-35)	P=0.13
Childhood trauma			MWU: -8.4
Median (IQR)	36 (31-42)	46 (38-59)	P<0.01
Cannabis use			
Yes	151 (56.3)	126 (72.0)	χ^2: 11.6
No	113 (42.2)	46 (26.3)	P<0.01
Paternal socioeconomic status			
Professional	130 (48.5)	55 (31.5)	χ^2: 27.8
Intermediate	59 (22.0)	36 (20.6)	P<0.01
Lower	42 (15.7)	40 (22.9)	
Routine	32 (11.9)	25 (14.3)	
Never worked	0 (0.0)	2 (1.1)	
Not classified	4 (1.5)	16 (9.1)	
Genetic distance			MWU: -3.5
Median (IQR)	-0.58 (-0.64- -0.13)	-0.58 (-0.64-1.46)	<0.01
Ethnicity			
White British	159 (59.4)	69 (39.4)	χ^2: 17.2
Black	56 (20.9)	59 (33.7)	P<0.01
Other	53 (19.8)	37 (29.9)	
Level of education			
Postgraduate	64 (23.9)	11 (6.3)	χ^2: 69.2
Undergraduate	86 (32.1)	22 (18.9)	P<0.01
Vocational	41 (15.3)	41 (23.4)	
Tertiary	47 (17.5)	25 (14.3)	
School qualifications	22 (8.2)	31 (17.7)	
No qualifications	7 (2.6)	32 (18.3)	
Relationship status			
Yes	240 (89.5)	118 (66.4)	χ^2: 35.4
No	28 (10.4)	55 (31.9)	P<0.01
Living arrangements			
Yes	241 (89.9)	145 (82.8)	χ^2: 6.1
No	19 (7.1)	25 (14.3)	P=0.05*
Cultural distance			
Yes	221 (82.5)	129 (73.7)	χ^2: 4.9
No	47 (17.5)	46 (26.3)	P=0.03
Discrimination			MWU: -4.9
Median (IQR)	0 (0-1)	1 (0-3)	P<0.01

[1]: MWU: Mann-Whitney U-test, used to assess differences in median across groups.

Broad ethnic group

A total of 159 controls (59.4%) and 69 cases (39.4%) were White British (Table 7.4). The ten (2.3%) participants who had a mixed ethnic background, the nineteen (4.3%) who were Asian, the nine 'Others' (2.0%) and the 59 (13.3%) people from any other White background were combined into one 'Other' category. In the original classification used in Chapter 6, 54 controls (20.1%) and 58 cases (33.1%) self-identified as Black. Four participants (0.9%) were from a North African background: three of these were Black African and assigned to the Black ethnic category, the final North African participant was of Arab

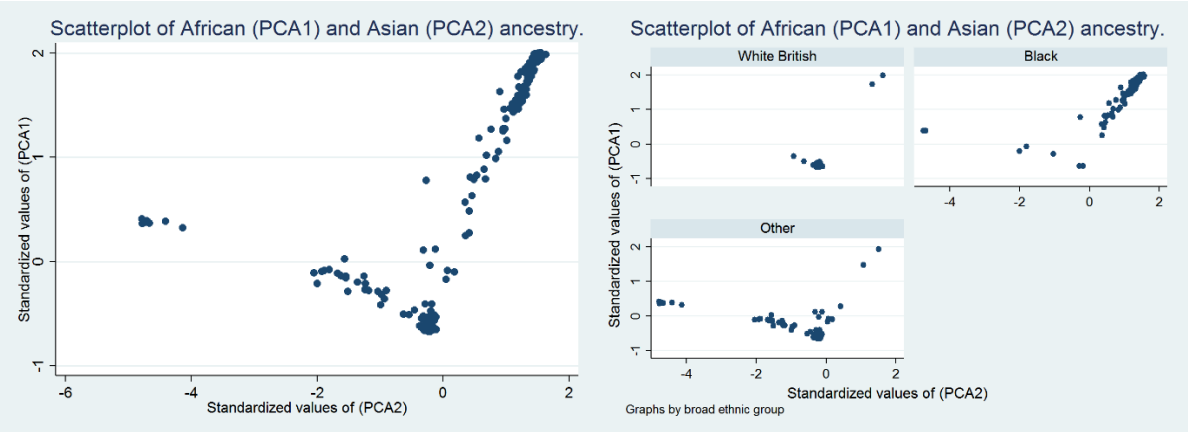
ethnic origin, and as such was assigned to the 'Other' category. The final distribution of broad ethnic group by case-control status can be found in Table 7.4 above. Cases were less likely to be White British than controls (χ^2: 17.2, p<0.01). The small Asian sample size meant that Asian ancestry could not be investigated as a risk factor.

Genetic distance

Figure 5A below plots the standardised values of African (PCA1) and Asian (PCA2) ancestry in this sample. There are two relatively distinct population groups (indicated by a high density of data points). One at the top right comprises both higher African and Asian ancestry, corresponding with the Black ethnic group. The other closer to the null value for both ancestries, corresponding with the White British group. The scattered dots represent participants of more mixed ethnic heritage, as can be seen in Figure 5B.

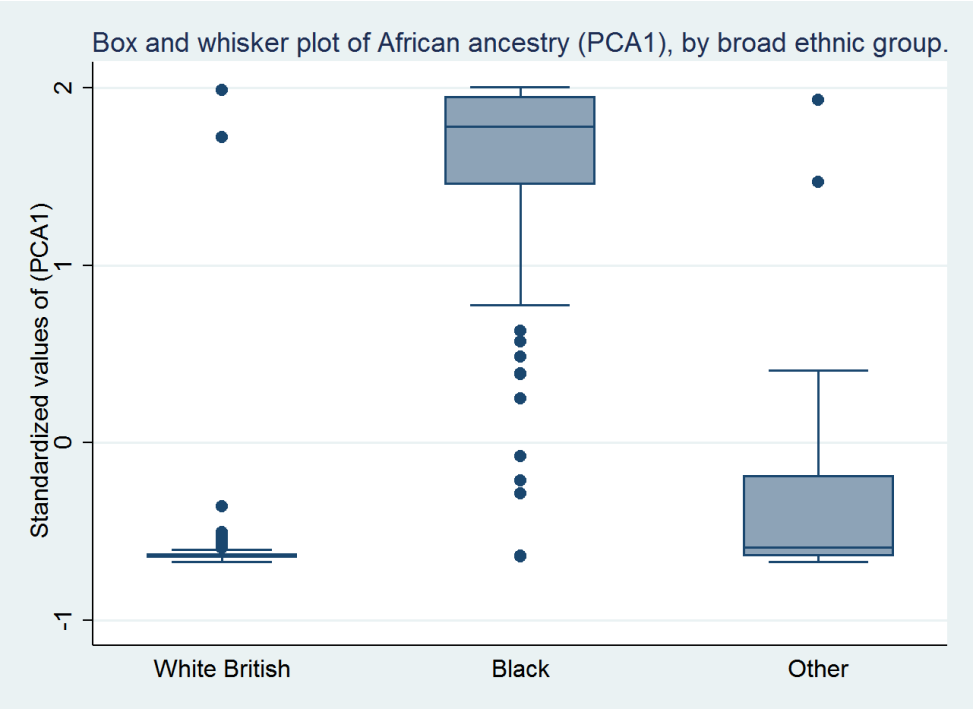
Figure 7.5: Scatterplot of African and Asian ancestry, by broad ethnic group.

A: Overall sample *B: By broad ethnic group*



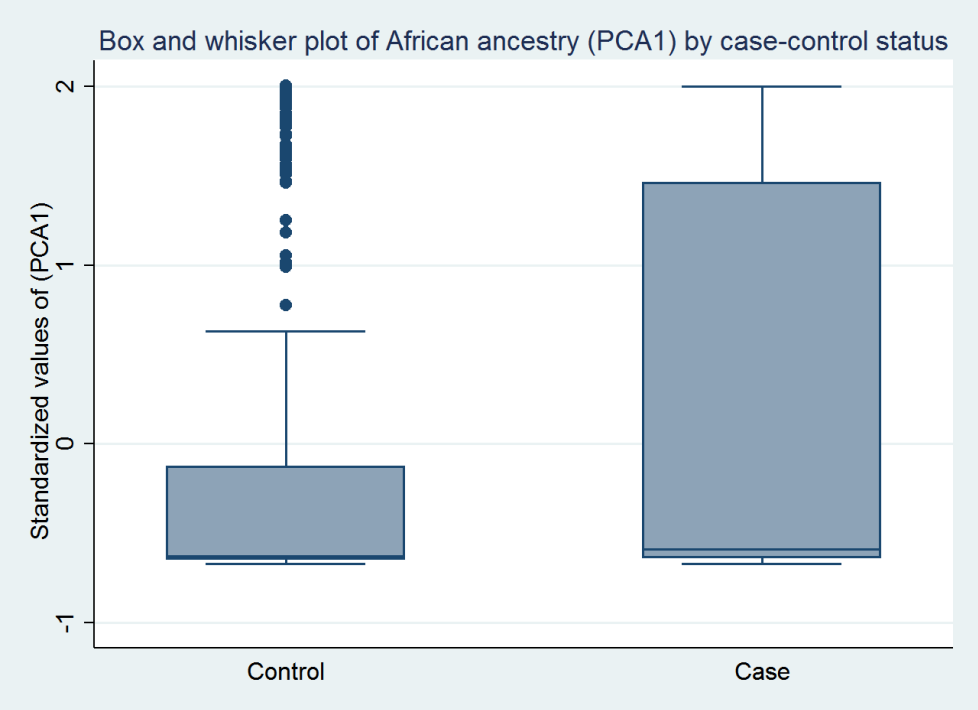
The overall median genetic distance was -0.62 (IQR: -0.64-0.48), and the Black group had a higher median genetic distance (1.78, IQR: 1.45-1.95) than the White British (-0.64, IQR:-0.65 - -0.63) and the 'other' (-0.59, IQR: -0.63 - -0.19) ethnic groups (Kruskal-Wallis χ^2: 296.3, 2df; p=0.01) (Figure 7.6 below). There is little variation in genetic distance within the White British group.

Figure 7.6: Box and whisker plot of the distribution of African ancestry by broad ethnic group



Based on the differential ethnic make-up of the case-control group, it was anticipated that there would also be differences in genetic distance between cases and controls. Controls had a median genetic distance of -0.62 (IQR: -0.63 – -0.13) and cases of -0.58 (IQR: -0.64-1.46) (Mann-Whitney U-test: -3.5, p<0.01) (Table 7.2, Figure 7.7 below)

Figure 7.7: Box and whisker plot of the distribution of African ancestry by case-status



7.4.2 Genetic distance and remaining exposures, and covariates

As can be seen in Table 7.5 below, those with a higher genetic distance tended to be slightly younger (β: 0.99, 95%CI: 0.98-0.99), experienced higher levels of childhood trauma (β: 1.02, 95%CI: 1.01-1.03) and were less likely to have ever smoked cannabis (β: 0.82, 95%CI: 0.68-1.00). Increased genetic distance was not associated with paternal socioeconomic status, but was associated with lower education (Table 7.5). Those with higher genetic distance were less likely to ever have been in a relationship (β: 0.71, 95%CI: 0.56-0.90) or lived with someone other than their parents (β: 0.58, 95%CI: 0.43-0.79), and experienced higher cultural distance (β: 1.52, 95%CI: 1.22-1.93) and levels of discrimination (β: 1.19, 95%CI: 1.12-1.26).

Table 7.5: Associations between genetic distance and remaining exposures and covariates

Variable	β (95% CI)	p-value
Age	**0.99 (0.98-0.99)**	**<0.01**
Sex		
Male	0.93(0.77-1.12)	0.43
Female	Reference	
Paternal age	1.01 (1.00-1.02)	0.18
Cannabis use		
Yes	**0.82 (0.68-1.00)**	**0.05**
No	Reference	
Childhood trauma	1.02 (1.01-1.02)	<0.01
Paternal socioeconomic status		
Professional	Reference	
Intermediate	1.11 (0.86-1.42)	0.42
Lower	1.07 (0.82-1.39)	0.62
Routine	1.22 (0.90-1.64)	0.20
Never worked	0.68 (0.14-2.31)	0.43
Not classified	1.34 (0.84-2.12)	0.22
Level of education		
Postgraduate	Reference	
Undergraduate	**1.37 (1.02-1.82)**	**0.04**
Vocational	**1.42 (1.03-1.93)**	**0.03**
Tertiary	**1.62 (1.18-2.25)**	**<0.01**
School qualifications	**1.44 (1.01-2.04)**	**0.04**
No qualifications	1.23 (0.84-1.81)	0.30
Relationship status		
Yes	**0.71 (0.56-0.90)**	**<0.01**
No	Reference	
Living arrangements		
Yes	**0.58 (0.43-0.79)**	**<0.01**
No	Reference	
Cultural distance		
Yes	**1.52 (1.22-1.92)**	**<0.01**
No	Reference	
Discrimination (0-12)	**1.19 (1.12-1.26)**	**<0.01**

Results in **bold** are significant (p<0.05).

7.4.3 Psychotic disorders and genetic distance

As hypothesised, a higher genetic distance was associated with increased odds of developing a psychotic disorder (OR for a one standard deviation increase: 1.32; 95%CI: 1.09 – 1.60), and genetic distance explained 1.4% of variance in psychosis risk (Table 7.6 below).

As soon as covariates (independent predictors of age, sex, their interaction and cannabis use, and confounders of childhood trauma and paternal socioeconomic status as well as paternal age) were introduced into the model (Model A), the association between psychotic disorders and genetic distance was no longer significant (OR: 1.06, 95%CI: 0.84-1.34). The further addition of indicators of social distance (level of education, relationship status, living arrangements: Model B) did little to change the point estimate (OR: 1.10, 95%CI: 0,85-1.41). It dropped only slightly with the inclusion of cultural distance (Model C, OR: 1.04, 95%CI: 0.80-1.35) and the further inclusion of self-reported levels of discrimination (Model D, OR: 0.96, 95%CI: 0.74-1.26). In this final model (Model D, Table 7.6), lower levels of education (OR for no qualifications: 26.31, 95%CI: 7.16-96.88), higher cultural distance (OR:2.16, 9%%CI: 1.16-4.04) and increased levels of discrimination (OR: 1.24, 95%CI: 1.03-1.51) were associated with increased odds of developing a psychotic disorder, and ever having been in a relationship was associated with lower odds (OR: 0.30, 95%CI: 0.15-0.60). This model explained 33.5% of variance in psychosis risk (Table 7.6), which was higher than other models (A-C).

Figure 7.8: Associations between genetic distance and FEP risk, by statistical model

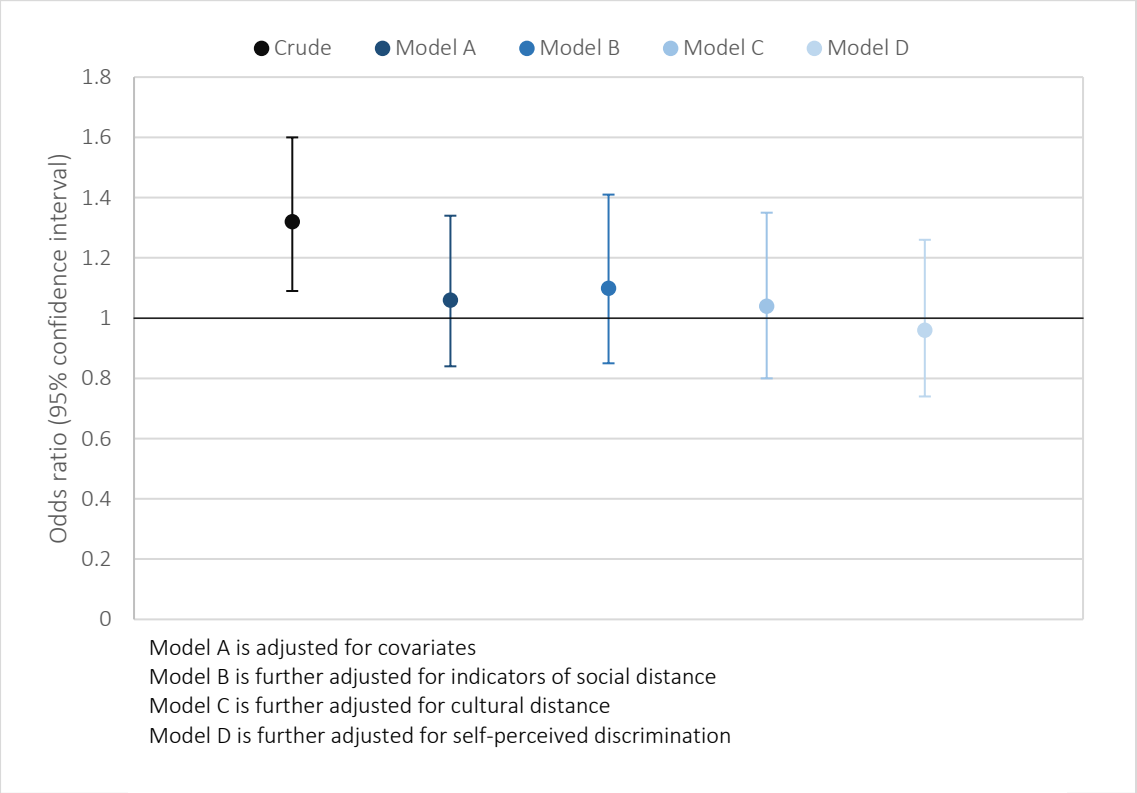


Table 7.6: Associations between FEP risk and exposures of interest, by statistical model

	Crude		Model A[1]		Model B[2]		Model C[3]		Model D[4]	
Variable	OR (95%CI)	P-value	OR (95%CI)	P-value	OR (95%CI)	P-value	OR (95%CI)	P-value	OR (95%CI)	P-value
Genetic distance	**1.32 (1.09-1.60)**	**<0.01**	1.06 (0.84-1.34)	0.61	1.10 (0.85-1.41)	0.48	1.04 (0.80-1.35)	0.77	0.96 (0.74-1.26)	0.78
Level of education										
Postgraduate	Reference		n/a		Reference		Reference		Reference	
Undergraduate	**2.24 (1.05-4.77)**	**0.04**			1.81 (0.78-4.20)	0.17	1.91 (0.82-4.45)	0.14	1.97 (0.83-4.65)	0.12
Vocational	**5.80 (2.68-12.55)**	**<0.01**			**5.19 (2.05-13.14)**	**<0.01**	**5.52 (2.15-14.15)**	**<0.01**	**5.53 (2.14-14.29)**	**<0.01**
Tertiary	**3.11 (1.40-6.94)**	**<0.01**			1.68 (0.66-4.22)	0.27	1.66 (0.66-4.19)	0.28	1.63 (0.64-4.15)	0.31
School qualifications	**8.25 (3.55-19.18)**	**<0.01**			**7.58 (2.74 (21.93)**	**<0.01**	**7.83 (2.73-22.40)**	**<0.01**	**7.50 (2.58-21.77)**	**<0.01**
No qualifications	**26.67 (9.44-75.35)**	**<0.01**			**22.57 (6.31-80.75)**	**<0.01**	**24.91 (6.87-90.28)**	**<0.01**	**26.31 (7.16-96.68)**	**<0.01**
Relationship status										
Yes	**0.25 (0.15-0.42)**	**<0.01**	n/a		**0.33 (0.16-0.64)**	**<0.01**	**0.31 (0.16-0.62)**	**<0.01**	**0.30 (0.15-0.60)**	**<0.01**
No	Reference				Reference		Reference		Reference	
Living arrangements										
Yes	**0.51 (0.27-0.95**)	**0.04**	n/a		**2.40 (1.02-5.64)**	**0.05**	2.12 (0.88-5.06)	0.09	2.13 (0.89-5.08)	0.09
No	Reference				Reference		Reference		Reference	
Cultural distance										
Yes	**1.67 (1.06-2.66)**	**0.03**	n/a		n/a		**2.04 (1.10-3.80)**	**0.02**	**2.16 (1.16-4.04)**	**0.02**
No	Reference						Reference		Reference	
Discrimination (0-12)	**1.40 (1.23-1.61)**	**<0.01**	n/a		n/a		n/a		**1.24 (1.03-1.51)**	**0.03**
Pseudo R^2		1.4%	22.9%		31.6%		32.4%		33.5%	

[1]Model A is adjusted for covariates (age, sex, their interaction, paternal age, childhood trauma, lifetime cannabis use and paternal socioeconomic status)
[2]Model B is further adjusted for indicators of social distance (level of education, relationship status, living arrangements)
[3]Model C is further adjusted for cultural distance
[4]Model D is further adjusted for self-perceived discrimination.
Odds ratios in **bold** are significant (p<0.05)

7.4.4 Genetic distance and Black ethnicity

In order to be able to assess if ethnicity or genetic distance was a more important predictor (hypothesis 3) I conducted logistic regression on the complete-case dataset. When adding genetic distance to a model already including ethnicity (and ethnicity only), neither genetic distance nor Black ethnicity was significantly associated with case-control status, although psychosis risk remained elevated in the 'other' ethnic group, and the odds ratio for the Black group increased (Table 7.7). This multivariable model did not fit the data better than a univariable model only including ethnicity (LRT χ^2: 0.04, p=0.83), but better than a univariable model only including genetic distance (LRT χ^2: 8.9, p=0.01).

Table 7.7: Associations between FEP risk and genetic distance and self-ascribed ethnicity.

	Univariable OR (95%CI)	Univariable p-value	Multivariable OR[1] (95%CI)	Multivariable p-value
Genetic distance	**1.32 (1.01-1.60)**	**<0.01**	0.95 (0.59-1.53)	0.83
Ethnicity				
White British	Reference		Reference	
Black	**2.43 (1.53-3.85)**	**<0.01**	2.70 (0.87-8.34)	0.08
Other	**2.04 (1.26-3.31)**	**<0.01**	**2.07 (1.26-3.40)**	**<0.01**

[1]Mutually adjusted for each other
Odds ratios in **bold** are significant (p<0.05).

7.4.5 Psychotic disorders and genetic distance within ethnic groups

Using univariable logistic regression, no association was found between genetic distance and psychosis risk in the White British group (OR: 0.43, 95%CI: 0.04-5.07), the Black group (OR: 0.90, 95%CI: 0.48-1.68) or in the 'Other' ethnic group (OR: 1.38, 95%CI: 0.54-3.50) (hypothesis 4).

7.4.6 Sensitivity analyses

By means of sensitivity analyses, I repeated models A-D on the non-imputed dataset. The results from the main sensitivity analyses are summarised in Table 7.8 below, and did not appear to differ significantly from the multivariable model building across the multiply imputed datasets, although there was substantial loss of observations (only 77.2%, n=342 of participants had data on all variables).

Table 7.8: Sensitivity analyses of the association between FEP risk and genetic distance, by statistical model.

Model	N (%)	OR (95%CI)	p-value
Crude	443 (100)	**1.32 (1.10-1.60)**	**<0.01**
Model A[1]	371 (83.7)	0.96 (0.73-1.26)	0.77
Model B[2]	357 (80.5)	0.93 (0.68-1.25)	0.61
Model C[3]	357 (80.5)	0.88 (0.64-1.21)	0.44
Model D[4]	342 (77.2)	0.80 (0.58-1.12)	0.20

[1]: Model A is adjusted for covariates (age, sex, their interaction, paternal age, childhood trauma, cannabis use, paternal socioeconomic status)
[2]: Model B is further adjusted for indicators of social distance
[3]: Model C is further adjusted for cultural distance
[4]: Model D is further adjusted for discrimination
Odds ratios in **bold** are significant (p<0.05).

7.5 Discussion

7.5.1 Summary of main results

This study demonstrated a crude association between genetic distance and psychosis (hypothesis 1, Table 7.9), which was accounted for by the inclusion of the covariates of age, sex, their interaction, paternal age, childhood trauma, cannabis use and paternal socio-economic status (hypothesis 2, Table 7.9). Indicators of social distance, cultural distance and psychosocial disempowerment were still associated with psychosis risk after adjusting for genetic distance. The present study suggested that self-ascribed ethnicity was a better predictor of psychosis risk than genetic distance (hypothesis 3). In this sample, no association was established between genetic distance and psychosis risk within individual ethnic groups (hypothesis 4, Table 7.9).

Table 7.9: Reappraisal of hypotheses

	Hypothesis	Outcome
1.	There is a crude association between genetic distance and psychosis	Confirmed
2.	This association is accounted for by inclusion of covariates, indicators of social distance, cultural distance and discrimination	Confirmed
3.	Ethnicity is a better predictor of psychosis risk than genetic distance	Confirmed
4.	There is no association between genetic distance and psychosis risk within individual ethnic groups	Confirmed

7.5.2 Strengths and limitations

Results from this exploratory study should be interpreted in light of its strengths and limitations. In terms of chance, I did not carry out a formal power calculation, but the present study appears to be underpowered on some tests. None of the models tested in Chapter 6 could explain excess odds in the Black ethnic group, but increased genetic distance was not significantly associated with psychosis odds following the introduction of covariates (Model A), suggesting that power is a major limitation, even after inclusion of the largest possible sample following multiple imputation. A future research direction is therefore to run these models across all five European EU-GEI samples, and subsequently meta-analyse results. This would boost sample size, and statistical power. It is unlikely this would facilitate the examination of Asian genetic heritage, considering the limited sample size of the Asian group in the EU-GEI study more broadly (n=60) and the high heterogeneity within this group.

A further limitation increasing the probability that chance was an explanation of results is the homogeneity of genetic distance in the White British group. Perhaps in a more heterogeneous White population, genetic distance would be associated with psychosis risk. The finding that here was no association between genetic distance and psychosis risk in the 'other' ethnic category (which was more heterogeneous in genetic distance), or in the Black group, tentatively suggest there might not be. Both groups were however limited in size (n=100 and n=115 respectively).

In addition to the biases resulting from the case-control study more widely (see Section 6.5.2), there is a possibility of a degree of bias resulting from sample selection. The control subsample recruited from the two English catchment areas was representative of the general population in terms of sex, but not in terms of age and binary (minority/majority) ethnic composition. There was also a degree of selection bias resulting from missing data. Cases, those from an ethnic minority and younger individuals were less likely to have been genotyped, and cases were more likely to have missing data on paternal age, childhood trauma, and discrimination. Genetic distance as operationalised in this dataset (Nigerian ancestry), appeared to be an acceptable, though not perfect, fit for this sample. Of the 60 individuals with Black ethnicity born outside the UK, thirteen (21.7%) were born in Nigeria, and a further 20 (33.3%) in wider West Africa. A further six (10%) were born in the Caribbean, and consequently were also likely to have West-African genetic heritage. Fourteen (23.3%) individuals were born in Eastern Africa, three (5%) in Southern Africa, three (5%) in other Western countries and one (1.7%) in South America.

Risk of bias due to missing data depends on the underlying reasons for this missing data (see Section 6.5.2). For this study, I have reported missing data by diagnostic outcome (Table 7.2), and any differences in genetic distance between those with complete and missing data (Table 7.3). Whilst there were some differences in missingness by case-control status, and differences in genetic distance by missingness, the similarity between the imputed analyses and sensitivity analyses suggest it is plausible that data is missing at random (MAR), meaning that estimates from imputed data would be unbiased (Sterne et al., 2009).

Whilst a large number of covariates were included, residual confounding could not be excluded (see Section 6.5.2). However, in the present study the main threats to validity and generalisability were considered to relate to its low power, and biases as described above.

7.5.3 Interpretation of main results in light of the existing literature

This is the first study exploring the role of genetic distance in explaining psychosis risk, and briefly comparing it with self-ascribed ethnicity. As such, interpretation should be done cautiously, and no previous literature was available to aid in contextualising findings.

In the present study, socio-environmental risk factors and covariates explained the excess psychosis risk in those with an increased genetic distance. In the final model (D) a lower level of education (Morgan et al., 2008), never having been in a relationship (Gayer-Anderson & Morgan, 2013), increased cultural distance and increased self-reported levels of discrimination (Karlsen et al., 2005; Veling et al., 2007) were associated with increased psychosis risk, as previously found in Chapter 6, and genetic distance was not. Whilst this, and the lack of association between genetic distance and psychosis risk within ethnic groups, is consistent with the role of the social environment in explaining excess odds of disorder in ethnic minorities found in Chapter 6, the low statistical power of the present study prohibits the drawing of any firm conclusions.

It appeared unlikely that genetic distance could be a useful proxy for ethnicity where no ethnic data is available. The present analyses suggested that a multivariable model including both genetic distance and ethnicity does not fit the data better than a univariable model only containing ethnicity. These results might be affected by the low statistical power of this study, and further research is needed before any statements can be made about the usefulness of genetic distance as proxy for self-ascribed ethnicity. More practically, it currently seems implausible that genotyping an individual to obtain genetic distance would be undertaken before noting a measure of self-ascribed ethnicity. To some extent this is a moot point: in the present study, both were statistically explained by the social environment.

The finding that ever having lived with someone other than one's parents became a risk factor rather than a protective factor following multivariable adjustment was unexplained and unexpected. Although this didn't persist into the final model, it is curious: it can't be explained by the housing situation in the two catchment areas, which both have such high housing prices that many people find themselves renting with others well into adulthood. This could mean that 'living with someone other than your parents' in itself becomes a marker of disadvantage (being forced out of the housing market into shared housing). This is in line with findings from Chapter 4, where I demonstrated that a lower percentage of owner-occupancy was associated with increased incidence, and is also supported by some literature. For instance, Morgan and colleagues found that Black Caribbean people were more likely to live in rented housing than their White British counterparts (Morgan et al., 2008), and Veling and colleagues found an association between increased residential mobility and increased incidence but not between percentage of single-person households and incidence (Veling et al., 2014). Furthermore, Bhavsar and colleagues did not find an association between poor housing and a deprived living environment and increased incidence (Bhavsar et al., 2014). However, this does not explain the crude protective effect of ever having lived with someone other than one's parents.

The finding that increased levels of discrimination were associated with increased odds of psychosis in this sample was unexpected. A positive association between discrimination and psychosis was hypothesised for the total sample, on the basis of a theoretical model, and some epidemiological literature (Karlsen et al., 2005; Veling et al., 2007; Veling, Hoek, et al., 2008). However, as it was not found in the overall EU-GEI sample, no effect was expected in the smaller subsample from the two English settings. This might be a chance finding, however there appeared to be higher levels of discrimination reported in England, compared with the overall EU-GEI sample. Only 45.8% reporting never to have been victims of discrimination, compared to 59.3% in the overall sample (χ^2: 46.88, $p<0.01$) so this might reflect a true effect (perhaps reflecting the higher percentage of individuals from an ethnic minority background in this sample).

More broadly, these results fit with a recent trend in using genetic data epidemiologically. As mentioned in the introduction to this Chapter, the PRS of schizophrenia has some limitations, and resent research from the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort, suggests that whilst the PRS is associated with negative symptoms at age 16.5 or anxiety at age 15.5, it is not associated with psychotic experiences at age 12 or 18 or depression at age 15.5 (Jones et al., 2016). This suggests that genetic risk for schizophrenia in adolescence does not manifest itself as psychotic symptoms, contrary to what would be expected on the basis of a 'psychosis continuum' model (Van Os et al., 2009), and perhaps that use of the PRS is more limited than initially hoped. The problems with using the PRS epidemiologically in ethnic minority groups are also documented in Section 7.1, and the results from this Chapter might give an indication as to why using PCAs does not substantially improve the predictive ability of the PRS in Black ethnic minorities (even if this would be a perfect predictor in the White majority population): genetic determinants of ancestry might simply not be as important a predictor as the sociocultural context of ethnic minorities.

7.5.4 Conclusion

In this sub-study, the sociocultural distance model explained excess odds of psychosis in those at increased genetic distance from the White British majority, tentatively underscoring the importance of the sociocultural context experienced by consequence of having an ethnic minority background (specifically, being Black in England) vis-à-vis a genetic determination of this minority status. However, this was an exploratory analysis in a small (n=443) sample only looking at African ancestry in England, so further research is needed and no firm conclusions could be drawn.

Chapter 8 - General Discussion

In this Chapter I summarise the main findings from this PhD in relation to its central theses of substantial heterogeneity in psychosis risk by place and person and the role of the social environment in explaining this (Section 8.1). I discuss any threats to their validity in terms of chance, bias, confounding, and reverse causality (Section 8.2). I also examine the main findings in context of the existing literature (Section 8.3), and on this basis make an estimate of how likely findings from this thesis are to be causal (Section 8.4). I discuss the implications in terms of aetiology of psychosis, and service provision, policy and public health (Section 8.5). Finally, I outline areas of further work (Section 8.6) and provide a general conclusion (Section 8.7).

8.1 Summary of main findings

In this thesis, I researched the substantial variation in the incidence of psychotic disorders existing across both person and place. In Chapter 2, I established support for this using a systematic review and meta-analysis of existing literature from the six countries included in the EU-GEI study. I demonstrated high heterogeneity in incidence rates across a range of diagnostic outcomes including all psychotic disorders and non-affective and affective disorders. A trend towards higher incidence in men compared with women was observed for both all psychotic disorders and schizophrenia, and rates of all FEP were higher in ethnic minorities and in younger age groups. I demonstrated that increased urbanicity and latitude were associated with increased incidence of all psychotic disorders, although variation remained and I could not explain heterogeneity across other diagnostic outcomes.

It was difficult to ascertain if this heterogeneity reflected true variance in treated incidence, or if this was due to methodological differences between studies, which appeared to play a role. To reduce methodological heterogeneity, I analysed incidence data from the EU-GEI study using a consistent methodology across 17 settings in six countries (Chapter 4). Here I found substantial variation by person and place: after standardisation for age, sex and ethnic minority status, incidence of psychotic disorders varied eight-fold between settings. Similar variation was found independently for non-affective and affective disorders. As in my systematic review, the EU-GEI study found rates were highest in young men, and were higher in ethnic minorities compared with the White majority. Whilst in the meta-analysis the social environment (as measured by inequality, self-perceived trust and freedom) was not associated with the incidence of psychotic disorders, in the EU-GEI study lower owner-occupancy was associated with increased incidence of all psychotic disorders and non-affective disorders. Lower unemployment was associated with increased incidence of affective disorders.

In the remainder of the thesis, I sought to explain one component of the observed heterogeneity in psychosis risk: the elevated rates among some ethnic minority groups. One of the most consistent

epidemiological findings, both in the international literature and in the first part of my thesis, was the increased risk of psychosis in migrants and their descendants. Drawing on epidemiological observations and using literature from social science and philosophy (Chapter 5), I suggested the sociocultural context of minorities could be a putative risk factor for their excess psychosis risk. I argued that culturally distant minorities were at particularly increased risk of social exclusion or falling outside the group of 'fellow citizens', leading to increased social distance. Such an outsider experience can lead to increased psychosocial disempowerment, and this is how I hypothesised this sociocultural distance model would increase psychosis risk.

I tested this explanatory framework on 1,100 cases and 1,473 controls from the EU-GEI study (Chapter 6). Most ethnic minority groups had excess odds of developing any psychotic disorder, but after the inclusion of covariates, indicators of social distance, cultural distance and self-perceived discrimination (as a proxy for psychosocial disempowerment) only the Black group remained at excess risk. Even in this group, excess risk was substantially attenuated (OR reduced from 2.95 to 1.53), suggesting an important role for the social environment. Level of education (one of the indicators of social distance) and cultural distance were particularly important in explaining excess risk, but self-perceived discrimination was no longer associated with psychosis risk in a multivariable model. Similar results were found independently for non-affective and affective disorders. When exploring religious minority status (Jewish, Muslim or 'other' religions), it appeared that following any religion (including Christianity) was a strong predictor of psychosis risk, and this was not fully explained by including covariates (including ethnicity), indicators of social distance, cultural distance, and psychosocial disempowerment in the model. This model attenuated excess risk in the Muslim and 'other' religious groups, but increased risk in the Christian group.

I proposed that these exposures could explain any crude association between genetic distance (African ancestry) and excess psychosis risk. I tested this using the subsample of the EU-GEI study recruited in England (Chapter 7). I established a crude association between genetic distance and risk of developing a psychotic disorder, but explained this by inclusion of covariates, indicators of social distance and cultural distance. Self-ascribed Black ethnicity appeared to be a better explanation for variance than genetic distance, and no association between genetic distance and psychosis risk was found within ethnic groups. These results tentatively support the important role of the sociocultural context in increasing psychosis risk in ethnic minorities.

8.2 Potential threats to validity

The findings resulting from this thesis should be interpreted in light of the methodological strengths and limitations inherent in this work. Strengths and limitations specifically pertaining to each empirical Chapter were discussed in the discussion of these Chapters, and this Section reviews the major threats to validity of the overall findings, in terms of chance, bias, confounding and reverse causality.

8.2.1 Chance

I employed a number of strategies to minimise chance being an explanation of the main findings of this thesis. I had clearly formulated *a priori* specified hypotheses, which meant my modelling approach was focussed on the relevant questions to avoid data-mining or dredging. This reduced the probability of a Type-I error (the incorrect rejection of the null-hypothesis or a false positive finding) by limiting the number of exposures included in the model, and subsequent statistical tests. This was clearest in Chapters 6 and 7, which built on the sociological and philosophical work outlined in Chapter 5, but was also apparent in Chapters 2 and 4, where *a priori* hypotheses were included. The use of appropriate random-effects (Chapter 2) or multi-level (Chapters 4 and 6) modelling meant that standard errors were larger and confidence intervals wider. This not only reduced the possibility of a Type I error, but also was a more realistic depiction of the hierarchical nature of the data.

The EU-GEI study yielded a large sample size, allowing precise estimates of effect across both the incidence study (Chapter 4) and the case-control study (Chapter 6). I also used multiple imputation to avoid loss of precision due to missing data (Chapters 6 and 7). The interpretation of findings was careful. The similarity of findings yielding from the case-control data across diagnostic outcomes and ethnic groups (Chapters 6 and 7) increased confidence they were not due to chance; any further empirical support for these findings yielding from the wider literature will be assessed in the next Section. Interpretation of findings did not rely on p-values alone, but I examined (changes in) effect sizes and confidence intervals. Owing to the large sample size, p-values are always likely to be small and don't convey any information on the magnitude of the effect. Assessing effect sizes alongside p-values protects against false confidence in the strength of findings due to small p-values (or large effect sizes) alone (Sullivan & Feinn, 2012).

Nonetheless, chance may have played a role. Despite the formulation of *a priori* hypotheses and the similarity of results, I did carry out a substantial number of statistical tests increasing the probability of a Type I error. When accepting a p-value of 0.05 as indicating a significant result, one in 20 tests will be due to chance (Johnson, 2013). Recent developments to address the reproducibility crisis in (medical) science have included proposal to move the significance threshold for discovery to $p<0.005$ (Singh Chawla, 2017), and although this is no panacea, it would substantially reduce the possibility of false positives. Most of the findings in this thesis have a p-value of <0.01 (for instance the higher incidence in catchment areas with low owner-occupancy in Chapter 4, and the association between cultural distance and all psychotic disorders in Chapter 6), so Type-I errors appear to be rare.

Importantly, the genetic distance study in Chapter 7 does not benefit from such a large sample size (n=443), and as such is more prone to chance variation. In particular, the lack of statistical power in this Chapter may have led to a Type-II error (the incorrect failure to reject the null hypothesis, or false negative), as the excess risk of psychosis in those with increased genetic distance was readily explained in

this study, despite the fact it could not be explained in Black group in the larger study in Chapter 6. Findings from this Chapter will need to be replicated in a larger sample to reduce the possible explanation of results by chance, before any conclusions can be drawn.

8.2.2 Bias

A number of biases should be considered when interpreting this work. The first two relate to the data used from the case-control arm of the EU-GEI study. First, there is likely to be a degree of information bias resulting from imperfect measurements. The exposures defined in Chapter 5 (cultural distance, social distance and psychosocial disempowerment) based on sociological and philosophical literature were broad constructs that were measured using much narrower proxies. This thesis also introduced several novel exposures to the field (cultural distance and genetic distance, as well as the concept of social distance – though not the indicators of it), and there are therefore limited precedents on which to model the measurement of these exposures. Cultural distance, for instance, has mainly been used in a business context (Shenkar, 2001; West & Graham, 2004), and self-perceived discrimination might be a poor measure of psychosocial disempowerment. Whilst this was likely to have led to a degree of information bias, it was not possible to determine the direction of this bias (away from the null hypothesis or not, depending on whether or not the bias operated differentially by ethnicity). Findings pertaining to these exposures are therefore interpreted carefully, and the extent to which they can be generalised to the broader construct is limited.

Second, as with any retrospective study, the possibility of recall bias cannot be excluded. Cognitive impairments including memory deficits are common in individuals with psychosis (Aleman, Hijman, De Haan, & Kahn, 1999), and as such differential recall bias is a serious consideration: it is possible that there is a systematic difference between cases and controls in accuracy or completeness of the recollection of, for instance, lifetime cannabis use, experiences of discrimination or lifetime relationship status. Whilst other measures (such as first language) are considered less prone to recall bias, the effect of recall bias overall is difficult to determine.

In relation to the incidence work, there might be a degree of ecological bias or fallacy (Sections 2.5 and 4.5). In the meta-regression (Chapter 2) urbanicity was a relatively crude measure and income inequality, self-perceived trust and freedom were only estimated at country level. In the EU-GEI study (Chapter 4) unemployment, single-person households and owner-occupancy were estimated at provincial (NUTS-2) level, and I don't know if the individuals included in the EU-GEI study live in neighbourhoods with levels representative of their NUTS-2 region. It is not possible to assess the influence of this ecological fallacy.

However, the EU study overall sought to minimise a number of biases through its comparable methodology across catchment areas, and I considered a number of biases when designing the studies included in this thesis. In the overall study, ascertainment bias was minimised by using a comparable

recruitment strategy across catchment areas (see Chapter 3). Participants were recruited and assessed via a standardised protocol, and questionnaires administered were identical across catchment areas, reducing the possibility of bias through measurement error. Furthermore, I chose lifetime exposures to estimate cannabis use, relationship status and living arrangements to minimise recall bias, and any bias associated with missing data (lifetime data was more complete than data from a specific time point). Onset of psychosis is also difficult to determine (Singh et al., 2005), so using lifetime measures also avoids any spurious accuracy associated asking for exposure five years prior to onset, for example.

Where possible, I have sought to minimise information bias by using well-validated measures to estimate my exposures (see Section 4.3 and 6.3). I also used multiple imputation (Chapters 6 and 7) to avoid introduction of bias through missing data. The validity of the approach used is discussed in the discussion Sections of Chapter 6 and to a lesser extent Chapter 7, but broadly indicate that the assumption of data being Missing At Random is plausible, and if this holds results yielding from analyses based on imputed date would be unbiased.

8.2.3 Confounding

Despite the fact that I adjusted for a large number of confounders (particularly in Chapters 6 and 7), confounding cannot be ruled out as having influenced the results presented in this thesis. There might have been unobserved confounders: in Chapters 6 and 7, analyses were based on literature from the social sciences and it is possible that despite best efforts a theoretical confounder has been overlooked. Such 'known unknowns', if they exist, could theoretically be discovered by further careful research, but 'unknown unknowns' by their definition are elusive (Pearl, 1984). A known unknown for Chapters 6 and 7, for instance, is family history of psychosis. By controlling for as many known confounders, the likelihood of there being 'unknown unknowns' is reduced (Cartwright, 1989), but one can never know how close one is to eliminating them (Worrall, 2007). Furthermore, in the incidence study we were unable to test for a number of known individual confounders as these were not routinely available at population-level (such as cannabis use, childhood trauma and socio-economic status, see Section 4.5). In the case-control study I have not been able to include urban birth and upbringing, and family risk of psychosis as this data was not available. These confounders could potentially have further attenuated odds ratios in ethnic minority groups, and provide an alternative explanation for their excess risk, provided they do confound the relationship between ethnic minority status and psychotic disorders.

Residual confounding resulting from imperfect measurements likely plays a role, but is unlikely to fully explain findings. As can be read in Section 4.5, there is likely to be a degree of residual confounding in the incidence study as a consequence of using a binary ethnic minority status measure. However, this is unlikely to change the main findings of substantial variation across person and place, and a role for the sociocultural context in explaining this heterogeneity (Susser & Martínez-Alés, 2017). In Chapters 6 and 7,

residual confounding is an important consideration. The exposures are broad constructs (social distance, cultural distance and psychosocial disempowerment) and their measurement is imperfect. This does not only mean there is a risk of information bias (see above), but also means that residual confounding is a potential explanation of results. This is a possible explanation for the strong association found between cultural distance and psychosis risk in ethnic minorities in Chapters 6 and 7. For the former, whilst language distance and fluency in the majority language appear to be a satisfactory proxy for cultural distance through their central role in identity formation (see Section 5.4), culture encompasses more than just language, and language might not be an equally important element of it for everyone.

8.2.4 Reverse causality

When using cross-sectional data, reverse causality is always an important consideration. In Chapter 4, the possibility of reverse causality is acknowledged regarding the association between lower owner-occupancy and higher incidence of all psychotic disorders and non-affective disorders: individuals at increased risk of a psychotic disorder may have moved to more unstable areas. In Chapters 6 and 7, the potential for reverse causality was most evident for indicators of social distance and self-perceived discrimination, and perhaps less so for cultural distance and genetic distance. Genetic distance or ancestry is fixed over the life-course, and reverse causality is at best a partial explanation for findings on cultural distance. It is possible that gaining fluency in the majority language is hampered by a prodromal psychotic disorder. However, this would only apply to those who have not obtained fluency in the majority language prior to developing psychosis. Furthermore, reverse causality cannot explain the proportion of cultural distance that is accounted for by having a first language that isn't the majority language. Language acquisition typically starts in infancy – well before any prodromal symptoms might appear.

There is a stronger potential for reverse causality acknowledged for the indicators of social distance. The relationship between childhood IQ and later schizophrenia risk is well-established and robustly replicated (Khandaker et al., 2011). This reverse causality (where lower IQ is a marker of neurodevelopmental origins of disease, rather than a risk factor) might extend to educational attainment (which is associated with, but not solely determined by IQ, see Section 6.5), and interpretation of the strong association between educational attainment and psychosis risk in this study should not extend to the former causing the latter. For the remaining two indicators of social distance (relationship status and living arrangements), I examined life-time exposures and this reduced the possibility of reverse causality compared to examining current exposures, but did not eliminate it. The univariable association between psychosis risk and self-perceived discrimination could be due, at least in part, to increased paranoid ideation, and as such this finding needs to be interpreted cautiously. This association did not persist in multivariable models.

Reverse causality is also a possible alternative explanation for findings on religion. As discussed in Chapter 6, there is some evidence for a bias towards external attribution styles in psychosis (Garety et al., 2001;

Van Os et al., 2009), and this might also lend itself well to religious explanations of daily-life phenomena, particularly when they become increasingly otherworldly. Furthermore, being religious is not a fixed characteristic or identity and I had no data on duration of religious affiliation, so it could be that those who are more prone to psychosis are more likely to stay or become religious (perhaps as a coping mechanism), or that the latter marks a shift in their beliefs related to religion.

8.3 Interpretation and contextualisation of findings

This thesis is written within a wider tradition of social psychiatry (see Section 1.4), arguing that the psychotic disorders should be understood as disorders of social functioning (Abed & Abbas, 2011; Clark, 2016; Fletcher & Frith, 2009; Gold & Gold, 2014; Van Os et al., 2010). This allows for using a sociological framework which, as was discussed in Chapter 5, facilitates the identification of structural social drivers of mental illness more broadly and psychotic disorders specifically (Horwitz, 2017; Scambler, 2010; Thoits, 2017; Wilson, 2010). It is in the context of this literature from social psychiatry, sociology and neuroscience that I have interpreted the main findings resulting from the empirical work presented in this thesis: the variance in incidence or risk of psychotic disorders across place and person, and the explanatory role of the social environment. Epidemiological support for empirical findings is discussed in the discussion Sections of Chapters 2, 4, 6, and 7.

8.3.1 Social psychiatry

The variance in incidence rates between studies (Chapter 2) and catchment areas (Chapter 4) can, to some extent, be explained in the contact of a dysfunctional social threat response (Abed & Abbas, 2011; Gold & Gold, 2014). Gold and Gold (2014) propose that any association between urbanicity or population density and increased incidence of psychotic disorders is due to the social threat of malicious intentions being more frequent. Although there is strong epidemiological evidence for such an association (Vassos et al., 2012), this thesis found limited support for an association between population density and risk of psychosis. This might indicate that Gold and Gold's position benefits from refinement: these excessive social threat levels might be specific to certain characteristics of the social (and perhaps built) environment which are more common in densely populated urban areas, and the association is not with population density as such. A long-standing tradition of research into neighbourhood-level risk-factors concerning social fragmentation, disadvantage and lack of security (Allardyce et al., 2005; Bhavsar et al., 2014; Dohrenwend & Dohrenwend, 1969; Drukker et al., 2006; Faris & Dunham, 1939; Giggs & Cooper, 1987; Hare, 1956; Hollingshead & Redlich, 1958; Kirkbride et al., 2014; Maylaih et al., 1989; Omer et al., 2014; Veling et al., 2015) can be interpreted in this way: the social threat of malicious intentions might be more common or more salient in areas where there is low social cohesion or sense of belonging and associated increased feelings of unfamiliarity, and this unfamiliarity is associated with increased threat levels. The association between lower levels of owner-occupancy and higher incidence of (non-affective) psychotic

disorders are interpreted in this tradition: as indicating a high degree of instability, which would make the social threat of malicious intentions more frequent.

This same paradigm was applied when interpreting higher rates of psychotic disorders in ethnic minorities (variance by person). Sociocultural distance can make social threat of malicious intentions more salient, by making it more difficult to assess intentions. The default option when encountering an ambiguous situation is a threat response (Gold & Gold, 2014). This is alleviated by a sense of familiarity based on previous experiences, and being socioculturally distant might make achieving this familiarity more difficult. Social isolation or lack of social support can make it less likely to develop a wide notion of what are normal social interactions, and this in turn can make it more likely that any interaction is interpreted as threatening, possibly in part due to the unfamiliar homogeneity effect whereby everything that is unfamiliar becomes homogenous and threatening (Malinowska, 2016). These findings also fit within a socio-developmental cognitive model of psychosis (Howes & Murray, 2014), where early life experiences can bias cognitive schemas and subsequent interpretation of the world, as well as within the framework of predictive processing (Clark, 2016). We build our model of the world in cooperation with others, and if there is less contact with the wider world, this might be less successful. A low education can link to a low cognitive reserve (Khandaker et al., 2011), and this might simply make it more difficult to assess others' intentions. For survival, the default option of assuming danger is crucial, but this might lead to increased paranoid ideations in a modern world where survival is not continuously threatened.

Increased cultural distance could also contribute to increased difficulty in determining intentions. In Chapter 5, I argued that, acknowledging that everyone exhibits multiple identities, the forming of an identity might be more difficult for (ethnic) minorities. This is chiefly for two reasons: the boundaries of their identity are to a larger extent determined by the majority (who will mostly see someone as 'Black' or a 'good Black student'), and there are fewer role models or positive stereotypes available to base identities on. It is speculative, but it could be that these two factors also make it harder for more culturally distant minorities to determine intentions. Language plays an obvious role in this: English, for example, is fraught with euphemisms and expressions that convey an entirely different message than what is actually being said, and this is only open to those with a near-native level of fluency ('Interesting…'). This is supported by predictive processing, where language plays a crucial role as the scaffolding from which we construct our views about the world (Clark, 2016). However, Abed and Abbas argue for a broader conception of communication as their central thesis is that limited communication with in-group members, excessive communication with out-group members or a combination of both is crucial in the excessive threat response in ethnic minorities (Abed & Abbas, 2011).

As I mentioned in Chapters 1 and 5, I hypothesised that the association with discrimination was linked to the world being more threatening if you are an outsider. The social defeat theory (Selten et al., 2013) has

the central preposition that an individuals' risk of psychosis is increased when they are subject to the negative experience of being excluded from the majority group and experience this as defeating. Results from Chapter 6 suggest however, that the deep-rooted sociocultural drivers of inequality are important in driving excess psychosis risk. These are discriminatory, regardless of whether the are cognitively appraised as such by those affected. The crude association between discrimination and psychoses risk appears to be perfectly predicted by sociocultural distance. Outsider status still appears to be important (and threatening), but is not well captured by self-perceived discrimination.

Whilst findings on religion are much more tentative due to their deviance from prior hypotheses, lack of empirical support elsewhere and possible alternative explanations (see Section 6.5), they could be interpreted in a similar tradition. Depending on specific religious beliefs, and degree of religiosity, religious individuals might be more likely to ascribe external or religious meaning, and thereby salience, to perceived social threats. However, this is purely speculative, and does not explain why the statistical model attenuated risk in minority religions (Muslim and other), but increased risk in Christian groups, after controlling for ethnicity. Perhaps there is a cumulative or interactive effect to both being of an ethnic minority background and ascribing to a minority religion.

8.3.2 Neuroscience

The neuroscience of how epidemiological risk factors impact on the brain is still in its infancy, and largely focuses on perceived social stress (Akdeniz, Tost, & Meyer-Lindenberg, 2014). For instance, urban birth was found to be associated with increased amygdala activity, and urban upbringing was associated with changes in the perigenual anterior cingulate cortex, which is important in regulating amygdala activity, negative affect, and stress (Lederbogen et al., 2011). Current work is trying to disentangle which aspects of urban living cause these changes, but results of this are not yet available[6].

The role of social stress in the development of psychotic disorders in ethnic minorities is researched slightly more widely. One study investigating neural social stress processing associated with perceived discrimination in ethnic minority individuals also demonstrated increased connectivity in the perigenual anterior cingulate cortex, and in ethnic minorities this was mediated by chronic social stress (Akdeniz, Tost, Streit, et al., 2014). A second study demonstrated a higher amygdala reaction to out-group faces, particularly in Black participants in response to White faces, and particularly if they lived in more segregated or less ethnically dense neighbourhoods (McCutcheon et al., 2017). A further study showed that racially primed social exclusion leads to particularly high stress in minorities exhibiting a high number of attenuated psychotic symptoms (stress was measured physiologically), compared with neutral social ostracism, suggesting particular salience for race-related stress in developing psychosis (Anglin, 2017). A

[6] Professor Andreas Meyer-Lindenberg presented the design of this study at the 6th European Conference on Schizophrenia Research in Berlin (September 2017).

final study demonstrated, across two samples, increased striatal stress-induced dopamine release and synthesis capacity in minority participants compared with ethnic majority participants, regardless of their clinical status (healthy control, ultra-high-risk or suffering from psychosis)(Egerton et al., 2017).

Overall, this neuroscientific evidence suggests a role for social stress, although it is difficult to determine if this is a higher baseline (or chronic) stress level, or higher reactive or acute stress. To an extent, this fits in with the idea of 'allostatic load'. Hormones associated with a chronic stress burden protect the body in the short run, and promote adaption (allostasis), but in the long run the burden of chronic stress causes changes in the brain and body that can lead to disease (allostatic load)(McEwen, 2012). Minority groups appear to have higher stress responses than the general population, although this is yet to be established longitudinally. This, I think, increases the likelihood that the discrimination measure used was an inadequate indicator of psychosocial disempowerment or social stress, and the role of social stress continues to be important.

8.3.3 Sociology

The evidence cited above to indicate that the social threat of malicious intentions might be more frequent in areas of greater social disorganisation and more salient in minority groups, also supports a sociological structural strain argument. The broader organisation of society operates in such a way that disadvantage is concentrated in certain areas and population groups (Thoits, 2017). Findings in relation to high incidence in areas with lower owner-occupancy could be seen as support for the idea that areas with lower social capital are associated with increased mortality, morbidity, and stress levels (Putnam, 2007; 2015), as high residential turnover and low homeownership were indicators of social capital.

Findings on the increased risk of psychotic disorders in ethnic minority groups, and the subsequent role of the sociocultural distance hypothesis in explaining this add further weight to the argument that disadvantage tends to cluster in minority groups (Wilson, 2010), and that this contributes to their poorer health (Williams et al., 2017), specifically, their increased risk of psychosis (Morgan et al., 2008). In Chapter 5 I argued that this would lead to high levels of stress proliferation and psychosocial disempowerment. It seems likely that his is both to an extent already captured in the operationalisation of sociocultural distance, and difficult to disentangle in an observational study. It appears to be easier to research psychosocial disempowerment or social stress in neuroscience (see above) and animal models and studies (McEwen, 2012; Sapolsky, 2005). Sociological theories would also support the findings from Chapter 7, where the sociocultural context was tentatively found to be more important in predicting psychosis risk than a genetic determinant of ancestry. They would further support the concept of allostatic load, suggesting that daily hassles or stressors of day-to-day life are at least as important as event stressors (Thoits, 2017; Wheaton & Montazer, 2017). More generally, sociology takes the organisation of society as

its starting point, and not any subjective interpretation of it meaning that, ultimately, structural disadvantage and exclusion increase morbidity.

8.4 Are these findings causal?

Whilst there might be some alternative explanations for the variance in incidence rates observed in the EU-GEI incidence study (see, for instance Susser 2017), I do not believe these will explain all variance. I think the second part of this thesis (from Chapter 5 onwards) is perhaps more open to alternative explanations, and I therefore comment on the likelihood of causality of the evidence presented in Chapters 6 and 7 in light of the strengths and limitations and wider evidence as presented. Inferring causality from observational studies is fraught with difficulty, and I do not believe that any causal claims are justified on the basis of this thesis alone. As I mentioned in Chapter 5, I understand causes to be mechanistic explanations underpinned by probabilistic associations. Psychotic disorders are multi-causal disorders, whereby any cause identified is at best an INUS-condition: an *Insufficient* but *Necessary* part of an *Unnecessary* but *Sufficient* condition. In other words: any cause is only one ingredient of the causal cake. Not everyone from an ethnic minority background who experiences the adverse social circumstances described in this thesis will develop psychosis, and not everyone who has developed psychosis is from an ethnic minority background or has experienced these adverse circumstances (at least not to the same extent). Identifying causes from 'enabling conditions' can be difficult and whereas Woodward's criteria of specificity and stability are philosophically elegant, this thesis is insufficient in breadth and depth to satisfy these. I have therefore attempted to assess the likelihood of causality of the role of social distance, cultural distance and psychosocial disempowerment against the Bradford-Hill criteria (Table 8.1 below).

Table 8.1: Strength of evidence assessed against Bradford-Hill criteria

Criterion	Strength of evidence
Strength	Strong for relationship status, level of education and cultural distance, weaker for living arrangements and weak for discrimination.
Consistency	Strong for relationship status and level of education, weaker for living arrangements and discrimination, and absent for cultural distance.
Specificity	Absent; they are very broad risk factors, which could be aetiologically relevant to more disorders.
Temporality	Moderate for cultural distance, weaker for indicators of social distance and discrimination
Biological gradient	Strong for level of education and cultural distance, impossible to assess for discrimination, living arrangements and relationship status.
Plausibility	Informed by social psychiatry and sociology, and some evidence of a biological mechanism.
Coherence	To a degree: consistent with social psychiatry and sociology, different from neuroscience on the role of social stress/discrimination.
Experiment	No experimental evidence available
Analogy	There are no analogous risk factors we know are causal.

Overall, the strength of the evidence as assessed against the Bradford-Hill criteria was limited: cultural distance had a strong association, preceded the outcome and there was some evidence of a biological

gradient[7], but this finding has not been replicated previously, and no experimental evidence was available to corroborate findings. Level of education and relationship status have strong epidemiological precedents and associations, but it is impossible to distinguish temporality clearly due to the neurodevelopmental origins of psychotic disorders. The findings on discrimination (a crude association, but not within a multivariable model) are to an extent incoherent with the wider literature: whilst epidemiological findings are equipoised, neuroscientific and animal-based research, as well as more theoretical work based on sociology and social psychiatry, point towards a strong role for subjective stress. Whereas Chapter 7 adds tentative evidence on the importance of the sociocultural distance model versus genetics, this was a very small study, with no wider supporting evidence or precedent. The work presented in this thesis is also not very specific: the risk factors presented here could also be aetiologically relevant to very different disease outcomes, such as the association between stress and cardiovascular disease. Furthermore, I can't appeal to experimental evidence or to analogy to strengthen any potential causal claims. This thesis, a best, provides new and testable hypotheses.

The theory outlined in Chapter 5 needs refining on the basis of the presented evidence in Chapters 6 and 7. Whereas the central prepositions have not been falsified, the operationalisation of the concept of psychosocial disempowerment would benefit from refinement. It needs to be operationalised more closely in line with its actual meaning, and reflect the hassles of daily life, or chronic stress. Cultural distance and social distance do appear to play an important role, though their measurement was imperfect (see Section 6.5). Furthermore, the correlation between the concepts was not as strong as anticipated. In Figure 5.3, I proposed that cultural distance would influence both social distance and psychosocial disempowerment, and that the latter two would influence each other too. Whereas all exposures were associated with (certain) ethnic minority groups, the correlations between cultural distance, social distance and psychosocial disempowerment were typically low, meaning they measure largely different constructs.

8.5 Implications

In this Section, I will discuss the implications of the findings in this thesis as if they were true and causal. In a way, it is a thought experiment: what are the implications for the aetiology of psychotic disorders and for service provision, policy and public health if there really is heterogeneity in risk, and the sociocultural distance hypothesis can explain excess risk in minority groups.

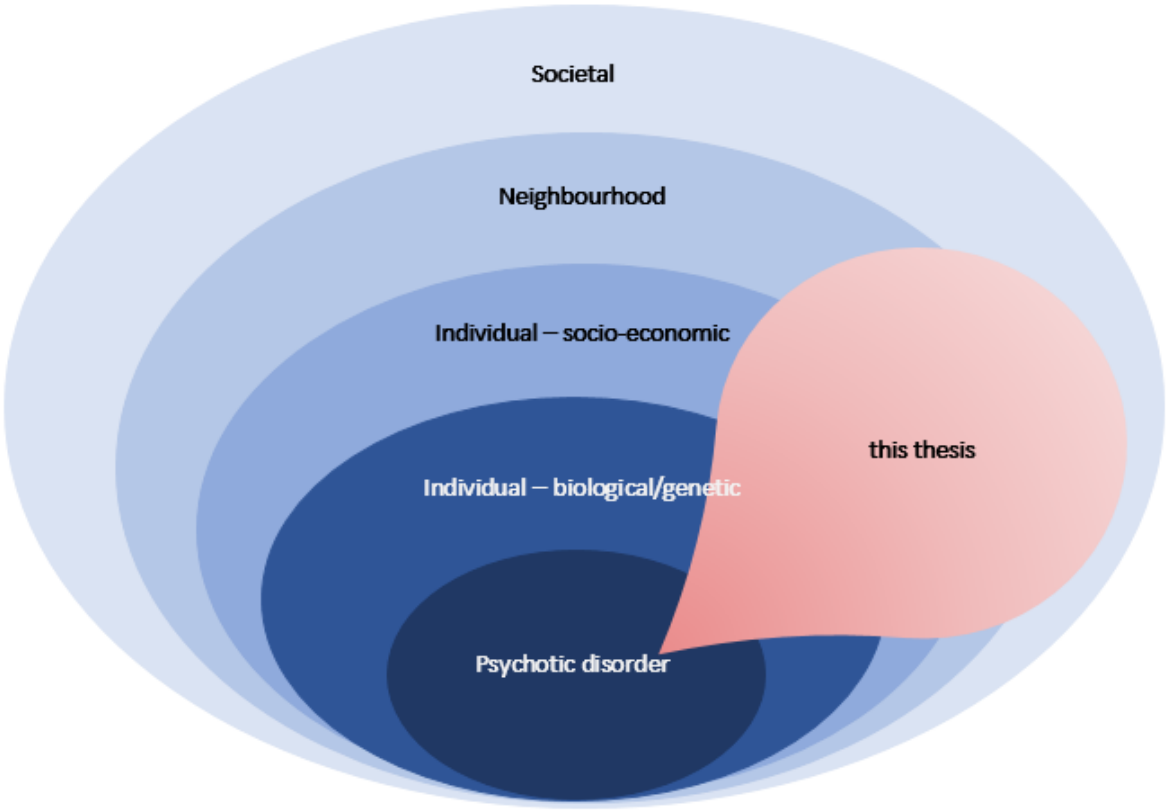
8.5.1 Implications of the aetiology of psychotic disorders

As discussed in Chapter 5, the aetiology of psychotic disorders is multi-causal and these causes are set across multiple layers or systems of causation. I illustrate this schematically in Figure 8.1 below. This figure is intended to illustrate the contributions of my thesis and not to give a complete overview of the

[7] I initially tested the association between language distance (0-3) and psychotic disorders separately, and found a dose-response relationship for language distance.

complexities inherent in psychosis aetiology generally. As such is it highly simplified. In reality the systems of causation also overlap and mutually influence each other, and the more distant (societal/neighbourhood) systems are not less salient, as I indicated in Chapter 5.

Figure 8.1: Levels of causation of psychotic disorders and how this thesis fits in.



The societal level has not been discussed in much detail in this thesis, or elsewhere. In Chapter 2, I have attempted to include a number of country-level exposures (inequality, self-perceived trust and freedom) in the meta-regression, but they were unsuccessful in explaining variance in incidence rates across diagnostic categories. The finding in Chapters 2 and 4 that incidence rates, even in rural and otherwise low-risk areas, were consistently higher in some countries compared with others, indicates that this might be a valuable area for further research (Kirkbride, Hameed, Ankireddypalli, et al., 2017; Richardson, Hameed, Perez, Jones, & Kirkbride, 2017).

However, the main contributions of this thesis to our understanding of psychosis aetiology sits within the neighbourhood (catchment-area) and socio-economic individual levels. The heterogeneity in incidence yielding from a large study using comparable methodology across catchment areas means that the social environment is crucial in the aetiology of psychosis at a more local level (the neighbourhood system in Figure 8.1). Whereas the previous international incidence study (Jablensky et al., 1992) accelerated a search for the 'schizophrenia gene', this study should spur a localised search for modifiable environmental risk factors. In Northern Europe for instance, identifying those elements associated with city living that are

associated with increased incidence would be crucial, whereas in Southern Europe there does not appear to a link between urbanicity and psychotic disorders (at least not in the way seen in Northern Europe), and local knowledge will have to inform research there. The findings on owner-occupancy and the null-finding on latitude indicate that the social environment is more important in determining risk than geographical location and any supposed associated biological mechanism (such as vitamin D deficiency (Kinney et al., 2009)).

This thesis also contributes to our understanding of psychosis on individual socio-economic or 'interpersonal' level. I have researched this in ethnic and religious minorities, and found a central role for the sociocultural distance hypothesis. This demonstrates the importance of socio-economic circumstances, and the unique pressures and challenges faced by (ethnic) minorities, and it has gone some way in identifying what these are. There appears to be something aetiologically relevant about being an outsider, both culturally and socially, and this social outsider status tends to accumulate in those groups that are already somehow different from the majority, be that through their ethnic heritage or, perhaps, through their religion. Neuroscience and primate studies suggest an important role for social stress (Akdeniz, Tost, Streit, et al., 2014; Akdeniz, Tost, & Meyer-Lindenberg, 2014; Lederbogen et al., 2011; McCutcheon et al., 2017; McEwen, 2012; Sapolsky, 2005), but this remains a concept that is difficult to capture in observational epidemiological studies. However, I have shown that the sociocultural distance hypothesis is plausible candidate, and the social environment a fruitful area for more research.

I also don't believe epidemiology alone will be able to identify what is aetiologically relevant about being an outsider, and I think herein lays a further contribution of this thesis. I would never have been able to arrive at the epidemiological hypotheses put forward and tested in Chapters 6 and 7, had I not taken an interdisciplinary approach. I think in order to further understand the mechanism behind the findings presented in this thesis, they will need to be supplemented by qualitative research with ethnic minorities, and experimental studies.

These findings support the central idea of social psychiatry: our brain and its associated disorders are, at least in part, functions of our environment (Abed & Abbas, 2011; Fletcher & Frith, 2009; Ventriglio et al., 2016). The outsider status captured in sociocultural distance hypothesis at the socio-economic individual level may well influence downstream biological processes, meaning that that the probability of a neurobiological deficiency crucially hinges on individual socio-economic circumstances. I have only touched on the individual biological or genetic system, in Chapter 7, and here I found that the sociocultural distance hypothesis could statistically explain any excess risk associated with genetic distance. Whilst this study was underpowered, if this finding were to be replicated in a larger sample, it would support the importance of socio-economic characteristics in explaining differences in biological or genetic risk.

There are many aspects of the aetiology of psychotic disorders this thesis does not touch upon. For instance, a strong epidemiological finding from the incidence study (Chapter 4) supported by wider literature is that risk of disorder (particularly of non-affective disorders) is highest in young men, and this thesis has not examined this in further detail. What this thesis does suggest however, is the appropriate way to research this would be to take a step back from epidemiology and examine what it is about being a young man that increases their risk of psychosis, starting with looking at what pressures they might face resulting from their sociocultural context.

8.5.2 Implications for service provision, policy and public health

The incidence part of this thesis supported one strong implication for service delivery: early intervention in psychosis services globally should follow the most recent National Institute for health and Care Excellence (NICE) guideline (NICE, 2014), and abolish any upper age limit (Lappin et al., 2016). As I discussed in Section 4.3, only 67.9% of male cases and 50.8% of female cases presented before the age of 35. Furthermore, as the overall population distribution is shifting towards older age, it will mean that there simply are more individuals and thereby more cases in the over-35 group. In the UK, all services should follow NICE guidelines, and other countries should follow NICE's example to prevent gender- and age inequitable early intervention services.

The incidence investigation also underscores the importance of good-quality local data collection for service planning. The variation in treated incidence found across the studies included in the systematic review and meta-analyses and the catchment areas included in the EU-GEI study indicates that the healthcare burden varies strongly between local areas. To an extent, incidence can be predicted on the basis of routinely available socio-demographic data, as it varies by known socio-demographic correlates. Perhaps this should be updated to include owner-occupancy or a similar indicator of social cohesion. For England and Wales, a free online prediction tool is available in the form of PsyMaptic (Kirkbride, 2017; Kirkbride et al., 2013), and it would be beneficial if a similar service would be rolled out more widely.

Any policy recommendation concerning the structural drivers of higher rates of psychotic disorders in ethnic minorities is perhaps easy to conceive: reducing structural and cultural barriers to participation of (ethnic) minorities would reduce inequalities. This is, however, difficult to implement. A policy problem like this is known as a 'wicked problem': a social situations where there is no obvious solution, many individuals and stakeholders are necessarily involved, there is disagreement among stakeholders about the appropriate solution, and desired behaviour change is part of the solution (Rittel & Webber, 1973). Solving such problems goes beyond the reach of one agency, and in fact unaligned systemic interventions often have adverse side-effects. As such, these problems are often characterised by chronic policy failure. To succeed, they require a broad systemic response, engaging a wide range of stakeholders and citizens and working across institutional boundaries (Ferlie, Fitzgerald, McGivern, Dopson, & Bennett, 2011; Sullivan &

Skelcher, 2002). The social experience of minorities is subject to many influences, but the main player is the ethnic majority in the host country, and the extent to which they welcome minorities (Collier, 2013).

Not all policy problems specifically pertaining to minority health are difficult to solve. Some physical disorders are particularly common in certain ethnic minority groups, and this can lead to well-tailored services. An example from this in the United Kingdom is sickle cell anaemia, which mainly affects people of African, Caribbean, Middle Eastern, Eastern Mediterranean and Asian descent (NHS Choices, 2016). Recently, for instance, NHS Blood and Transplant has launched a campaign to increase the number of black donors, partially to be able to supply well-matched blood for anaemia-related transfusions (NHS Blood and Transplant, 2017). It is hard to see how such a racialised campaign would work in the area of mental health, and in particular psychotic disorders. Mental illness, and in particular psychosis, still carries a high level of high stigma (Mental Health Foundation, 2017), and specifically tailoring services to minorities would run the risk of playing into (old) stereotypes such as 'big, Black and dangerous'. One of the problems in solving wicked problems is that often policy solutions specifically targeted towards the population group that experiences the problem becomes associated with that group and as such becomes tainted – 'services for the poor become poor services' (Rittel & Webber, 1973). Any concrete policy solution would have to reflect an inclusive approach to minorities across areas of government, and a broader organisation of society that doesn't single out people on the basis of one of their identities ('Black' or 'Muslim') but first and foremost treats everyone as people.

Perhaps one concrete policy example resulting from the case-control data is the importance of acquisition of the majority language for migrants and their descendants for their integration. This should be approached carefully however, and with the objective of inclusion rather than a thinly veiled attempt to reduce migration. Furthermore, this will always only be one element of integration and will not in itself ensure good integration of all minority groups. Language acquisition should be supported regardless of any immigration policies that countries may have: governments should support immigrants into their country to acquire the majority language. An example of how not to do this is given by the current situation in the Netherlands: language acquisition is a component of mandatory integration, which is a prerequisite for obtaining long-term residency or citizenship (for most non-EU immigrants). This has been mandatory since 2007, and had previously successfully been arranged centrally by the government. Unfortunately, in 2013 this has been deregulated and left to private providers, and since the system has become dysfunctional. The responsibility for arranging language tuition (for a language they don't speak) now lies with the migrants themselves rather than with the government, with the predictable consequence of the number of successful language tests passed dropping by 80% compared with the pre-privatisation period (Algemene Rekenkamer, 2017). This leaves many more immigrants without adequate language training and ultimately in a less secure position, as their right to permanent residence is contingent on successful completion of a language test.

8.6 Future research directions

The findings from this thesis perhaps raise more questions than they have answered. In this Section, I outline some of the future research avenues that could follow on from this work. These are divided into two broad themes: delving into more detail and exploring new avenues.

8.6.1 Further details

First, I have only examined broad diagnostic categories (all psychotic disorders, non-affective psychoses and affective psychoses), which reflects current trends in psychiatric epidemiology (see Chapter 3). More recent research has focussed on the idea of symptom domains however, in recognition that traditional DSM or ICD-based diagnostic categories are both too narrow and too broad (Insel et al., 2010). Whilst there is no consensus yet as to what these symptom domains are, positive and negative symptoms, as well as cognitive deficits are widely recognised in psychotic disorders (Chapter 1). A lot of the neuroscientific thinking regarding predictive processing and the brain and Bayesian inference has focussed on the positive symptoms of hallucinations and delusions (Clark, 2016; Fletcher & Frith, 2009), and it would be interesting to see the extent to which social and cultural distance, as well as psychosocial disempowerment, are associated with these psychotic experiences in the general population. This could be tested in the EU-GEI study by using the Community Assessment of Psychic Experience (CAPE) or Structured Interview for Schizotypy-Revised (SIS-R), which assess psychopathology and schizotypy in the general population, as outcome measures, using the same modelling strategy as was used in this thesis.

Second, the focus of the incidence study was to get an overview of the epidemiological landscape, which meant it ran the risk of overlooking more nuanced findings (such as the differential effect of population density within countries) and should be complimented by further research at a more detailed level. As participants' postcodes (or equivalent) were gathered as part of routine data-collection included in the EU-GEI study, neighbourhood-level analysis within catchment areas is possible. It is unlikely that the same neighbourhood-level data is available for all of the six countries, but in England data could be analysed by Census variables such as levels and domains of deprivation, and voter turnout (for instance) as an indicator of social capital.

A similar argument can be made for the results yielding from the case-control analysis. I have appropriately allowed for the hierarchical nature of the dataset, but that does not fully capture the difference in what it is like to be part of an ethnic minority in each catchment area. As I mentioned in the introduction, I am very cautious about making any bold claims about 'ethnic minorities' as effects are likely to differ by ethnic group and by catchment areas. This is illustrated by the intricacies in the effects of ethnic density (see Section 1.2), which appear to vary by ethnic group.

Also, I have not stratified results from the case-control analysis by gender, and the differences in incidence across the life-course for men and women indicate that one size might not fit all. I have simply included sex as a covariate, but this might not fully capture any gendered effects. In Chapter 5, I built the argument that including socio-economic status as a confounder doesn't fully capture the chronic cumulative difference in experience for ethnic minorities (socio-economic status is not simply a confounder but is on the causal pathway from ethnic minority status to ill health). It is possible, although speculative, this might extend to gender too. Gender is an important social identity to which set expectations are attached (although perhaps increasingly less so), so the cumulative social experience of men and women might be very different, and this ought to be researched in more detail.

I have only touched on the finding that being religious is associated with an increased risk of developing psychosis and it would be interesting to explore this in more detail. Potential explanations for this finding are varied, but one suggestion might be that it is mediated by low IQ. A previous study indicates that religious individuals also have a lower IQ (Dutton & Van der Linden, 2017), which is also a known risk factor for psychosis (Khandaker et al., 2011). As a standardised IQ-test (the Wechsler Adult Intelligence Scale) was administered, it would be possible to, within the EU-GEI dataset, repeat or disconfirm this finding, and to test if this would explain the excess risk of psychosis in religious minorities, which was left unexplained by the current model.

8.6.2 New avenues

The way in which I have operationalised some of the constructs included in Chapter 5 was by no means perfect and was largely determined by availability of data within the EU-GEI study. It would therefore be ideal to be able to design a longitudinal study which purposely samples different minority groups, and is able to test for the constructs of cultural distance, social distance and psychosocial disempowerment explicitly. A longitudinal design would diminish the plausibility of reverse causality being an explanation of results. In order to bridge the gap between sociology and epidemiology, such a study would require substantial further thought on how these concepts should ideally be operationalised, and on study design. It would also require collaboration with social science researchers who have experience in this area, and public involvement from ethnic minority individuals with and without a psychotic disorder

I would also like to examine if the theory outlined in Chapter 5 is an explanatory framework for any excess risk in non-heterosexuals, as they too form a marginalised minority group. Whilst there are indications that non-heterosexuals have poorer mental health in general (Chakraborty, McManus, Brugha, Bebbington, & King, 2011; King et al., 2008), are at increased risk of suicide (King et al., 2008; Meader & Chan, 2017) and have a higher prevalence of psychotic experiences (Gevonden et al., 2013), no study has yet tested if non-heterosexuals are at increased risk of psychotic disorders and, if so, what the drivers of this are. This is by

no means meant to pathologise non-heterosexuality, but to adequately understand the burden of disease in a marginalised group with a view to offering appropriate services and preventative strategies.

Finally, due to the limitations inherent to using the PRS in examining excess risk in ethnic minorities, the extent to which gene-environment interactions have been investigated is limited. This is an important area for further research, but I have no concrete suggestions on how to do this. Mendelian randomisation involves using genetic variants in observational epidemiology to make causal inferences about non-genetic modifiable risk factors (Lawlor, Harbord, Sterne, Timpson, & Smith, 2008), by using this genetic variant as a robust proxy for the environmental risk factor. This technique has been heralded as promising in psychosis research (McGrath et al., 2013), but application has been limited. It is difficult to see how this would complement the current investigation: genetic ancestry is not randomly distributed across the population, neither is cultural distance (both are always relative to the majority population). Whereas there are known genetic associations with intelligence (Sniekers et al., 2017), this is not a feasible proxy for educational attainment, as it is not the sole determinant of educational success, and no genetic instruments are available to act as proxies for relationship status and living arrangements.

8.7 Conclusion

This thesis has presented a varied collection of work with two common denominators: substantial variation in psychosis risk, and a role for the sociocultural context in explaining this heterogeneity. Psychosis risk varied by place, and by person. Incidence varied across studies included in the systematic review and meta-analyses as well as across the catchment areas represented in the EU-GEI study. Incidence also varied between groups of people: it was higher in young men, and in ethnic minorities.

When examining this variance by place and person, a role for the sociocultural context emerged. There was some evidence for a population density effect in the systematic review, and within some countries in the EU-GEI study. However, the strongest putative environmental determinant emerging from the incidence study was the percentage of owner-occupied housing. This is a distinctly social characteristic and has been linked to social capital. When investigating the increased risk of psychosis in ethnic minorities, an even more central role for the sociocultural context emerged. The sociocultural distance hypotheses developed in this thesis appears to explain a large proportion of excess risk in ethnic minority groups: a statistical model including indicators of social distance, cultural distance and psychosocial disempowerment explained most of the variance in psychosis risk after appropriately controlling for known covariates. Only the Black group remained at increased risk of all psychotic disorders, but even this was substantially attenuated. One of the strongest predictors was the cultural distance of ethnic minorities. Furthermore, this same model statistically explained the excess risk in those with increased African ancestry in the EU-GEI subsample recruited in England, tentatively indicating it is more important

than a genetic determinant of ethnic minority status. These novel findings are a significant advance in our understanding of inequities in psychosis risk in minority groups.

Causality in psychosis remains multi-faceted, it is too early to draw any conclusions on causality and my findings should be subject to replication and scrutiny elsewhere. Nevertheless, my thesis adds further evidence that the sociocultural context plays a crucial role in explaining heterogeneity in psychosis risk.

References

1000 Genomes Project Consortium. (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature*, *491*(7422), 56–65. http://doi.org/10.1038/nature11632

1000 Genomes Project Consortium. (2015). A global reference for human genetic variation. *Nature*, *526*(7571), 68–74. http://doi.org/10.1038/nature15393

Abed, R. T., & Abbas, M. J. (2011). A reformulation of the social brain theory for schizophrenia: the case for out-group intolerance. *Perspectives in Biology and Medicine*, *54*(2), 132–151. http://doi.org/10.1353/pbm.2011.0020

Abraham, J. E., Maranian, M. J., Russell, R., Ingle, S., Luccarini, C., Earl, H. M., … Caldas, C. (2012). Saliva samples are a viable alternative to blood samples as a source of DNA for high throughput genotyping. *BMC Medical Genomics* , *5*(19). http://doi.org/10.1016/S0076-6879(06)10003-8

Afshari, R., & Bhopal, R. S. (2002). Changing pattern of use of "ethnicity" and "race" in scientific literature. *International Journal of Epidemiology*, *31*(5), 1074–1074. http://doi.org/10.1093/ije/31.5.1074

Ahmad, W. I. U., & Bradby, H. (2007). Locating ethnicity and health: exploring concepts and contexts. *Sociology of Health & Illness*, *29*(6), 795–810. http://doi.org/10.1111/j.1467-9566.2007.01051.x

Akdeniz, C., Tost, H., & Meyer-Lindenberg, A. (2014). The neurobiology of social environmental risk for schizophrenia: an evolving research field. *Social Psychiatry and Psychiatric Epidemiology*, *49*(4), 507–517. http://doi.org/10.1007/s00127-014-0858-4

Akdeniz, C., Tost, H., Streit, F., Haddad, L., Wüst, S., Schäfer, A., … Meyer-Lindenberg, A. (2014). Neuroimaging Evidence for a Role of Neural Social Stress Processing in Ethnic Minority-Associated Environmental Risk. *JAMA Psychiatry*, *71*(6), 1–9. http://doi.org/10.1001/jamapsychiatry.2014.35

Akerlof, G., & Kranton, R. (2011). *Identity Economics: How Our Identities Shape Our Work, Wages and Well-being.* Princeton, New Jersey: Princeton University Press.

Akiyama, T. (1996). Onset study of english-speaking temporary residents in Japan. *Social Psychiatry and Psychiatric Epidemiology*, *31*(3–4), 194–198. http://doi.org/10.1007/BF00785767

Alamanos, Y., & Drossos, A. A. (2005). Epidemiology of adult rheumatoid arthritis. *Autoimmunity Reviews*, *4*(3), 130–136. http://doi.org/10.1016/j.autrev.2004.09.002

Aleman, A., Hijman, R., De Haan, E. H. F., & Kahn, R. S. (1999). Memory impairment in schizophrenia: a meta-analysis. *Am J Psychiatry*, *156*(9), 1358–1366.

Aleman, A., Kahn, R. S., & Selten, J.-P. (2003). Sex differences in the risk of schizophrenia. *Archives of General Psychiatry*, *60*(6), 565. http://doi.org/10.1001/archpsyc.60.6.565

Algemene Rekenkamer. (2017). Inburgering: Eerste resultaten van de Wet Inburgering 2013.

Allardyce, J., Gilmour, H., Atkinson, J., Rapson, T., Bishop, J., & McCreadie, R. G. (2005). Social fragmentation, deprivation and urbanicity : relation to first-admission rates for psychoses. *British Journal of Psychiatry*, *187*(May), 401–406.

Alonso, J., Lépine, J.-P., Brugha, S., De Girolamo, G., De Graaf, R., Demyttenaere, K., … Alonso, J. (2007). Overview of key data grom the European Study of the Epidemiology of Mental Disorders

(ESEMeD) on behalf of the ESEMeD/MHEDEA 2000 Scientific Committee. *J Clin Psychiatry Italy*, *6868*(20028), 3–9.

Amaral, E. F., & Fusco, W. (2005). Shaping Brazil: The Role of International Migration. Retrieved August 8, 2017, from http://www.migrationpolicy.org/article/shaping-brazil-role-international-migration

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders: DSM-IV* (4th editio). Washington (DC): American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and Statisictal Manual of Mental Disorders* (5th editio). Arlington, VA: American Psychiatric Association.

Anderson, K. K., Cheng, J., Susser, E., McKenzie, K. J., & Kurdyak, P. (2015). Incidence of psychotic disorders among first-generation immigrants and refugees in Ontario. *Canadian Medical Association Journal*, *187*(9), E279–E286. http://doi.org/10.1503/cmaj.150494

Anderson, K. K., Fuhrer, R., Abrahamowicz, M., & Malla, A. K. (2012). The incidence of first-episode schizophrenia-spectrum psychosis in adolescents and young adults in Montreal: An estimate from an administrative claims database. *Canadian Journal of Psychiatry*, *57*(10), 626–633. http://doi.org/10.1177/070674371205701007

Andreasen, N. C., Flaum, M., & Arndt, S. (1992). The Comprehensive Assessment of Symptoms and History (CASH). An instrument for assessing diagnosis and psychopathology. *Archives of General Psychiatry*, *49*(8), 615–23.

Anglin, D. (2017). SA59. The psychophysiological effects of racially-primed social exclusion among ethnic minorities with attenuated psychotic symptoms. In *ICOSR*.

Arseneault, L., Cannon, M., Poulton, R., Murray, R., Caspi, A., & Moffitt, T. E. (2002). Cannabis use in adolescence and risk for adult psychosis: longitudinal prospective study. *BMJ (Clinical Research Ed.)*, *325*(7374), 1212–1213. http://doi.org/10.1136/bmj.325.7374.1212

Azevedo, M. H., Soares, M. J., Coelho, I., Dourado, A., Valente, J., Macedo, A., … Pato, C. (1999). Using consensus OPCRIT diagnoses. An efficient procedure for best-estimate lifetime diagnoses. *The British Journal of Psychiatry*, *175*(2).

Babiker, I. E., Cox, J. L., & Miller, P. M. (1980). Social Psychiatry The Measurement of Cultural Distance and Its Relationship to Medical Consultations, Symptomatology and Examination Performance of Overseas Students at Edinburgh University*. *Social Psychiatry*, *15*, 109–116.

Baker, K. D., & Skuse. (2005). Adolescents and young adults with 22q11 deletion syndrome: psychopathology in an at-risk group. *The British Journal of Psychiatry*, *186*(2).

Barnett, J. H., Sahakian, B. J., Werners, U., Hill, K. E., Brazil, R., Gallagher, O., … Jones, P. B. (2005). Visuospatial learning and executive function are independently impaired in first-episode psychosis. *Psychological Medicine*, *35*(7), 1031–41. http://doi.org/10.1017/S0033291704004301

Bassett, A. S., Chow, E. W. C., AbdelMalik, P., Gheorghiu, M., Husted, J., & Weksberg, R. (2003). The Schizophrenia Phenotype in 22q11 Deletion Syndrome. *American Journal of Psychiatry*, *160*(9), 1580–1586. http://doi.org/10.1176/appi.ajp.160.9.1580

Bassett, A. S., Chow, E. W. C., Husted, J., Weksberg, R., Caluseriu, O., Webb, G. D., & Gatzoulis, M. A. (2005). Clinical features of 78 adults with 22q11 deletion syndrome. *American Journal of Medical Genetics Part A*, *138A*(4), 307–313. http://doi.org/10.1002/ajmg.a.30984

Bécares, L., Nazroo, J., & Stafford, M. (2009). The buffering effects of ethnic density on experienced

racism and health. *Health & Place*, *15*, 700–708. http://doi.org/10.1016/j.healthplace.2008.10.008

Benabou, R., & Tirole, J. (2011). Identity, Morals and Taboos: Beliefs as Assets. *The Quarterly Journal of Economics*, *126*, 805–855.

Bendall, S., Jackson, H. J., Hulbert, C. A., & McGorry, P. D. (2008). Childhood trauma and psychotic disorders: a systematic, critical review of the evidence. *Schizophrenia Bulletin*, *34*(3), 568–79. http://doi.org/10.1093/schbul/sbm121

Berg, A. O., Melle, I., Rossberg, J. I., Romm, K. L., Larsson, S., Lagerberg, T. V, … Hauff, E. (2011). Perceived discrimination is associated with severity of positive and depression/anxiety symptoms in immigrants with psychosis: a cross-sectional study. *BMC Psychiatry*, *11*(1), 77. http://doi.org/10.1186/1471-244X-11-77

Bernstein, D. P., Stein, J. A., Newcomb, M. D., Walker, E., Pogge, D., Ahluvalia, T., … Zule, W. (2003). Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child Abuse & Neglect*, *27*(2), 169–190. http://doi.org/10.1016/S0145-2134(02)00541-0

Bhavsar, V., Boydell, J., Murray, R., & Power, P. (2014). Identifying aspects of neighbourhood deprivation associated with increased incidence of schizophrenia. *Schizophrenia Research*, *156*(1), 115–121. http://doi.org/10.1016/j.schres.2014.03.014

Bhopal, R. (1997). Is research into ethnicity and health racist, unsound, or important science? *BMJ (Clinical Research Ed.)*, *314*(7096), 1751–6. http://doi.org/10.1136/BMJ.314.7096.1751

Bhopal, R. (2004). Glossary of terms relating to ethnicity and race: for reflection and debate. *Journal of Epidemiology and Community Health*, *58*(6), 441–5. http://doi.org/10.1136/JECH.2003.013466

Bhugra, D., Gupta, S., Bhui, K., Craig, T., Dogra, N., Ingleby, J. D., … Tribe, R. (2011). WPA guidance on mental health and mental health care in migrants. *World Psychiatry*, *10*(1), 2–10. http://doi.org/10.1002/j.2051-5545.2011.tb00002.x

Bhugra, D., Hilwig, M., Hossein, B., Marceau, H., Neehall, J., Leff, J., … Der, G. (1996). First-contact incidence rates of schizophrenia in Trinidad and one-year follow-up. *British Journal of Psychiatry*, *169*(5), 587–592. http://doi.org/10.1192/bjp.169.5.587

Bleuler, E. (1911). *Dementia Praecox oder Gruppe der Schizophrenien*. Tubingen: Diskord.

Bogardus, E. S. (1926). Social Distance in the City. *Proceedings and Publications of the American Sociological Society*, *20*, 40–46.

Bonifazi, C., Heins, F., Strozza, S., & Vitiello, M. (2009). Italy: The Italian transition from an emigration to immigration country. *IDEA Working Papers*, (5).

Boonstra, N., Wunderink, L., De Wit, P. H. M., Noorthoorn, E., & Wiersma, D. (2008). De administratieve incidentie van niet-affectieve psychosen in Friesland en Twente. *Tijdschrift Voor Psychiatrie*, *50*(10), 637–643.

Boonstra, N., Wunderink, L., de Wit, P., Noorthoorn, E., & Wiersma, D. (2008). De administratieve incidentie van niet-affective psychosen in Friesland en Twente. *Tijdschrift Voor Psychiatrie*, *50*(10), 637–643.

Bora, E., Yucel, M., & Pantelis, C. (2009). Theory of mind impairment in schizophrenia: Meta-analysis. *Schizophrenia Research*, *109*(1–3), 1–9. http://doi.org/10.1016/j.schres.2008.12.020

Borgenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009). *Introduction to Meta-Analysis*. Chichester, West Sussex: John Wiley & Sons Ltd.

Bourque, F., van der Ven, E., & Malla, A. (2011). A meta-analysis of the risk for psychotic disorders among first- and second-generation immigrants. *Psychological Medicine*, *41*(5), 897–910. http://doi.org/10.1017/S0033291710001406

Bover, O., & Velilla, P. (1997). *Migrations in Spain: Historical Background and Current Trends*.

Boydell, J., Van Os, J., Lambri, M., Castle, D., Allardyce, J., McCreadie, R. G., & Murray, R. M. (2003). Incidence of schizophrenia in south-east London between 1965 and 1997. *British Journal of Psychiatry*, *182*(JAN.), 45–49. http://doi.org/10.1192/bjp.182.1.45

Boydell, J., van Os, J., McKenzie, K., & Murray, R. M. (2004). The association of inequality with the incidence of schizophrenia - An ecological study. *Social Psychiatry and Psychiatric Epidemiology*, *39*(8), 597–599. http://doi.org/10.1007/s00127-004-0789-6

Bradford Hill, A. (1965). The Environment and Disease: Association or Causation? *Proceedings of the Royal Society of Medicine*, *58*(5), 295–300.

Bramon, E., Walshe, M., Mcdonald, C., Martín, B., Toulopoulou, T., Wickham, H., … Murray, R. M. (2005). Dermatoglyphics and Schizophrenia: A meta-analysis and investigation of the impact of obstetric complications upon a–b ridge count. *Schizophrenia Research*, *75*, 399–404. http://doi.org/10.1016/j.schres.2004.08.022

Bromet, E. J., Naz, B., Fochtmann, L. J., Carlson, G. A., & Tanenberg-Karant, M. (2005). Long-term diagnostic stability and outcome in recent first-episode cohort studies of schizophrenia. *Schizophrenia Bulletin*, *31*(3), 639–649. http://doi.org/10.1093/schbul/sbi030

Bromet, E. J., Schwartz, J. E., Fennig, S., Geller, L., Jandorf, L., Kovasznay, B., … Ram, R. (1992). The epidemiology of psychosis: the Suffolk County Mental Health Project. *Schizophrenia Bulletin*, *18*(2), 243–55.

Brown, A. S., & Derkits, E. J. (2010). Prenatal infection and schizophrenia: a review of epidemiologic and translational studies. *The American Journal of Psychiatry*, *167*(3), 261–80. http://doi.org/10.1176/appi.ajp.2009.09030361

Brown, S. (1997). Excess mortality of schizophrenia. A meta-analysis. *The British Journal of Psychiatry*, *171*(6).

Burns, J. K., & Esterhuizen, T. (2008). Poverty, inequality and the treated incidence of first-episode psychosis. An ecological study from South Africa. *Social Psychiatry and Psychiatric Epidemiology*, *43*(4), 331–335. http://doi.org/10.1007/s00127-008-0308-2

Candelo, N., Croson, R. T. A., & Li, S. X. (2016). Identity and social exclusion: an experiment with Hispanic immigrants in the U.S. *Experimental Economics*. http://doi.org/10.1007/s10683-016-9492-1

Cannon, M., Jones, P. B., & Murray, R. M. (2002). Obstetric complications and schizophrenia: historical and meta-analytic teview. *American Journal of Psychiatry*, *159*(7), 1080–1092. http://doi.org/10.1176/appi.ajp.159.7.1080

Cantor-Graae, E., & Selten, J.-P. (2005). Schizophrenia and migration: a meta-analysis and review. *American Journal of Psychiatry*, *162*(1), 12–24. http://doi.org/10.1176/appi.ajp.162.1.12

Carlborg, A., Ferntoft, L., Thuresson, M., & Bodegard, J. (2015). Population study of disease burden, management, and treatment of bipolar disorder in Sweden: A retrospective observational

registry study. *Bipolar Disorders*, *17*(1), 76–85. http://doi.org/10.1111/bdi.12234

Cartwright, N. (1989). *Nature's Capacities and their Measurement*. Oxford: Oxford University Press.

Centraal Bureau voor de Statistiek. (2016). Bevolking naar migratieachtergrond. Retrieved August 8, 2017, from https://www.cbs.nl/nl-nl/achtergrond/2016/47/bevolking-naar-migratieachtergrond

Central Intelligence Agency. (2017). The World Factbook - Brazil. Retrieved August 8, 2017, from https://www.cia.gov/library/publications/the-world-factbook/geos/br.html

Chakraborty, A., McManus, S., Brugha, T. S., Bebbington, P., & King, M. (2011). Mental health of the non-heterosexual population of England. *The British Journal of Psychiatry*, *198*(2).

Chandola, T., Brunner, E., & Marmot, M. (2006). Chronic stress at work and the metabolic syndrome: prospective study. *BMJ (Clinical Research Ed.)*, *332*(7540), 521–5. http://doi.org/10.1136/bmj.38693.435301.80

Chang, C.-K., Hayes, R. D., Perera, G., Broadbent, M. T. M., Fernandes, A. C., Lee, W. E., … Stewart, R. (2011). Life expectancy at birth for people with serious mental illness and other major disorders from a secondary mental health care case register in London. *PloS One*, *6*(5), e19590. http://doi.org/10.1371/journal.pone.0019590

Chen, D. T., Jiang, X., Akula, N., Shugart, Y. Y., Wendland, J. R., Steele, C. J. M., … McMahon, F. J. (2013). Genome-wide association study meta-analysis of European and Asian-ancestry samples identifies three novel loci associated with bipolar disorder. *Molecular Psychiatry*, *18*(2), 195–205. http://doi.org/10.1038/mp.2011.157

Cheng, F., Kirkbride, J. B., Lennox, B. R., Perez, J., Masson, K., Lawrence, K., … Jones, P. B. (2011). Administrative incidence of psychosis assessed in an early intervention service in England: first epidemiological evidence from a diverse, rural and urban setting. *Psychological Medicine*, *41*, 949–958. http://doi.org/10.1017/S0033291710002461

Chiswick, B. R., & Miller, P. W. (2004). Linguistic distance: a quantitative measure of the distance between English and other languages. *IZA Discussion Paper Series*, (1246).

Clark, A. (2016). *Surfing Uncertainty: Prediction, Action and the Embodied Mind*. Oxford: Oxford University Press.

Cocchi, A., Balbi, A., Corlito, G., Ditta, G., Di Munzio, W., Nicotera, M., … Preti, A. (2015). Early intervention in psychosis: A feasibility study financed by the Italian Center on Control of Maladies. *Early Intervention in Psychiatry*, *9*(2), 163–171. http://doi.org/10.1111/eip.12135

Cohen, A. N., Hamilton, A. B., Saks, E. R., Glover, D. L., Glynn, S. M., Brekke, J. S., & Marder, S. R. (2017). How Occupationally High-Achieving Individuals With a Diagnosis of Schizophrenia Manage Their Symptoms. *Psychiatric Services*, *68*(4), 324–329. http://doi.org/10.1176/appi.ps.201600031

Coid, J. W., Kirkbride, J. B., Barker, D., Cowden, F., Stamps, R., Yang, M., & Jones, P. B. (2008). Raised incidence rates of all psychoses among migrant groups: findings from the East London First Episode Psychosis Study. *Arch.Gen.Psychiatry*, *65*(11), 1250–1258. http://doi.org/10.1001/archpsyc.65.11.1250

Collier, P. (2013). *Exodus: How Migration is Changing Our World.* New York: Oxford University Press.

Colodro-Conde, L., Couvy-Duchesne, B., Whitfield, J., Streit, F., Gordon, S., Rietschel, M., … Martin, N. G. (2017). Higher genetic risk for schizophrenia is associated with living in urban and populated areas. *BioRxiv*. http://doi.org/10.1101/179432

Cooper, B. (2005). Immigration and schizophrenia: the social causation hypothesis revisited. *The British Journal of Psychiatry*, *186*(5).

Cooper, H. (2002). Investigating socio-economic explanations for gender and ethnic inequalities in health. *Social Science and Medicine*, *54*, 693–706.

Cooper, J. E., Goodhead, D., Craig, T., Harris, M., Howat, J., & Korer, J. (1987). The incidence of schizophrenia in Nottingham. *The British Journal of Psychiatry*, *151*(5), 619–626. http://doi.org/10.1192/bjp.151.5.619

Cooper, R., & David, R. (1986). The Biological Application to Epidemiology Concept of Race and its Public Health and. *Journal of Health Politics, Policy & Law*, *11*(1), 97–116.

Craddock, M., Asherson, P., Owen, M. J., Williams, J., McGuffin, P., & Farmer, A. E. (1996). Concurrent validity of the OPCRIT diagnostic system. Comparison of OPCRIT diagnoses with consensus best-estimate lifetime diagnoses. *The British Journal of Psychiatry*, *169*(1), 58–63. http://doi.org/10.1192/bjp.169.1.58

Craddock, N., & Sklar, P. (2013). Bipolar disorder 1: Genetics of bipolar disorder. *The Lancet*, *381*, 1654–62. http://doi.org/10.1016/S0140-6736(13)60855-7

Crebbin, K., Mitford, E., Paxton, R., & Turkington, D. (2008). First-episode psychosis: An epidemiological survey comparing psychotic depression with schizophrenia. *Journal of Affective Disorders*, *105*(1–3), 117–124. http://doi.org/10.1016/j.jad.2007.04.025

Crebbin, K., Mitford, E., Paxton, R., & Turkington, D. (2009). First-episode drug-induced psychosis: A medium term follow up study reveals a high-risk group. *Social Psychiatry and Psychiatric Epidemiology*, *44*(9), 710–715. http://doi.org/10.1007/s00127-008-0490-2

Cross-Disorder Group of the Psychiatric Genetics Consortium. (2013). Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *The Lancet*, *381*, 1371–1379. http://doi.org/10.1016/S0140-6736(12)62129-1

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. *Archives of General Psychiatry*, *63*(3), 267–272. http://doi.org/10.1001/archpsyc.63.3.267

Davies, A. A., Basten, A., & Frattini, C. (2006). *Migration: A Social Determinant of the Health of Migrants*.

de Candia, T. R., Lee, S. H., Yang, J., Browning, B. L., Gejman, P. V., Levinson, D. F., … Keller, M. C. (2013). Additive genetic variation in schizophrenia risk is shared by populations of African and European descent. *The American Journal of Human Genetics*, *93*(3), 463–470. http://doi.org/10.1016/j.ajhg.2013.07.007

Dean, L. (2012). *Schizophrenia*. *Medical Genetics Summaries*. National Center for Biotechnology Information (US).

Dein, S., & Littlewood, R. (2011). Religion and psychosis: A common evolutionary trajectory? *Transcultural Psychiatry*, *48*(3), 318–335. http://doi.org/10.1177/1363461511402723

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, *7*, 177–188.

Dijksterhuis, A. (2005). Why we are social animals: the high road to imitation as social glue. In S. Hurley & N. Chater (Eds.), *Imitation, Human Development, and Culture*. Cambridge, Massachusetts: The MIT Press.

Dohrenwend, B. P., & Dohrenwend, B. S. (1969). *Social Status and Psychological Disorder: A Causal Inquiry*. Hoboken, New Jersey: Wiley-Interscience.

Drukker, M., Krabbendam, L., Driessen, G., & van Os, J. (2006). Social disadvantage and schizophrenia: A combined neighbourhood and individual-level analysis. *Social Psychiatry and Psychiatric Epidemiology*, *41*(8), 595–604. http://doi.org/10.1007/s00127-006-0081-z

Dudbridge, F. (2013). Power and predictive accuracy of polygenic risk scores. *PLoS Genet*, *9*(310). http://doi.org/10.1371/journal.pgen.1003348

Dutton, E., & Van der Linden, D. (2017). Why is intelligence negatively associated with religiousness? *Evolutionary Psychological Science*, 1–12. http://doi.org/10.1007/s40806-017-0101-0

Egerton, A., Howes, O. D., Houle, S., Mckenzie, K., Valmaggia, L. R., Bagby, M. R., … Mizrahi, R. (2017). Elevated striatal dopamine function in immigrants and their children: a risk mechanism for psychosis. *Schizophrenia Bulletin*, *43*(2), 293–301. http://doi.org/10.1093/schbul/sbw181

Ehrenreich, B. (2008). *Nickel and Dimed: On (not) getting by in America*. New York: Henry Holt and Company, LLC.

Encyclopaedia Brittanica. (2017). Latitude and longitude | geography. Retrieved December 16, 2017, from https://www.britannica.com/science/latitude

Errazuriz, A. (2013). *The Mental Health of Peruvian Immigrants in Santiago, Chile.* University of Cambridge.

Esterberg, M. L., Trotman, H. D., Holtzman, C., Compton, M. T., & Walker, E. F. (2010). The impact of a family history of psychosis on age-at-onset and positive and negative symptoms of schizophrenia: A meta-analysis. *Schizophrenia Research*, *120*(1–3), 121–130. http://doi.org/10.1016/j.schres.2010.01.011

EuroHIV. (2007). *HIV/AIDS surveillance in Europe. End-year report 2006*.

European Centre for Disease Surveillance and Control. (2010). *Tuberculosis surveillance in Europe 2008*.

European Commission. (2013). Population and Housing Census. Census Database. Retrieved August 11, 2016, from http://ec.europa.eu/eurostat/web/population-and-ho

European Network of National Schizophrenia Networks Studying Gene-Environment Interactions. (2009a). Identifying genetic and environmental interactions in schizophrenia. Annex I: Description of Work.

European Network of National Schizophrenia Networks Studying Gene-Environment Interactions. (2009b). Research Programme. Retrieved from http://www.eu-gei.eu/about-the-project/research-programme

European Network of National Schizophrenia Networks Studying Gene-Environment Interactions. (2013). EU-GEI study protocol (Cambridgeshire and Peterborough).

European Parliament. (2017). Free movement of persons | EU fact sheets. Retrieved August 8, 2017, from http://www.europarl.europa.eu/atyourservice/en/displayFtu.html?ftuId=FTU_2.1.3.html

European Union Agency for Fundamental Rights. (2017). *Second European Union Minorities and Discrimination Survey. Muslims – Selected findings*. http://doi.org/10.2811/443810

Faris, R., & Dunham, H. (1939). *Mental disorders in urban areas: An ecological study of schizophrenia*

and other psychoses. Chicago/London: The University of Chicago Press.

Fearon, P., Kirkbride, J. B., Morgan, C., Dazzan, P., Morgan, K., Lloyd, T., … Murray, R. M. (2006). Incidence of schizophrenia and other psychoses in ethnic minority groups: results from the MRC AESOP Study. *Psychological Medicine*, *36*(11), 1541–50. http://doi.org/10.1017/S0033291706008774

Fenton, S., & Charlsley, K. (2000). Epidemiology and sociology as incommensurate games: accounts from the study of health and ethnicity. *Health*, *4*(4), 403–425.

Ferlie, E., Fitzgerald, L., McGivern, G., Dopson, S., & Bennett, C. (2011). Public policy networks and "wicked problems": a nascent solution? *Public Administration*, *89*(2), 307–324.

Fioravanti, M., Carlone, O., Vitale, B., Cinti, M. E., & Clare, L. (2005). A meta-analysis of cognitive deficits in adults with a diagnosis of schizophrenia. *Neuropsychology Review*, *15*(2), 73–95. http://doi.org/10.1007/s11065-005-6254-9

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. (1996). *Structured Clinical Interview for DSM-IV Axis I disorders, Clinician Version (SCID-CV)*. Washington (DC): American Psychiatric Press, Inc.

Fisher, M., & Baum, F. (2010). The social determinants of mental health: implications for research and health promotion. *The Australian and New Zealand Journal of Psychiatry*, *44*(January), 1057–1063. http://doi.org/10.3109/00048674.2010.509311

Fletcher, P. C., & Frith, C. D. (2009). Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews Neuroscience*, *10*(1), 48–58. http://doi.org/10.1038/nrn2536

Focus Migration. (2007a). Country Profile - Brazil. Retrieved August 8, 2017, from http://focus-migration.hwwi.de/uploads/tx_wilpubdb/CP_15_brazil_01.pdf

Focus Migration. (2007b). Country Profile - France. Retrieved August 8, 2017, from http://focus-migration.hwwi.de/typo3_upload/groups/3/focus_Migration_Publikationen/Laenderprofile/LP02_Frankreich_v2.pdf

Focus Migration. (2007c). Country Profile - Spain. Retrieved August 8, 2017, from http://focus-migration.hwwi.de/typo3_upload/groups/3/focus_Migration_Publikationen/Laenderprofile/CP_06_Spain-08.pdf

Focus Migration. (2007d). Country Profile - The Netherlands. Retrieved August 8, 2017, from http://focus-migration.hwwi.de/The-Netherlands.2644.0.html?L=1

Furman, K. (2016). *AIDS Denialism in South Africa: A case study in the rationality and ethics of science policy*. London School of Economics and Political Sciences.

Garety, P. A., Kuipers, E., Fowler, D., Freeman, D., Bebbington, & P. E. (2001). A cognitive model of the positive symptoms of psychosis. *Psychological Medicine*, *31*(2), 189–195. http://doi.org/10.1017/S0033291701003312

Gayer-Anderson, C., & Morgan, C. (2013). Social networks, support and early psychosis: a systematic review. *Epidemiology and Psychiatric Sciences*, *22*, 131–146. http://doi.org/10.1017/S2045796012000406

Gearing, R. E., Alonzo, D., Smolak, A., McHugh, K., Harmon, S., & Baldwin, S. (2011). Association of religion with delusions and hallucinations in the context of schizophrenia: Implications for engagement and adherence. *Schizophrenia Research*, *126*(1–3), 150–163. http://doi.org/10.1016/j.schres.2010.11.005

Gevonden, M. J., Selten, J. P., Myin-Germeys, I., de Graaf, R., ten Have, M., van Dorsselaer, S., … Veling, W. (2013). Sexual minority status and psychotic symptoms: findings from the Netherlands Mental Health Survey and Incidence Studies (NEMESIS). *Psychological Medicine*, *44*(2), 421–33. http://doi.org/10.1017/S0033291713000718

Giegling, I., Hosak, L., Mössner, R., Serretti, A., Bellivier, F., Claes, S., … Rujescu, D. (2017). Genetics of schizophrenia: A consensus paper of the WFSBP Task Force on Genetics. *The World Journal of Biological Psychiatry*, 1–14. http://doi.org/10.1080/15622975.2016.1268715

Gigantesco, A., Lega, I., & Picardi, A. (2012). The Italian SEME surveillance system of severe mental disorders presenting to community mental health services. *Clinical Practice and Epidemiology in Mental Health : CP & EMH*, *8*, 7–11. http://doi.org/10.2174/1745017901208010007

Giggs, J. A., & Cooper, J. E. (1987). Ecological structure and the distribution of schizophrenia and affective psychoses in Nottingham. *The British Journal of Psychiatry*, *151*(5).

Girton College. (2017). Girton's Past. Retrieved July 12, 2017, from https://www.girton.cam.ac.uk/girtons-past

Gissler, M., Alexander, S., MacFarlane, A., Small, R., Stray-Pedersen, B., Zeitlin, J., … for the ROAM collaboration. (2009). Stillbirths and infant deaths among migrants in industrialized countries. *Acta Obstetricia et Gynecologica Scandinavica*, *88*(2), 134–148. http://doi.org/10.1080/00016340802603805

Gold, J., & Gold, I. (2014). *Suspicious Minds: How Culture Shapes Madness*. New York: Free Press.

Goldberg, D., & Huxley, P. (1980). *Mental Illness in the Community: The Pathway to Psychiatric care*. London: Tavistock Publications.

Gould, M., Theodore, K., Pilling, S., Bebbington, P., Hinton, M., & Johnson, S. (2006). Initial treatment phase in early psychosis: Can intensive home treatment prevent admission? *Psychiatric Bulletin*, *30*(7), 243–246. http://doi.org/http://dx.doi.org/10.1192/pb.30.7.243

Grigoriadis, S., & Seeman, M. V. (2002). The role of estrogen in schizophrenia: implications for schizophrenia practice guidelines for women. *Canadian Journal of Psychiatry*, *47*(5), 437–442.

Häfner, H., Riecher-Rössler, A., An Der Heiden, W., Maurer, K., Fätkenheuer, B., Löffler, W., … VASSAR, P. (1993). Generating and testing a causal explanation of the gender difference in age at first onset of schizophrenia. *Psychological Medicine*, *23*(4), 925. http://doi.org/10.1017/S0033291700026398

Hanoeman, M., Selten, J.-P., & Kahn, R. S. (2002). Incidence of schizophrenia in Surinam. *Schizophrenia Research*, *54*(3), 219–21. http://doi.org/10.1016/S0920-9964(01)00269-9

Harbord, R., Harris, R., & Sterne, J. (2009). Updated tests for small-study effects in meta-analyses. In J. A. Sterne (Ed.), *Meta-Analysis in Stata: An Updated Collection from the Stata Journal.* (pp. 138–150). College Station, TX: Stata Press.

Hardoon, S., Hayes, J. F., Blackburn, R., Petersen, I., Walters, K., Nazareth, I., & Osborn, D. P. J. (2013). Recording of severe mental illness in United Kingdom primary care, 2000-2010. *PLoS ONE*, *8*(12). http://doi.org/10.1371/journal.pone.0082365

Hare, E. H. (1956). Mental illness and social conditions in Bristol. *Journal of Mental Science*, *102*, 349–357. http://doi.org/10.1192/bjp.102.427.349

Harrison, E., & Rose, D. (2006). *The European Socio-economic Classification (ESeC) User Guide*. Colchester, Essex.

Harrison, P. J. (2015). Recent genetic findings in schizophrenia and their therapeutic relevance. *Journal of Psychopharmacology*, *29*(2), 85–96. http://doi.org/10.1177/0269881114553647

Hartz, S. M., Horton, A. C., Oehlert, M., Carey, C. E., Agrawal, A., Bogdan, R., … Beirut, L. J. (2017). Association between substance use disorder and polygenic liability to schizophrenia. *Biological Psychiatry*, *82*(10), 709–715. http://doi.org/10.1016/J.BIOPSYCH.2017.04.020

Hayes, J. F., Marston, L., Walters, K., King, M. B., & Osborn, D. P. (2017). Mortality gap for people with bipolar disorder and schizophrenia: UK-based cohort study 2000-2014. *The British Journal of Psychiatry*.

Hayes, J. F., Miles, J., Walters, K., King, M., & Osborn, D. P. J. (2015). A systematic review and meta-analysis of premature mortality in bipolar affective disorder. *Acta Psychiatrica Scandinavica*, *131*(6), 417–425. http://doi.org/10.1111/acps.12408

Hennekens, C. H., Hennekens, A. R., Hollar, D., & Casey, D. E. (2005). Schizophrenia and increased risks of cardiovascular disease. *American Heart Journal*, *150*(6), 1115–1121. http://doi.org/10.1016/j.ahj.2005.02.007

Henquet, C., Murray, R., Linszen, D., & Van Os, J. (2005). The environment and schizophrenia: The role of cannabis use. *Schizophrenia Bulletin*, *31*(3), 608–612. http://doi.org/10.1093/schbul/sbi027

Hertzman, C., & Boyce, T. (2010). How experience gets under the skin to create gradients in developmental health. *Annu. Rev. Public Health*, *31*, 329–47. http://doi.org/10.1146/annurev.publhealth.012809.103538

Heslin, M., Lomas, B., Lappin, J. M., Donoghue, K., Reininghaus, U., Onyejiaka, A., … Doody, G. A. (2015). Diagnostic change 10 years after a first episode of psychosis. *Psychological Medicine*, *45*(13), 2757–2769. http://doi.org/10.1017/S0033291715000720

Hickling, F. W. (1995). The incidence of first-contact schizophrenia in Jamaica. *British Journal of Psychiatry*, *167*(December 1992), 193–196.

Hickling, F. W., McKenzie, K., Mullen, R., & R, M. (1999). A Jamaican psychiatrist evaluates diagnoses at a London psychiatric hospital. *British Journal of Psychiatry*, *175*, 283–285.

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. http://doi.org/10.1002/sim.1186

Hjorthøj, C., Stürup, A. E., McGrath, J. J., & Nordentoft, M. (2017). Years of potential life lost and life expectancy in schizophrenia: a systematic review and meta-analysis. *The Lancet Psychiatry*. http://doi.org/10.1016/S2215-0366(17)30078-0

Hoang, U., Stewart, R., & Goldacre, M. J. (2011). Mortality after hospital discharge for people with schizophrenia or bipolar disorder: retrospective study of linked English hospital episode statistics, 1999-2006. *BMJ* , *343*, d5422.

Hogerzeil, S. J., van Hemert, A. M., Rosendaal, F. R., Susser, E., & Hoek, H. W. (2014). Direct comparison of first-contact versus longitudinal register-based case finding in the same population: early evidence that the incidence of schizophrenia may be three times higher than commonly reported. *Psychological Medicine*, *44*(16), 3481–90. http://doi.org/10.1017/S003329171400083X

Hollander, A.-C., Dal, H., Lewis, G., Magnusson, C., Kirkbride, J. B., & Dalman, C. (2016). Refugee migration and risk of schizophrenia and other non-affective psychoses: cohort study of 1.3 million people in Sweden. *BMJ*, *352*.

Hollingshead, A. B., & Redlich, F. C. (1958). *Social Class and Mental Illness: A Community Study*. New York: Wiley.

Horwitz, A. V. (2017). An Overview of Sociological Perspectives on the Definitions, Causes, and Responses to Mental Health and Illness. In T. L. Scheid & T. N. Brown (Eds.), *A Handbook for the Study of Mental Health* (2nd ed., pp. 6–19). Cambridge: Cambridge University Press.

Howes, O. D., & Kapur, S. (2009). The dopamine hypothesis of schizophrenia: Version III - The final common pathway. *Schizophrenia Bulletin*, *35*(3), 549–562. http://doi.org/10.1093/schbul/sbp006

Howes, O. D., & Kapur, S. (2014). A neurobiological hypothesis for the classification of schizophrenia: Type a (hyperdopaminergic) and type b (normodopaminergic). *British Journal of Psychiatry*, *205*(1), 1–3. http://doi.org/10.1192/bjp.bp.113.138578

Howes, O. D., McCutcheon, R., Owen, M. J., Murray, R. M., Huttunen, M., Aalto, S., … Myin-Germeys, I. (2017). The role of genes, stress, and dopamine in the development of schizophrenia. *Biological Psychiatry*, *81*(1), 9–20. http://doi.org/10.1016/j.biopsych.2016.07.014

Howes, O. D., & Murray, R. M. (2014). Schizophrenia: An integrated sociodevelopmental-cognitive model. *The Lancet*, *383*(9929), 1677–1687. http://doi.org/10.1016/S0140-6736(13)62036-X

Hutchinson, J., & Smith, A. D. (1996). *Ethnicity*. Oxford: Oxford University Press.

Hutnik, N. (1992). *Ethnic Minority Identity: A Social Psychological Perspective*. Oxford: Clarendon Press.

Illumina. (2017). HumanCoreExcome-24. San Diego.

Ingleby, D. (2006). *European Research on Migration and Health*.

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., … Wang, P. (2010). Research Domain Criteria (RDoC): toward a new classification framework for research on mental disorders. *American Journal of Psychiatry*, *167*(7), 748–751. http://doi.org/10.1176/appi.ajp.2010.09091379

Institut national de la statistique et des études économiques. (2015). Populations française, étrangère et immigrée en France depuis 2006 - Insee Focus - 38. Retrieved August 8, 2017, from https://www.insee.fr/fr/statistiques/1410693

Institut National de la statistique et des études économiques. (2014). Recent immigrants to France - Increasingly European immigration - Insee Première - 1524. Retrieved August 8, 2017, from https://www.insee.fr/en/statistiques/1283872

Institut national de la statistuque et des études économiques. (2016). IMG1B - Population immigrée par sexe, âge et pays de naissance en 2013−France−Étrangers - Immigrés en 2013 | Insee. Retrieved August 8, 2017, from https://www.insee.fr/fr/statistiques/2020942?sommaire=2106113&geo=FRANCE-1

Instituto Brasileiro de Geografia e Estatística. (2017). 2010 Popolation Census. Retrieved January 4, 2017, from http://www.ibge.gov.br/english/estatistica/populacao/censo2010/

Instituto Brasileiro de Geographica e Estatistica. (2011). *Censo Demographafico 2010. Caraceristicas da populacao e dos domicilios. Resultatados do universo.* Rio de Janeiro .

Instituto Nacional de Estadística. (2015). *Population Figures at 1 January 2015. Migrations Statistics 2014*.

Instituto Nacional de Estadística. (2017). *Cifras de Población 1-1-2017, Estadística Migraciones 2016* (Vol. 4495654405).

Ioannidis, J. P., Ntzani, E. E., & Trikalinos, T. A. (2004). "Racial" differences in genetic effects for complex diseases. *Nature Genetics*, *36*, 1312–1318. http://doi.org/10.1038/ng1474

Istat. (2016a). International and internal migration. Retrieved August 8, 2017, from https://www.istat.it/en/archive/193774

Istat. (2016b). Non-EU citizens: presence, new inflows and acquisition of citizenship. Retrieved August 8, 2017, from https://www.istat.it/en/archive/190679

Istat. (2017). National demographic balance. Retrieved August 8, 2017, from https://www.istat.it/en/archive/201143

Jablensky, A., Sartorius, N., Ernberg, G., Anker, M., Korten, A., Cooper, J. E., … Bertelsen, A. (1992). *Schizophrenia: manifestations, incidence and course in different cultures. A World Health Organization ten-country study. Psychological medicine. Monograph supplement* (Vol. 20). http://doi.org/10.1017/S0264180100000904

Jackson, D., Kirkbride, J., Croudace, T., Morgan, C., Boydell, J., Errazuriz, A., … Jones, P. B. (2013). Meta-analytic approaches to determine gender differences in the age-incidence characteristics of schizophrenia and related psychoses. *International Journal of Methods in Psychiatric Research*, *22*(1), 36–45. http://doi.org/10.1002/mpr.1376

Janssens, A. C. J., Ioannidis, J. P., van Duijn, C. M., Little, J., Khouriy, M. J., & GRIPS Group. (2011). Strengthening the reporting of genetic risk prediction studies: the GRIPS statement. *Genome Medicine*, *3*(16).

Jennissen, R. (2011). *De Nederlandse migratiekaart*. Den Haag.

Johnson, V. E. (2013). Revised standards for statistical evidence. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(48), 19313–7. http://doi.org/10.1073/pnas.1313476110

Jones, H. J., Stergiakouli, E., Tansey, K. E., Hubbard, L., Heron, J., Cannon, M., … Zammit, S. (2016). Phenotypic Manifestation of Genetic Risk for Schizophrenia During Adolescence in the General Population. *JAMA Psychiatry*, *73*(3), 221. http://doi.org/10.1001/jamapsychiatry.2015.3058

Jones, P. B., & Bhugra, D. (2010). Migration and Mental Health. In R. Bhattacharya, S. Cross, & D. Bhugra (Eds.), *Clinical Topics in Cultural Psychiatry* (pp. 15–26). London: RCPscyh Publications.

Jones, P., Rodgers, B., Murray, R., & Marmot, M. (1994). Child developmental risk factors for adult schizophrenia in the British 1946 birth cohort. *The Lancet*, *344*(8934), 1398–1402. http://doi.org/10.1016/S0140-6736(94)90569-X

Jordanova, V., Crawford, M. J., McManus, S., Bebbington, P., & Brugha, T. (2015). Religious discrimination and common mental disorders in England: a nationally representative population-based study. *Social Psychiatry and Psychiatric Epidemiology*, *50*(11), 1723–1729. http://doi.org/10.1007/s00127-015-1110-6

Karakayali, N. (2009). Social distance and affective orientations. *Sociological Forum*, *24*(3), 538–562. http://doi.org/10.1111/j.1573-7861.2009.01119.x

Karlsen, S., & Nazroo, J. Y. (2002). Relation between racial discrimination, social class, and health among ethnic minority groups. *American Journal of Public Health*, *92*(4), 624–631.

Karlsen, S., Nazroo, J. Y., McKenzie, K., Bhui, K., & Weich, S. (2005). Racism, psychosis and common mental disorder among ethnic minority groups in England. *Psychological Medicine*, *35*, 1795–1803. http://doi.org/10.1017/S0033291705005830

Kelleher, I., Keeley, H., Corcoran, P., Ramsay, H., Wasserman, C., Carli, V., … Cannon, M. (2013). Childhood trauma and psychosis in a prospective cohort study: cause, effect, and directionality. *American Journal of Psychiatry*, *170*(7), 734–741. http://doi.org/10.1176/appi.ajp.2012.12091169

Kelly, M., & Nazroo, J. (2008). Ethnicity and Health. In G. Scambler (Ed.), *Sociology as Applied to Medicine* (8th ed., pp. 159–175). Croydon: Saunders Elsevier.

Kennedy, N., Boydell, J., Kalidindi, S., Fearon, P., Jones, P. B., Van Os, J., & Murray, R. M. (2005). Gender differences in incidence and age at onset of mania and bipolar disorder over a 35-year period in Camberwell, England. *American Journal of Psychiatry*, *162*(2), 257–262. http://doi.org/10.1176/appi.ajp.162.2.257

Kennedy, N., Everitt, B., Boydell, J., Van Os, J., Jones, P. B., & Murray, R. M. (2005). Incidence and distribution of first-episode mania by age: results from a 35-year study. *Psychological Medicine*, *35*(6), 855–863. http://doi.org/10.1017/S0033291704003307

Khandaker, G. M., Barnett, J. H., White, I. R., & Jones, P. B. (2011). A quantitative meta-analysis of population-based studies of premorbid intelligence and schizophrenia. *Schizophrenia Research*, *132*, 220–227. http://doi.org/10.1016/j.schres.2011.06.017

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. *JAMA Psychiatry*, *71*(10), 1121–1128. http://doi.org/10.1001/jamapsychiatry.2014.1332

Khandaker, G. M., Stochl, J., Zammit, S., Lewis, G., & Jones, P. B. (2014). A population-based longitudinal study of childhood neurodevelopmental disorders, IQ and subsequent risk of psychotic experiences in adolescence. *Psychological Medicine*, *44*, 3229–3238. http://doi.org/10.1017/S0033291714000750

Khandaker, G. M., Zammit, S., Lewis, G., & Jones, P. B. (2014). A population-based study of atopic disorders and inflammatory markers in childhood before psychotic experiences in adolescence. *Schizophrenia Research*, *152*(1), 139–145. http://doi.org/10.1016/j.schres.2013.09.021

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. *Psychological Medicine*, *43*(2), 239–57. http://doi.org/10.1017/S0033291712000736

King, M., Semlyen, J., Tai, S. S., Killaspy, H., Osborn, D., Popelyuk, D., & Nazareth, I. (2008). A systematic review of mental disorder, suicide, and deliberate self harm in lesbian, gay and bisexual people. *BMC Psychiatry*, *8*(1), 70. http://doi.org/10.1186/1471-244X-8-70

Kinney, D. K., Teixeira, P., Hsu, D., Napoleon, S. C., Crowley, D. J., Miller, A., … Huang, E. (2009). Relation of schizophrenia prevalence to latitude, climate, fish consumption, infant mortality, and skin color: a role for prenatal vitamin D deficiency and infections? *Schizophrenia Bulletin*, *35*(5), 582–595. http://doi.org/10.1093/schbul/sbp023

Kirkbride, J. B. (2009). Impact of contextual-environmental mechanisms on the incidence of schizophrenia and other psychoses. In W. Gattaz & G. Busatto (Eds.), *Advances in Schizophrenia Research.* New York: Springer.

Kirkbride, J. B. (2017). Psymaptic — Psychiatric Mapping Translated into Innovations for Care.

Retrieved August 9, 2017, from http://www.psymaptic.org/

Kirkbride, J. B., Barker, D., Cowden, F., Stamps, R., Yang, M., Jones, P. B., & Coid, J. W. (2008). Psychoses, ethnicity and socio-economic status. *British Journal of Psychiatry*, *193*(1), 18–24. http://doi.org/10.1192/bjp.bp.107.041566

Kirkbride, J. B., Boydell, J., Ploubidis, G. B., Morgan, C., Dazzan, P., McKenzie, K., … Jones, P. B. (2008). Testing the association between the incidence of schizophrenia and social capital in an urban area. *Psychological Medicine*, *38*(8), 1083–94. http://doi.org/10.1017/S0033291707002085

Kirkbride, J. B., Croudace, T., Brewin, J., Donoghue, K., Mason, P., Glazebrook, C., … Jones, P. B. (2009). Is the incidence of psychotic disorder in decline? Epidemiological evidence from two decades of research. *International Journal of Epidemiology*, *38*(5), 1255–1264. http://doi.org/10.1093/ije/dyn168

Kirkbride, J. B., Errazuriz, A., Croudace, T. J., Morgan, C., Jackson, D., Boydell, J., … Jackson, D. (2012). Incidence of schizophrenia and other psychoses in England, 1950–2009: a systematic review and meta-analyses. *PLoS ONE*, *7*(3), e31660. http://doi.org/10.1371/journal.pone.0031660

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. *Social Psychiatry and Psychiatric Epidemiology*, *42*(6), 438–445. http://doi.org/10.1007/s00127-007-0193-0

Kirkbride, J. B., Fearon, P., Morgan, C., Dazzan, P., Morgan, K., Tarrant, J., … Jones, P. B. (2006). Heterogeneity in incidence rates of schizophrenia and other psychotic syndromes: findings from the 3-center AeSOP study. *Archives of General Psychiatry*, *63*(3), 250–258. http://doi.org/10.1001/archpsyc.63.3.250

Kirkbride, J. B., Hameed, Y., Ankireddypalli, G., Ioannidis, K., Crane, C. M., Nasir, M., … Jones, P. B. (2017). The epidemiology of first-episode psychosis in early ontervention in psychosis services: findings From the Social Epidemiology of Psychoses in East Anglia [SEPEA] study. *American Journal of Psychiatry*, *174*(2), 143–153. http://doi.org/10.1176/appi.ajp.2016.16010103

Kirkbride, J. B., Hameed, Y., Ioannidis, K., Ankireddypalli, G., Crane, C. M., Nasir, M., … Jones, P. B. (2017). Ethnic minority status, age-at-immigration and psychosis risk in rural environments: evidence From the SEPEA Study. *Schizophrenia Bulletin*. http://doi.org/10.1093/schbul/sbx010

Kirkbride, J. B., Jackson, D., Perez, J., Fowler, D., Winton, F., Coid, J. W., … Jones, P. B. (2013). A population-level prediction tool for the incidence of first-episode psychosis: translational epidemiology based on cross-sectional data. *BMJ Open*, *3*(2), e001998. http://doi.org/10.1136/bmjopen-2012-001998

Kirkbride, J. B., Jones, P. B., Ullrich, S., & Coid, J. W. (2014). Social deprivation, inequality, and the neighborhood-level incidence of psychotic syndromes in East London. *Schizophrenia Bulletin*, *40*(1), 169–180. http://doi.org/10.1093/schbul/sbs151

Kirkbride, J. B., Morgan, C., Fearon, P., Dazzan, P., Murray, R. M., & Jones, P. B. (2007). Neighbourhood-level effects on psychoses: re-examining the role of context. *Psychological Medicine*, *37*(10), 1413–25. http://doi.org/10.1017/S0033291707000499

Kirkbride, J. B., Stubbins, C., & Jones, P. B. (2012). Psychosis incidence through the prism of early intervention services. *British Journal of Psychiatry*, *200*(2), 156–157. http://doi.org/10.1192/bjp.bp.111.094896

Kirkwood, B. R., & Sterne, J. A. C. (2003). *Essential Medical Statistics*. Oxford: Blackwell Publishing Ltd.

Kirmayer, L. J., Narasiah, L., Munoz, M., Rashid, M., Ryder, A. G., Guzder, J., … Pottie, K. (2011). Common mental health problems in immigrants and refugees: General approach in primary care. *CMAJ*, *183*(12), 959–967. http://doi.org/10.1503/cmaj.090292

Koch, R. (1884). Die Aetiologie der Tuberculose. *Mitt. Kaiser. Gesundh.*, *2*(1).

Koch, R. (1890). Ueber bakteriologische Forschung. In *Ver. X. Int. Med. Congr.* (p. 35). Berlin, Germany.

Koczan, Z. (2016). Does identity matter? *Migration Studies*, *4*(1), 116–145. http://doi.org/10.1093/migration/mnv021

Kohn, M. L., & Clausen, J. A. (1955). Social isolation and schizophrenia. *American Sociological Review*, *20*(3), 265. http://doi.org/10.2307/2087384

Kraepelin, E. (1899). *Psychiatrie* (6th editio). Leipzig: Barth.

Kroon, J. S., Wohlfarth, T. D., Dieleman, J., Sutterland, A. L., Storosum, J. G., Denys, D., … Sturkenboom, M. C. (2013). Incidence rates and risk factors of bipolar disorder in the general population: A population-based cohort study. *Bipolar Disorders*, *15*(3), 306–313. http://doi.org/10.1111/bdi.12058

Lappin, J. M., Heslin, M., Jones, P. B., Doody, G. A., Reininghaus, U. A., Demjaha, A., … Morgan, C. (2016). Outcomes following first-episode psychosis – Why we should intervene early in all ages, not only in youth. *Australian & New Zealand Journal of Psychiatry*, *50*(5011), 1055–1063. http://doi.org/10.1177/0004867416673454

Lasalvia, A., Bonetto, C., Tosato, S., Zanatta, G., Cristofalo, D., Salazzari, D., … Santi, M. (2014). First-contact incidence of psychosis in north-eastern Italy: Influence of age, gender, immigration and socioeconomic deprivation. *British Journal of Psychiatry*, *205*(2), 127–134. http://doi.org/10.1192/bjp.bp.113.134445

Laursen, T. M., Wahlbeck, K., Hällgren, J., Westman, J., Ösby, U., Alinaghizadeh, H., … Nordentoft, M. (2013). Life expectancy and death by diseases of the circulatory system in patients with bipolar disorder or schizophrenia in the Nordic countries. *PLoS ONE*, *8*(6), e67133. http://doi.org/10.1371/journal.pone.0067133

Lawlor, D. A., Harbord, R. M., Sterne, J. A. C., Timpson, N., & Smith, G. D. (2008). Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Statistics in Medicine*, *37*, 1133–1163. http://doi.org/10.1002/sim

Leão, T. S., Sundquist, J., Frank, G., Johansson, L.-M., Johansson, S.-E., & Sundquist, K. (2006). Incidence of schizophrenia or other psychoses in first- and second-generation immigrants: a national cohort study. *The Journal of Nervous and Mental Disease*, *194*(1), 27–33. http://doi.org/10.1097/01.nmd.0000195312.81334.81

Leão, T., Sundquist, J., Johansson, L. M., Johansson, S.-E., & Sundquist, K. (2005). Incidence of mental disorders in second-generation immigrants in Sweden: a four-year cohort study. *Ethnicity & Health*, *10*(3), 243–56. http://doi.org/10.1080/13557850500096878

Lederbogen, F., Kirsch, P., Haddad, L., Streit, F., Tost, H., Schuch, P., … Meyer-Lindenberg, A. (2011). City living and urban upbringing affect neural social stress processing in humans. *Nature*, *474*(7352), 498–501. http://doi.org/10.1038/nature10190

Leucht, S., Cipriani, A., Spineli, L., Mavridis, D., Örey, D., Richter, F., … Davis, J. M. (2013). Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *The Lancet*, *382*(9896), 951–962. http://doi.org/10.1016/S0140-6736(13)60733-3

Lewis, G., Croft-Jeffreys, C., & David, A. (1990). Are British psychiatrists racist? *British Journal of Psychiatry*, *157*, 410–415.

Lindert, J., Von Ehrenstein, O. S., Priebe, S., Mielck, A., & Brähler, E. (2009). Depression and anxiety in labor migrants and refugees – A systematic review and meta-analysis. http://doi.org/10.1016/j.socscimed.2009.04.032

Lindström, M., Sundquist, J., & Östergren, P.-O. (2001). Ethnic differences in self reported health in Malmö in southern Sweden. *J Epidemiol Community Health*, *55*, 97–103.

Link, B. G., & Phelan, J. (1995). Social conditions as fundamental causes of disease. *J Health Soc Behav*, *29*, 80–94. http://doi.org/10.2307/2626958

Little, R., & Rubin, D. (2002). *Statistical analyses with Missing Data* (2nd Editio). New York: Wiley.

Lloyd, T., Kennedy, N., Fearon, P., Kirkbride, J., Mallett, R., Leff, J., … Jones, P. B. (2005). Incidence of bipolar affective disorder in three UK cities: Results from the ÆSOP study. *British Journal of Psychiatry*, *186*(FEB.), 126–131. http://doi.org/10.1192/bjp.186.2.126

Mackie, J. (1965). Causes and Conditions. *American Philosophical Quarterly*, *2*(4), 245–364.

Magdalene College. (2017). Magdalene in the Twenty-first Century. Retrieved from http://www.magd.cam.ac.uk/magdalene-in-the-twenty-first-century/

Mahmmood, M., & Fisher, H. (2006). 0548 The incidence of first episode psychosis in inner London: Findings from the Lambeth Early Onset (LEO) service. *Schizophrenia Research*, *86*, S66–S67.

Malinowska, J. K. (2016). Cultural neuroscience and the category of race: the case of the other-race effect. *Synthese*, *193*(12), 3865–3887. http://doi.org/10.1007/s11229-016-1108-y

Mallett, R. (1997). *MRC Sociodemographic Schedule.* London: Institute of Psychiatry.

Manrique-Garcia, E., Zammit, S., Dalman, C., Hemmingsson, T., Andreasson, S., & Allebeck, P. (2012). Cannabis, schizophrenia and other non-affective psychoses: 35 years of follow-up of a population-based cohort. *Psychological Medicine*, *42*(6), 1321–1328. http://doi.org/10.1017/S0033291711002078

March, D., Hatch, S. L., Morgan, C., Kirkbride, J. B., Bresnahan, M., Fearon, P., & Susser, E. (2008). Psychosis and Place. *Epidemiologic Reviews*, *30*(1), 84–100. http://doi.org/10.1093/epirev/mxn006

Marmot, M. (2015). *The Health Gap*. London: Bloomsbury Publishing.

Marmot, M., Allen, J., Bell, R., Bloomer, E., & Goldblatt, P. (2012). WHO European review of social determinants of health and the health divide. *The Lancet*, *380*, 1011–1029. http://doi.org/10.1016/S0140-6736(12)61228-8

Marwaha, S., & Johnson, S. (2004). Schizophrenia and Employment: A Review. *Soc Psychiatry Psychiatr Epidemiol*, *39*, 337–349. http://doi.org/10.1007/s00127-004-0762-4

Matheson, S. L., Shepherd, A. M., Pinchbeck, R. M., Laurens, K. R., & Carr, V. J. (2012). Childhood adversity in schizophrenia: a systematic meta-analysis. *Psychological Medicine*, (July 2016), 1–13. http://doi.org/10.1017/S0033291712000785

Maylaih, E., Weyerer, S., & Hafner, H. (1989). Spatial concentration of the incidence of treated psychiatric disorders in Mannheim. *Acta Psychiatrica Scandinavica*, *80*(6), 650–656. http://doi.org/10.1111/j.1600-0447.1989.tb03039.x

McCutcheon, R., Bloomfield, M. A. P., Dahoun, T., Terbeck, S., Mehta, M., & Howes, O. D. (2017). Increased amygdala reactivity in ethnic minority individuals and its relationship with the social environment. *Psychological Medicine*.

McEwen, B. S. (2012). Brain on stress: How the social environment gets under the skin. *Proceedings of the National Academy of Sciences*, *109*(Supplement_2), 17180–17185. http://doi.org/10.1073/pnas.1121254109

McGrath, J. J., Mortensen, P. B., Visscher, P. M., Wray, N. R., Sullivan, P., & Hultman, C. (2013). Where GWAS and Epidemiology Meet: Opportunities for the Simultaneous Study of Genetic and Environmental Risk Factors in Schizophrenia. *Schizophrenia Bulletin*, *39*(5), 955–959. http://doi.org/10.1093/schbul/sbt108

McGrath, J., Saha, S., Chant, D., & Welham, J. (2008). Schizophrenia: a concise overview of incidence, prevalence, and mortality. *Epidemiologic Reviews*, *30*(1), 67–76. http://doi.org/10.1093/epirev/mxn001

McGrath, J., Saha, S., Welham, J., El Saadi, O., MacCauley, C., & Chant, D. (2004). A systematic review of the incidence of schizophrenia: the distribution of rates and the influence of sex, urbanicity, migrant status and methodology. *BMC Medicine*, *2*, 13. http://doi.org/10.1186/1741-7015-2-13

McKenzie, K., Bhui, K., Nanchahal, K., & Blizard, B. (2008). Suicide rates in people of South Asian origin in England and Wales: 1993–2003. *The British Journal of Psychiatry*, *193*(5).

Meader, N., & Chan, M. K. Y. (2017). Sexual orientation and suicidal behaviour in young people. *The British Journal of Psychiatry*, *211*(2).

Menezes, P., & Scazufca, M. (2007). Incidence of first-contact psychosis in Sao Paulo, Brazil. *The British Journal of Psychiatry*, *191*, s102-106.

Mental Health Foundation. (2017). Stigma and discrimination. Retrieved November 29, 2017, from https://www.mentalhealth.org.uk/a-to-z/s/stigma-and-discrimination

Migration Watch UK. (2014). A summary history of immigration to Britain.

Missinne, S., & Bracke, P. (2012). Depressive symptoms among immigrants and ethnic minorities: a population based study in 23 European countries. *Soc Psychiatry Psychiatrc Epidemiol*, *47*, 97–109. http://doi.org/10.1007/s00127-010-0321-0

Mizrahi, R., Addington, J., Rusjan, P. M., Suridjan, I., Ng, A., Boileau, I., … Wilson, A. A. (2012). Increased stress-induced dopamine release in psychosis. *Biological Psychiatry*, *71*(6), 561–567. http://doi.org/10.1016/j.biopsych.2011.10.009

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, T. P. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine*, *6*(7), e1000097. http://doi.org/10.1371/journal.pmed.1000097

Mohr, S., Borras, L., Nolan, J., Gillieron, C., Brandt, P.-Y., Eytan, A., … Huguelet, P. (2012). Spirituality and religion in outpatients with schizophrenia: a multi-site comparative study of Switzerland, Canada, and the United States. *The International Journal of Psychiatry in Medicine*, *44*(1), 29–52. http://doi.org/10.2190/PM.44.1.c

Montalvo, J. G., & Reynal-Querol, M. (2010). Ethnic polarization and the duration of civil wars. *Econ Gov*, *11*, 123–143. http://doi.org/10.1007/s10101-010-0077-8

Moore, T. H. M., Zammit, S., Lingford-Hughes, A., Barnes, T. R. E., Jones, P. B., Burke, M., & Lewis, G. (2007). Cannabis use and risk of psychotic or affective mental health outcomes: a systematic

review. *Lancet*, *370*, 319–328. http://doi.org/10.1016/S0140-6736(07)61162-3

Morgan, C., Charalambides, M., Hutchinson, G., & Murray, R. M. (2010). Migration, ethnicity, and psychosis: Toward a sociodevelopmental model. *Schizophrenia Bulletin*, *36*(4), 655–664. http://doi.org/10.1093/schbul/sbq051

Morgan, C., Dazzan, P., Morgan, K., Jones, P., Harrison, G., Leff, J., … Fearon, P. (2006). First episode psychosis and ethnicity: initial findings from the AESOP study. *World Psychiatry : Official Journal of the World Psychiatric Association (WPA)*, *5*(1), 40–6.

Morgan, C., John, S., Esan, O., Hibben, M., Patel, V., Weiss, H., … Cohen, A. (2017). The incidence of psychoses in diverse settings, INTREPID (2): a feasibility study in India, Nigeria, and Trinidad. http://doi.org/10.1017/S0033291716000441

Morgan, C., Kirkbride, J., Hutchinson, G., Craig, T., Morgan, K., Dazzan, P., … Fearon, P. (2008). Cumulative social disadvantage, ethnicity and first-episode psychosis: a case-control study. *Psychological Medicine*, *38*(12), 1701. http://doi.org/10.1017/S0033291708004534

Morgan, C., Mallett, R., Hutchinson, G., & Leff, J. (2004). Negative pathways to psychiatric care and ethnicity: the bridge between social science and psychiatry. *Social Science & Medicine*, *58*, 739–752. http://doi.org/10.1016/S0277-9536(03)00233-8

Morrell, S., Taylor, R., Slaytor, E., & Ford, P. (1999). Urban and rural suicide differentials in migrants and the Australian-born, New South Wales, Australia 1985-1994. *Social Science & Medicine (1982)*, *49*(1), 81–91.

Mühleisen, T. W., Leber, M., Schulze, T. G., Strohmaier, J., Degenhardt, F., Treutlein, J., … Cichon, S. (2014). Genome-wide association study reveals two new risk loci for bipolar disorder. *Nature Communications*, *5*. http://doi.org/10.1038/ncomms4339

Mulè, A., Sideli, L., Capuccio, V., Fearon, P., Ferraro, L., Kirkbride, J. B., … Murray, R. M. (2017). Low incidence of psychosis in Italy: confirmation from the first epidemiological study in Sicily. *Social Psychiatry and Psychiatric Epidemiology*, *52*(2), 155–162. http://doi.org/10.1007/s00127-016-1322-4

Murray, R. M., Sham, P., van Os, J., Zanelli, J., Cannon, M., & McDonald, C. (2004). A developmental model for similarities and dissimilarities between schizophrenia and bipolar disorder. *Schizophrenia Research*, *71*(2–3), 405–416. http://doi.org/10.1016/J.SCHRES.2004.03.002

Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, *78*(3), 691–692.

National Collaborating Centre for Methods and Tools. (2011). AMSTAR: assessing methodological quality of systematic reviews. Hamilton, ON: McMaster University.

Nazroo, J., & Karlsen, S. (2003). Patterns of identity amoung ethnic minority people: Diversity and commonality. *Ethnic and Racial Studies*, *26*(5), 902–930.

Ngozi Adichie, C. (2013). *Americanah*. New York: Albert A. Knopf.

NHS Blood and Transplant. (2017). Black donors | Blood Donation. Retrieved November 29, 2017, from https://www.blood.co.uk/why-give-blood/the-need-for-blood/why-we-need-more-black-donors/

NHS Choices. (2016). Sickle cell disease. Retrieved November 29, 2017, from https://www.nhs.uk/conditions/sickle-cell-disease/

NHS Digital. (2017). *Mental Health Bulletin: 2016-2017 Annual Report*.

NICE. (2014). *Psychosis and schizophrenia in adults: prevention and management. CG178.* NICE.

Nicolaas, H., & Sprangers, A. (2007). Buitenlandse migratie in Nederland 1795–2006: de invloed op de bevolkings- samenstelling. *Bevolkingstrends*, 32–46.

Nixon, N. L., & Doody, G. A. (2005). Official psychiatric morbidity and the incidence of schizophrenia 1881-1994. *Psychological Medicine*, *35*(8), 1145–1153. http://doi.org/http://dx.doi.org/10.1017/S0033291705004939

Nurnberger, J. I., Blehar, M. C., Kaufmann, C. A., York-Cooler, C., Simpson, S. G., Harkavy-Friedman, J., … Reich, T. (1994). Diagnostic interview for genetic studies. Rationale, unique features, and training. NIMH Genetics Initiative. *Archives of General Psychiatry*, *51*(11), 849–864.

Ødegaard, Ø. (1932). Emigration and insanity. *Acta Psychiatrica Neurologica Scandinavica (Suppl.)*, *4*, 1–206.

OECD. (2017). *Education at a Glance*.

Office for National Statistics. (2012). International Migrants in England and Wales - Office for National Statistics. Retrieved August 8, 2017, from https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/articles/internationalmigrantsinenglandandwales/2012-12-11

Office for National Statistics. (2013a). 2011 Census analysis: Immigration Patterns of Non-UK Born Populations in England and Wales in 2011. Retrieved August 8, 2017, from https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/articles/immigrationpatternsofnonukbornpopulationsinenglandandwalesin2011/2013-12-17

Office for National Statistics. (2013b). Ethnic group by sex by age. Retrieved August 11, 2016, from https://www.nomisweb.co.uk/census/2011/dc2101ew

Oh, H., Abe, J., Negi, N., & Devylder, J. (2015). Immigration and psychotic experiences in the United States: Another example of the epidemiological paradox? *Psychiatry Research*, *229*, 784–790. http://doi.org/10.1016/j.psychres.2015.08.002

Omer, S., Kirkbride, J. B., Pringle, D. G., Russell, V., O'Callaghan, E., & Waddington, J. L. (2014). Neighbourhood-level socio-environmental factors and incidence of first episode psychosis by place at onset in rural Ireland: The Cavan-Monaghan First Episode Psychosis Study [CAMFEPS]. *Schizophrenia Research*, *152*(1), 152–157. http://doi.org/10.1016/j.schres.2013.11.019

ONS Digital. (2015). International migration: a recent history | Visual.ONS. Retrieved August 8, 2017, from http://visual.ons.gov.uk/uk-perspectives-a-recent-history-of-international-migration/

Owen, M. J., Sawa, A., & Mortensen, P. B. (2016). Schizophrenia. *Lancet (London, England)*, *388*(10039), 86–97. http://doi.org/10.1016/S0140-6736(15)01121-6

Oxford University Press. (2016). English Oxford Living Dictionaries. Retrieved from https://en.oxforddictionaries.com/definition/minority

Parascandola, M., & Weed, D. L. (2001). Causation in epidemiology. *J Epidemiol Community Health*, *55*, 905–912.

Pearl, J. (1984). *Heuristics: Intelligent search strategies for computer problem solving.* Reading, MA.: Addison-Wesledy Pub. Co. Inc.

Pedersen, C. B., & Cantor-Graae, E. (2012). Age at migration and risk of schizophrenia among immigrants in Denmark: A 25-year incidence study. *American Journal of Psychiatry*, *169*(10), 1117–1118. http://doi.org/10.1176/appi.ajp.2012.12050614

Pelayo-Terán, J. M., Pérez-Iglesias, R., Ramírez-Bonilla, M. L., González-Blanch, C., Martínez-García, O., Pardo-García, G., … Crespo-Facorro, B. (2008). Epidemiological factors associated with treated incidence of first-episode non-affective psychosis in Cantabria: Insights from the Clinical Programme on Early Phases of Psychosis. *Early Intervention in Psychiatry*, *2*(3), 178–187. http://doi.org/10.1111/j.1751-7893.2008.00074.x

Perälä, J., Suvisaari, J., Saarni, S. I., Kuoppasalmi, K., Isometsä, E., Pirkola, S., … Lönnqvist, J. (2007). Lifetime Prevalence of Psychotic and Bipolar I Disorders in a General Population. *Archives of General Psychiatry*, *64*(1), 19. http://doi.org/10.1001/archpsyc.64.1.19

Porta, M. (Ed.). (2014). *A dictionary of epidemiology* (6th editio). Oxford: Oxford University Press.

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. *Nature Genetics*, *38*(8), 904–909. http://doi.org/10.1038/ng1847

Proctor, S. E., Mitford, E., & Paxton, R. (2004). First episode psychosis: A novel methodology reveals higher than expected incidence; a reality-based population profile in Northumberland, UK. *Journal of Evaluation in Clinical Practice*, *10*(4), 539–547. http://doi.org/10.1111/j.1365-2753.2003.00474.x

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., … Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–75. http://doi.org/10.1086/519795

Putnam, R. D. (2007). Diversity and community in the twenty-first century the 2006 Johan Skytte prize lecture. *Scandinavian Political Studies*, *20*(137–174), 207–217. http://doi.org/10.1080/10511482.2008.9521631

Putnam, R. D. (2015). *Our Kids: The American Dream in Crisis*. New York: Simon & Schuster.

Reay, R., Mitford, E., McCabe, K., Paxton, R., & Turkington, D. (2010). Incidence and diagnostic diversity in first-episode psychosis. *Acta Psychiatrica Scandinavica*, *121*(4), 315–319. http://doi.org/10.1111/j.1600-0447.2009.01505.x

Rechel, B., Mladovsky, P., Ingleby, D., Mackenbach, J. P., & Mckee, M. (2013). Health in Europe 5 Migration and health in an increasingly diverse Europe. *The Lancet*, *381*, 1235–1245. http://doi.org/10.1016/S0140-6736(12)62086-8

Reichenberg, A., Caspi, A., Harrington, H., Houts, R., Keefe, R. S. E., Murray, R. M., … Moffitt, T. E. (2010). Static and dynamic cognitive deficits in childhood preceding adult schizophrenia: a 30-year study. *The American Journal of Psychiatry*, *167*(2), 160–9. http://doi.org/10.1176/appi.ajp.2009.09040574

Reininghaus, U. A., Morgan, C., Simpson, J., Dazzan, P., Morgan, K., Doody, G. A., … Craig, T. K. J. (2008). Unemployment, social isolation, achievement–expectation mismatch and psychosis: findings from the AESOP Study. *Soc Psychiatry Psychiatr Epidemiol*, *43*, 743–751. http://doi.org/10.1007/s00127-008-0359-4

Revier, C. J., Reininghaus, U., Dutta, R., Fearon, P., Murray, R. M., Doody, G. A., … Jones, P. B. (2015). Ten-Year outcomes of first-episode psychoses in the MRC ÆSOP-10 study. *The Journal of Nervous and Mental Disease*, *203*(5), 379–386.

http://doi.org/10.1097/NMD.0000000000000295

Richardson, L., Hameed, Y., Perez, J., Jones, P. B., & Kirkbride, J. B. (2017). Association of Environment With the Risk of Developing Psychotic Disorders in Rural Populations: Findings from the Social Epidemiology of Psychoses in East Anglia Study. *JAMA Psychiatry*. http://doi.org/10.1001/jamapsychiatry.2017.3582

Rittel, H., & Webber, M. (1973). Dilemmas in a General Theory of Planning. *Policy Sciences*, *4*(2), 155–169.

Rothman, K. J. (1976). Causes. *American Journal of Epidemiology*, *104*(6), 587–592.

Rothman, K. J. (2005). Causation and causal inference in epidemiology. *American Journal of Public Health*, *1*(S1). http://doi.org/10.2105/AJPH.2004.059204

Russo, F., & Williamson, J. (2007). Interpreting causality in the health sciences. *International Studies in the Philosophy of Science*, *21*(2), 157–170. http://doi.org/10.1080/02698590701498084

Rylander-Rudqvist, T., Håkansson, N., Tybring, G., & Wolk, A. (2006). Quality and quantity of saliva DNA obtained from the self-administrated Oragene method—a pilot study on the cohort of Swedish men. *Cancer Epidemiology and Prevention Biomarkers*, *15*(9).

Saha, S., Chant, D. C., Welham, J. L., & McGrath, J. J. (2006). The incidence and prevalence of schizophrenia varies with latitude. *Acta Psychiatrica Scandinavica*, *114*(1), 36–39. http://doi.org/10.1111/j.1600-0447.2005.00742.x

Saha, S., Chant, D., & McGrath, J. (2007). A systematic review of mortality in schizophrenia: is the differential mortality gap worsening over time? *Archives of General Psychiatry*, *64*(10), 1123–31. http://doi.org/10.1001/archpsyc.64.10.1123

Saha, S., Chant, D., Welham, J., McGrath, J., & Lapsley, H. (2005). A systematic review of the prevalence of schizophrenia. *PLoS Medicine*, *2*(5), e141. http://doi.org/10.1371/journal.pmed.0020141

Sapolsky, R. M. (2005). The influence of social hierarchy on primate health. *Science (New York, N.Y.)*, *308*(5722), 648–52. http://doi.org/10.1126/science.1106477

Sariah, A. E., Outwater, A. H., & Malima, K. I. (2014). Risk and protective factors for relapse among Individuals with Schizophrenia: A Qualitative Study in Dar es Salaam, Tanzania. *BMC Psychiatry*, *14*(1), 240. http://doi.org/10.1186/s12888-014-0240-9

Sariaslan, A., Larsson, H., D'Onofrio, B., Langstrom, N., Fazel, S., & Lichtenstein, P. (2015). Does population density and neighborhood deprivation predict schizophrenia? A nationwide Swedish family-based study of 2.4 million individuals. *Schizophrenia Bulletin*, *41*(2), 494–502. http://doi.org/10.1093/schbul/sbu105

Scambler, G. (Ed.). (2010). *Sociology as applied to medicine* (6th ed.). London: W.B. Saunders.

Schenker, M. B. (2008). Work-related injuries among immigrants: a growing global health disparity. *Occup Environ Med*, *65*(11), 717–718.

Schenker, M. B. (2010). A Global Perspective of Migration and Occupational Health. *American Journal of Industrial Medicine*. http://doi.org/10.1002/ajim.20834

Schizophrenia Working Group of the Psychiatric Genomic Consortium. (2015). Biological Insights From 108 Schizophrenia-Associated Genetic Loci. *Nature*, *511*(7510), 421–427. http://doi.org/10.1038/nature13595.Biological

Schneider, K. (1957). [Primary & secondary symptoms in schizophrenia]. *Fortschr Neurol Psychiatr*, *25*(9), 487–90.

Schofield, P., Ashworth, M., & Jones, R. (2011). Ethnic isolation and psychosis: re-examining the ethnic density effect. *Psychological Medicine*, *41*(6), 1263–9. http://doi.org/10.1017/S0033291710001649

Selten, J.-P., & Cantor-Graae, E. (2005). Social defeat: risk factor for schizophrenia? *British Journal of Psychiatry*, *187*, 101–102.

Selten, J.-P., Cantor-Graae, E., Slaets, J., & Kahn, R. S. (2002). Ødegaard's selection hypothesis revisited: schizophrenia in Surinamese immigrants to the Netherlands. *American Journal of Psychiatry*, *159*(4), 669–671. http://doi.org/10.1176/appi.ajp.159.4.669

Selten, J.-P., Zeyl, C., Dwark Asing, R., Lumsden, V., Kahn, R. S., & van Harten, P. N. (2005). First-contact incidence of schizophrenia in Surinam. *British Journal of Psychiatry*, *186*, 74–75.

Selten, J. P., Van Der Ven, E., Rutten, B. P. F., & Cantor-Graae, E. (2013). The social defeat hypothesis of schizophrenia: An update. *Schizophrenia Bulletin*, *39*(6), 1180–1186. http://doi.org/10.1093/schbul/sbt134

Selten, J. P., van Os, J., & Nolen, W. A. (2003). First admissions for mood disorders in immigrants to the Netherlands. *Social Psychiatry and Psychiatric Epidemiology*, *38*(10), 547–550. http://doi.org/10.1007/s00127-003-0673-9

Sen, A. (2006). *Identity and Violence: The Illusion of Destiny*. New York: W.W. Norton & Company.

Senior, P. A., & Bhopal, R. (1994). Ethnicity as a variable in epidemiological research. *BMJ*, *309*(6950).

Shenkar, O. (2001). Cultural distance revisited: towards a more rigorous conceptualization and measurement of cultural differences. *Journal of International Business Studies*, *32*(3), 519–535. http://doi.org/10.1057/palgrave.jibs.8490982

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: identity salience and shifts in quantitative performance. *Psychological Science*, *10*(1), 80–83.

Siegrist, J., & Marmot, M. (2004). Health inequalities and the psychosocial environment - Two scientific challenges. *Social Science and Medicine*, *58*(8), 1463–1473. http://doi.org/10.1016/S0277-9536(03)00349-6

Singh, S. P., Burns, tom, Amin, S., Jones, P. B., & Harrison, G. (2004). Acute and transient psychotic disorders: precursors, epidemiology, course and outcome. *British Journal of Psychiatry*, *185*, 452–459.

Singh, S. P., Cooper, J. E., Fisher, H. L., Tarrant, C. J., Lloyd, T., Banjo, J., … Jones, P. (2005). Determining the chronology and components of psychosis onset: The Nottingham Onset Schedule (NOS). *Schizophrenia Research*, *80*(1), 117–130. http://doi.org/10.1016/j.schres.2005.04.018

Singh, S., Wright, C., Joyce, E., Barnes, T., & Burns, T. R. E. (2003). Developing early intervention services in the NHS: A survey to guide workforce and training needs. *Psychiatric Bulletin*, *27*(7), 254–258. http://doi.org/10.1192/pb.27.7.254

Singh Chawla, D. (2017). Big names in statistics want to shake up much-maligned P value. Retrieved August 3, 2017, from https://www.nature.com/articles/d41586-017-02190-5

Singleton, N., Bumpstead, R., O'Brien, M., Lee, A., & Meltzer, H. (2003). Psychiatric morbidity among

adults living in private households, 2000. *International Review of Psychiatry*, *15*(1–2), 65–73. http://doi.org/10.1080/0954026021000045967

Sipos, A., Rasmussen, F., Harrison, G., Tynelius, P., Lewis, G., Leon, D. A., & Gunnell, D. (2004). Paternal age and schizophrenia: a population based cohort study. *BMJ*, *329*(7474), 1070. http://doi.org/10.1136/bmj.38243.672396.55

Smith Nielsen, S., & Krasnik, A. (2010). Poorer self-perceived health among migrants and ethnic minorities versus the majority population in Europe: a systematic review. *Int J Public Health*, *55*, 357–371. http://doi.org/10.1007/s00038-010-0145-4

Sniekers, S., Stringer, S., Watanabe, K., Jansen, P. R., Coleman, J. R. I., Krapohl, E., … Posthuma, D. (2017). Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nature Genetics*, *49*(7), 1107–1112. http://doi.org/10.1038/ng.3869

Solmi, F., Hayes, J. F., Lewis, G., & Kirkbride, J. B. (2017). Curiosity killed the cat: no evidence of an association between cat ownership and psychotic symptoms at ages 13 and 18 years in a UK general population cohort. *Psychological Medicine*, 1–9. http://doi.org/10.1017/S0033291717000125

StataCorp. (2013). Stata Statistical Software: Release 13. College Station, TX: StataCorp LP.

Steichen, T. (2009). Tests for publication bias in meta-analysis. In J. A. C. Sterne (Ed.), *Meta-Analysis in Stata: An Updated Collection from the Stata Journal.* (pp. 151–161). College Station, TX: Stata Press.

Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., … Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*, *338*.

Sterne, J., & Harbord, R. (2009). Funnel plots in meta-analyses. In J. A. C. Sterne (Ed.), *Meta-Analysis in Stata: An Updated Collection from the Stata Journal* (pp. 109–123). College Station, TX: Stata Press.

Sudhinaraset, M., Wigglesworth, C., & Takeuchi, D. T. (2016). Social and cultural contexts of alcohol use: influences in a social–ecological framework. *Alcohol Research: Current Reviews*, *38*(1), 35–45.

Sullivan, G. M., & Feinn, R. (2012). Using effect size—or why the P value is not enough. *Journal of Graduate Medical Education*, *4*(3), 279–282. http://doi.org/10.4300/JGME-D-12-00156.1

Sullivan, H., & Skelcher, C. (2002). *Working across Boundaries - Collaboration in Public Services* (Palgrave M). Basingstoke.

Sullivan, P. F., Daly, M. J., & O'Donovan, M. (2012). Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nature Reviews Genetics*, *13*(8), 537–551. http://doi.org/10.1038/nrg3240

Sullivan, S. A., Hollen, L., Wren, Y., Thompson, A. D., Lewis, G., & Zammit, S. (2016). A longitudinal investigation of childhood communication ability and adolescent psychotic experiences in a community sample. *Schizophrenia Research*, *173*(1–2), 54–61. http://doi.org/10.1016/j.schres.2016.03.005

Sundquist, K., Frank, G., & Sundquist, J. (2004). Urbanisation and incidence of psychosis and depression: Follow-up study of 4.4 million women and men in Sweden. *British Journal of*

Psychiatry, *184*(APR.), 293–298. http://doi.org/10.1192/bjp.184.4.293

Susser, E., & Martínez-Alés, G. (2017). Putting Psychosis Into Sociocultural Context: An International Study in 17 Locations. *JAMA Psychiatry*. http://doi.org/10.1001/jamapsychiatry.2017.3541

Susser, E. S., & Lin, S. P. (1992). Schizophrenia after prenatal exposure to the Dutch hunger winter of 1944-1945. *Archives of General Psychiatry*, *49*(12), 983. http://doi.org/10.1001/archpsyc.1992.01820120071010

Susser, M. (1973). *Causal Thinking in the Health Sciences: Concepts and Strategies of Epidemiology.* New York: Oxford University Press.

Sutterland, A. L., Dieleman, J., Storosum, J. G., Voordouw, B. A. C., Kroon, J., Veldhuis, J., … Sturkenboom, M. C. J. M. (2013). Annual incidence rate of schizophrenia and schizophrenia spectrum disorders in a longitudinal population-based cohort study. *Social Psychiatry and Psychiatric Epidemiology*, *48*(9), 1357–1365. http://doi.org/10.1007/s00127-013-0651-9

Szöke, A., Charpeaud, T., Galliot, A.-M., Vilain, J., Richard, J.-R., Leboyer, M., … Schürhoff, F. (2014). Rural-urban variation in incidence of psychosis in France: a prospective epidemiologic study in two contrasted catchment areas. *BMC Psychiatry*, *14*, 78. http://doi.org/10.1186/1471-244X-14-78

Tabak, N. T., & de Mamani, A. W. (2014). Religion's effect on mental health in schizophrenia. *Clinical Schizophrenia & Related Psychoses*, *8*(2), 91–100. http://doi.org/10.3371/CSRP.TUWE.021513

Tarricone, I., Mimmi, S., Paparelli, A., Rossi, E., Mori, E., Panigada, S., … Berardi, D. (2012). First-episode psychosis at the West Bologna Community Mental Health Centre: results of an 8-year prospective study. *Psychological Medicine*, *42*(11), 2255–64. http://doi.org/10.1017/S0033291712000335

The Cochrane Collaboration. (2008). *Cochrane Handbook for Systematic Reviews of Interventions Cochrane Handbook for Systematic Reviews of Interventions Cochrane Book Series*. (J. P. Higgins & S. Green, Eds.). Chichester, West Sussex: John Wiley & Sons, Ltd.

The International Schizophrenia Consortium. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, *460*(August). http://doi.org/10.1038/nature08185

The World Bank. (2017). DataBank: Poverty and Equity. Retrieved from http://databank.worldbank.org/data/reports.aspx?source=poverty-and-equity-database

Thoits, P. A. (2017). Sociological Approaches to Mental Illness. In T. L. Scheid & T. N. Brown (Eds.), *A Handbook for the Study of Mental Health* (2nd ed., pp. 106–124). Cambridge: Cambridge University Press.

Thorup, A., Waltoft, B. L., Pedersen, C. P., Mortensen, P. B., & Nordentoft, M. (2007). Young males have a higher risk of developing schizophrenia : a Danish register study. *Psychological Medicine*, *37*, 479–484. http://doi.org/10.1017/S0033291707009944

Tiihonen, J., Lönnqvist, J., Wahlbeck, K., Klaukka, T., Niskanen, L., Tanskanen, A., & Haukka, J. (2009). 11-year follow-up of mortality in patients with schizophrenia: a population-based cohort study (FIN11 study). *The Lancet*, *374*(9690), 620–627. http://doi.org/10.1016/S0140-6736(09)60742-X

Tishkoff, S. A., Reed, F. A., Friedlaender, F. R., Ehret, C., Ranciaro, A., Froment, A., … Williams, S. M. (2009). The genetic structure and history of Africans and African Americans. *Science* , *324*(5930), 1035–44. http://doi.org/10.1126/science.1172257

Torrey, E. F. (1987). Prevalence studies in schizophrenia. *The British Journal of Psychiatry* , *150*, 598–

608.

Tortelli, A., Morgan, C., Szoke, A., Nascimento, A., Skurnik, N., Monduit De Caussade, E., … Murray, R. M. (2014). Different rates of first admissions for psychosis in migrant groups in Paris. *Soc Psychiatry Psychiatr Epidemiol.*, *49*(7), 1109–1109. http://doi.org/10.1007/s00127-013-0795-7

Townsend, M., & Warsi, S. (2017, June 11). Britain needs to reset relations with its Muslims, insists Warsi. *The Guardian*.

Turner, M. A., Finch, P. J. C., McKechanie, A. G., Kiernan, M. D., Hawksley, O. J., Wadhwani, S., … Neal, L. A. (2006). Psychosis in the British army: a 2-year follow-up study. *Military Medicine*, *171*(12), 1215–9.

Turola, M. C., Comellini, G., Galuppi, A., Nanni, M. G., Carantoni, E., & Scapoli, C. (2012). Schizophrenia in real life: courses, symptoms and functioning in an Italian population. *International Journal of Mental Health Systems*, *6*(1), 22. http://doi.org/10.1186/1752-4458-6-22

UNICEF. (n.d.). Migration Profiles - Brazil.

University of Cambridge. (2017). Undergraduate Admissions Statistics.

van der Ven, E., Dalman, C., Wicks, S., Allebeck, P., Magnusson, C., van Os, J., & Selten, J. P. (2014). Testing Ødegaard's selective migration hypothesis: a longitudinal cohort study of risk factors for non-affective psychotic disorders among prospective emigrants. *Psychological Medicine*, 1–8. http://doi.org/10.1017/S0033291714001780

van der Werf, M., Hanssen, M., Köhler, S., Verkaaik, M., Verhey, F. R., van Winkel, R., … Allardyce, J. (2014). Systematic review and collaborative recalculation of 133,693 incident cases of schizophrenia. *Psychological Medicine*, *44*(1), 9–16. http://doi.org/10.1017/S0033291712002796

van Os, J., Kenis, G., & Rutten, B. P. (2010). The environment and schizophrenia. *Nature*, *468*(7321), 203–212. http://doi.org/10.1038/nature09563

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. *Psychological Medicine*, *39*(2), 179. http://doi.org/10.1017/S0033291708003814

van Os, J., Rutten, B. P., Myin-Germeys, I., Delespaul, P., Viechtbauer, W., Van Zelst, C., … Mirjanic, T. (2014). Identifying gene-environment interactions in schizophrenia: Contemporary challenges for integrated, large-scale investigations. *Schizophrenia Bulletin*, *40*(4), 729–736. http://doi.org/10.1093/schbul/sbu069

Van Winkel, R., Stefanis, N. C., & Myin-Germeys, I. (2008). Psychosocial Stress and Psychosis. A Review of the Neurobiological Mechanisms and the Evidence for Gene-Stress Interaction. *Schizophrenia Bulletin*, *34*(6), 1095–1105. http://doi.org/10.1093/schbul/sbn101

Vancampfort, D., Knapen, J., Probst, M., Scheewe, T., Remans, S., & De Hert, M. (2012). A systematic review of correlates of physical activity in patients with schizophrenia. *Acta Psychiatrica Scandinavica*, *125*(5), 352–362. http://doi.org/10.1111/j.1600-0447.2011.01814.x

Vandenheede, H., Deboosere, P., Stirbu, I., Agyemang, C. O., Harding, S., Knud, @bullet, … Mackenbach, J. P. (2012). Migrant mortality from diabetes mellitus across Europe: the importance of socio-economic change. *Eur J Epidemiol*, *27*, 109–117.

http://doi.org/10.1007/s10654-011-9638-6

Varese, F., Smeets, F., Drukker, M., Lieverse, R., Lataster, T., Viechtbauer, W., … Bentall, R. P. (2012). Childhood adversities increase the risk of psychosis: a meta-analysis of patient-control, prospective- and cross-sectional cohort studies. *Schizophrenia Bulletin*, *38*(4), 661–71. http://doi.org/10.1093/schbul/sbs050

Vassos, E., Di Forti, M., Coleman, J., Iyegbe, C., Prata, D., Euesden, J., … Breen, G. (2017). An Examination of Polygenic Score Risk Prediction in Individuals With First-Episode Psychosis. *Biological Psychiatry*, *81*(6), 470–477. http://doi.org/10.1016/j.biopsych.2016.06.028

Vassos, E., Pedersen, C. B., Murray, R. M., Collier, D. A., & Lewis, C. M. (2012). Meta-analysis of the association of urbanicity with schizophrenia. *Schizophrenia Bulletin*, *38*(6), 1118–1123. http://doi.org/10.1093/schbul/sbs096

Veen, N. D., Selten, J. P., Schols, D., Laan, W., Hoek, H. W., Van Der Tweel, I., & Kahn, R. S. (2004). Diagnostic stability in a Dutch psychosis incidence cohort. *British Journal of Psychiatry*, *185*(DEC.), 460–464. http://doi.org/10.1192/bjp.185.6.460

Veling, W., Hoek, H. W., & Mackenbach, J. P. (2008). Perceived discrimination and the risk of schizophrenia in ethnic minorities. *Social Psychiatry and Psychiatric Epidemiology*, *43*(12), 953–959. http://doi.org/10.1007/s00127-008-0381-6

Veling, W., Hoek, H. W., Selten, J.-P., & Susser, E. (2011). Age at migration and future risk of psychotic disorders among immigrants in the Netherlands: a 7 -year incidence study. *Am J Psychiatry*, *168*(12), 1278–1285.

Veling, W., Selten, J. P., Susser, E., Laan, W., Mackenbach, J. P., & Hoek, H. W. (2007). Discrimination and the incidence of psychotic disorders among ethnic minorities in The Netherlands. *International Journal of Epidemiology*, *36*(4), 761–768. http://doi.org/10.1093/ije/dym085

Veling, W., Selten, J. P., Veen, N., Laan, W., Blom, J. D., & Hoek, H. W. (2006). Incidence of schizophrenia among ethnic minorities in the Netherlands: A four-year first-contact study. *Schizophrenia Research*, *86*(1–3), 189–193. http://doi.org/10.1016/j.schres.2006.06.010

Veling, W., Susser, E., Selten, J.-P., & Hoek, H. W. (2014). Social disorganization of neighborhoods and incidence of psychotic disorders: a 7-year first-contact incidence study. *Psychological Medicine*, *45*(9), 1789–1798. http://doi.org/10.1017/S0033291714002682

Veling, W., Susser, E., Selten, J.-P., & Hoek, H. W. (2015). Social disorganization of neighborhoods and incidence of psychotic disorders: a 7-year first-contact incidence study. *Psychological Medicine*, *45*, 1789–1798. http://doi.org/10.1017/S0033291714002682

Veling, W., Susser, E., Van Os, J., Mackenbach, J. P., Selten, J. P., & Hoek, H. W. (2008). Ethnic density of neighborhoods and incidence of psychotic disorders among immigrants. *American Journal of Psychiatry*, *165*(1), 66–73. http://doi.org/10.1176/appi.ajp.2007.07030423

Ventriglio, A., Gupta, S., & Bhugra, D. (2016). Why do we need a social psychiatry? *The British Journal of Psychiatry : The Journal of Mental Science*, *209*(1), 1–2. http://doi.org/10.1192/bjp.bp.115.175349

Waraich, P., Goldner, E. M., Somers, J. M., & Hsu, L. (2004). Prevalence and Incidence Studies of Mood Disorders: A Systematic Review of the Literature. *Can J Psychiatry*, *49*(2), 124–138.

Weiner, M., Warren, L., & Fiedorowicz, J. G. (2011). Cardiovascular morbidity and mortality in bipolar disorder. *Annals of Clinical Psychiatry* , *23*(1), 40–7.

Werner, S., Malaspina, D., & Rabinowitz, J. (2007). Socioeconomic status at birth is associated with risk of schizophrenia: Population-based multilevel study. *Schizophrenia Bulletin*, *33*(6), 1373–1378. http://doi.org/10.1093/schbul/sbm032

West, J., & Graham, J. L. (2004). A Linguistic-based Measure of Cultural Distance and Its Relationship to Managerial Values. *Management International Review; Third Quarter*, *44*(443), 239–260.

Wheaton, B., & Montazer, S. (2017). Stressors, Stress and Distress. In T. L. Scheid & T. N. Brown (Eds.), *A Handbook for the Study of Mental Health* (2nd ed., pp. 171–199). Cambridge: Cambridge University Press.

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*, *30*(4), 377–399. http://doi.org/10.1002/sim.4067

Whiteford, H. A., Ferrari, A. J., Degenhardt, L., Feigin, V., & Vos, T. (2015). The global burden of mental, neurological and substance use disorders: an analysis from the Global Burden of Disease Study 2010. *PloS One*, *10*(2), e0116820. http://doi.org/10.1371/journal.pone.0116820

Wichmann, S., Holman, E. W., Bakker, D., & Brown, C. H. (2010). Evaluating linguistic distance measures. *Physica A: Statistical Mechanics and Its Applications*, *389*(17), 3632–3639. http://doi.org/10.1016/j.physa.2010.05.011

Wiking, E., Johansson, S.-E., Sundquist, J., & Wiking, E. (2004). Ethnicity, acculturation, and self reported health. A population based study among immigrants from Poland, Turkey, and Iran in Sweden. *J Epidemiol Community Health*, *58*, 574–582. http://doi.org/10.1136/jech.2003.011387

Williams, D. R., Costa, M., & Leavell, J. P. (2017). Race and Mental Health: Patterns and Challenges. In T. L. Scheid & T. N. Brown (Eds.), *A Handbook for the Study of Mental Health* (2nd ed., pp. 268–290). Cambridge: Cambridge University Press.

Wilson, W. J. (2010). *More than Just Race: Being Black and Poor in the Inner City*. New York: W.W. Norton & Company.

World Health Organization. (1992a). *International Statistical Classification of Diseases and Health Related Problems* (Tenth Revi).

World Health Organization. (1992b). *Schedules for Clinical Assessment in Neuropsychiatry*. Geneva, Switzerland.

World Values Survey Association. (2015). World Values Survey Wave 6 2010-2014 Official Aggegrate v.20150418. Retrieved from http://www.worldvaluessurvey.org/WVSDocumentationWV6.jsp

World Values Survey Association. (2017). Fieldwork and Sampling. Retrieved from http://www.worldvaluessurvey.org/WVSContents.jsp

Worrall, J. (2007). Why there's no cause to randomize. *Brit. J. Phil. Sci*, *58*, 451–488. http://doi.org/10.1093/bjps/axm024

Zammit, S., Allebeck, P., Dalman, C., Lundberg, I., Hemmingson, T., Owen, M. J., & Lewis, G. (2003). Paternal age and risk for schizophrenia. *The British Journal of Psychiatry*, *183*(5).

Zammit, S., Lewis, G., Rasbash, J., Dalman, C., Gustafsson, J.-E., & Allebeck, P. (2010). Individuals, schools, and neighborhood: a multilevel longitudinal study of variation in incidence of psychotic disorders. *Arch.Gen.Psychiatry*, *67*(9), 914–922.

Zandi, T., Havenaar, J. M., Smits, M., Limburg-Okken, A. G., van Es, H., Cahn, W., … van den Brink, W.

(2010). First contact incidence of psychotic disorders among native Dutch and Moroccan immigrants in the Netherlands: Influence of diagnostic bias. *Schizophrenia Research*, *119*(1–3), 27–33. http://doi.org/10.1016/j.schres.2010.02.1059

Zubin, J., & Spring, B. (1977). Vulnerability - a new view of schizophrenia. *J. Abnorm. Psychol.*, *86*, 103–26.

Appendices

Appendix 2A: Protocol as submitted to PROSPERO (CRD42015019276)

Version 1.3, 11/04/2015

A systematic review and meta-analysis of the incidence of psychotic disorders in Brazil, France, Italy, the Netherlands and Spain.

Hannah E Jongsma[8]*, Craig Morgan[9], Peter B Jones[6] and James B Kirkbride[10],

Contributions

One author (HEJ) will formulate the exact search strategy, and run the literature search. All authors will contribute to the selection process by scanning titles for exclusion of true negatives and identification of abstracts. One author (HEJ) will apply predefined inclusion criteria to all identified abstracts, cross-checking with the other authors where necessary. Two authors (HEJ and JBK) will independently carry out quality assessment on all studies approved for inclusion, differences will be solved by agreement. One author (HEJ) will be responsible for data collection, extraction and synthesis. This process will be monitored on a regular basis by a senior member of the research team with experience in systematic reviewing (CM, PBJ or JBK).

Support

No external funding was received for this project. JBK is funded by the Wellcome Trust, and HEJ is jointly funded by the EU-GEI project and the NIHR. PBJ is supported by the Wellcome Trust and the NIHR CLAHRC East of England

EU-GEI is the acronym of the project 'European network of National Schizophrenia Networks Studying Gene-Environment Interactions". The research leading to these results has received funding from the European Community's Seventh Framework Programme under grant agreement No.HEALTH-F2-2010-241909 (Project EU-GEI).

[8] Department of Psychiatry, University of Cambridge. Herchel Smith Building, Robinson Way, Cambridge, CB2 0SZ.
[9] Institute of Psychiatry, Psychology and Neuroscience, King's College London, East Wing, London, SE1 7EH
[10] Division of Psychiatry, University College London. Charles Bell House, 67-73 Riding House Street, London, W1W 7EJ
* Corresponding author: hej33@cam.ac.uk

1. Introduction

1.1 Rationale

Interpretation of Jablensky and colleagues' ten-country study led to the conclusion that incidence rates of schizophrenia were similar across countries (Jablensky 1992). However, the available methodology for epidemiological studies of this type has since been developed, with larger studies and the opportunity of meta-analysis to greatly enhance power. A substantial body of evidence has been accumulated to demonstrate the importance of the environment (McGrath 2004, Kirkbride 2006, van Os, Kenis et al. 2010, Bourque, van der Ven et al. 2011, Cheng, Kirkbride et al. 2011, Kirkbride, Errazuriz et al. 2012, Vassos, Pedersen et al. 2012, Bhavsar, Boydell et al. 2014). It is now accepted that the incidence of schizophrenia and other psychotic disorders varies across countries and settings.

However, the latest international meta-analysis on this subject dates from 2004, and includes only studies published up until 2002 (McGrath 2004). More recent meta-analyses have been carried out, for instance (Kirkbride, Errazuriz et al. 2012), but these have been commonly limited to a single country or special population group. In the last decade, interest in incidence studies has mushroomed, and it is therefore appropriate to synthesise research results across countries.

This systematic review and meta-analysis should be seen in the context of the EU-GEI study into gene-environment interactions in psychotic disorders (European Network of National Networks studying Gene-Environment Interactions in, van Os et al. 2014). A key aim of this study is to determine the incidence of psychotic disorders across thirteen centres in six countries. This systematic review and meta-analysis will inform the analysis of variation in incidence rates, will aid in assessing the findings' reliability and in contextualising these findings.

Participating countries and centres are as follows:

- Brazil: Sao Paulo, Ribeirão Preto
- England: Cambridge, London
- France: Creteil, Clermont-Ferrand
- Italy: Bologna, Palermo
- The Netherlands: Amsterdam, Gouda
- Spain: Barcelona, Oviedo, Valencia

A recent systematic review has already identified all relevant studies in England until 2009 (Kirkbride, Errazuriz et al. 2012). The identification of studies in the present systematic review will therefore exclude England. In order to aid comparability and reliability, the present study will follow the English methodology as closely as possible. The data from the English review will be made available for synthesis within the present review.

The chronological starting point of the systematic review and meta-analysis will follow the end-point of study identification in McGrath et al's international systematic review and meta-analysis (31 December 2001).

1.2 Objectives

- To assess the incidence of psychotic disorders in Brazil, France, Italy, the Netherlands and Spain;
- To provide an estimate of heterogeneity in incidence rates across these countries;
- To provide a comparison to the incidence in the UK as found in (Kirkbride, Errazuriz et al. 2012);
- To provide an informed background to the incidence rates derived from the EU-GEI study;
- To investigate variance by two exposure settings: urbanicity and latitude, and by study quality.

2. Methods

2.1 Inclusion criteria

Studies are eligible for inclusion if they meet the following criteria

a. Time period: published between 1 January 2002 and 31 December 2014
b. Extent: wholly or partially conducted in Brazil, France, Italy, the Netherlands or Spain
c. Scope: published and grey literature
d. Containing original data on the incidence of non-organic adult psychosis (16-64 years)
e. Were published in the English language. A separate list will be kept with identified abstracts published in Portuguese, French, Italian, Dutch or Spanish for which no English full-text article could be found but which appear to otherwise meet the inclusion criteria.

2.2 Information sources

This protocol is written after the electronic database search in PubMed and manual reference check of the yield from this database have been completed, but before searching other databases and contacting lead authors and centre coordinators

Our search was conducted in PubMed and PsycINFO. Reference lists of all articles included in the systematic review have been manually checked for further studies. All lead authors will be contacted with a list of studies for inclusion, and will be asked to identify any further studies, including those appearing in grey or unpublished literature. Collaborating partners from the EU-GEI centres in the various countries (see Section 1.1) will also be contacted and asked to identify further studies known to them which we may have missed from our initial database searches.

2.3 Search strategy

The search strategy was based on a search strategy previously used in England (Kirkbride, Errazuriz et al. 2012), based on Cochrane Systematic Reviewing guidelines. This strategy is known to be effective, and has been chosen to increase reliability in the systematic identification of studies between countries.

The search strategy that has been used was as follows (adapted to suit the various databases):

#1 schizo* .ti,ab

#2 (psychotic or psychosis or psychoses).ti,ab

#3 (bipolar disorder*_.ti,ab

#4 (delusion* disorder).ti,ab

#5 (((severe or serious or chronic) and mental and (illness* or disorder*)).ti,ab.

#6 SMI.ti,ab

#7 (mani* depressi*).ti,ab

#8 chronic psychosis

#9 exp psychosis

#10 schizoaffective disorder

#11 (#1 or #1 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10)

#12 (inciden* or epidemiolog*).ti,ab

#13 ((first* or 1st* or hospital*) and (episode* or contact* or admission* or admit*)).ti,ab

#14 (case and register*).ti,ab

#15 case control*.ti,ab

#16 (prospectiv* or population* or communit* or survey*).ti,ab

#17 (#12 or #13 or #14 or #15 or #16 or #17)

#18 (Brasil or Brazil)

#19 France

#20 (Italy or Sicily)

#21 (Spain)

#22 (Holland or the Netherlands)

#23 (#19 or #20 or #21 or #22)

#24 (Brasilian or Brazilian)

#25 (French)

#26 (Italian or Sicilian)

#27 (Dutch)

#28 (Spanish)

#29 (#24 or #25 or #26 or #27 or #28)

#30 (Sao Paulo or Rebeirão Preto)

#31 (Creteil or Clermont-Ferrand)

#32 (Bologno or Palermo)

#33 (Amsterdam or Gouda)

#34 (Barcelona or Oviedo or Valencia)

#35 (#30 or #31 or #32 or #33 or #34)

#36 (#23 or #29 or #35)

#37 (#11 and #17 and #36)

2.4 Data management

Citations will be managed in Endnote (version X7.3) and extracted data will be kept in an Excel spreadsheet. This will be stored on secure and shared server, so that all three authors have access and the data is backed up daily.

2.5 Selection process

All three authors assessed an equal share of records identified via the database search. Initially, they reviewed titles only with the objective of excluding true negatives. When there was uncertainty about potential for inclusion, the citation was included for abstract review. A ten per cent subsample of HEJ's yield was cross-checked by JBK to ensure consistency across reviewers.

Secondly, both PBJ and JBK sent HJ a list of citations potentially suitable for inclusion. HEJ reviewed all abstracts and scored them against inclusion criteria. Where no conclusion could be made, full text articles were searched. Uncertainties were resolved in agreement with JBK.

This process to date (PubMed only) is summarised in Appendix 1.

2.6 Data collection process

Data will be extracted from the eligible studies onto a standardised Excel spreadsheet. This spreadsheet is modelled on the systematic review carried out for the English Department of Health (Kirkbride, Errazuriz et al. 2012), and is available on request.

The spreadsheet includes study-level and rate-level variables, as well as meta-variables (see Section 2.7).

Where missing data is identified (no incidence rate given or could be calculated), attempts to contact the lead author will be made, and where possible they will be asked to provide any relevant data. Where this is applicable, the contact and outcome will be clearly indicated in the final report.

2.7 Data items

Data items will be divided into three types of variables: study-level variables, rate-level variables and meta-level variables. The former provide information about the design of the study, rate-level variables include information on estimates of incidence in each study, and meta-level variables are included to explore variation in rates by various additional factors such as latitude, study quality and urbanicity. They might not be explicitly measured in the original citation, but can be derived from it.

Where possible, data will be extracted for the following items:

Study level variables: authors, study title, publication source, year of publication, study type (incidence, special population), incidence type (first admission, first contact, first record/diagnosis, case register), geographical setting, study length (recruitment dates and duration), age range, diagnostic outcome studied, case finding methodology, usage of OPCRIT diagnosis, method of confirming diagnosis, diagnostic classification system used, source of denominator data, associated citations, and other relevant information.

Rate-level variables include numerical data such as incidence rates, size of the numerator and denominator population. Where available, this will be included stratified by age, sex, ethnicity or country of birth. The method of standardisation will also be noted.

Meta-variables included will be study quality (see Section 2.9 below), urbanicity and latitude. Both theoretical hypotheses and empirical evidence suggest a role for the latter two factors.

2.8 Outcomes and prioritisation

Outcomes

Our main outcome variables are the incidence rates, expressed per 100,000 person years at-risk, of the following (categories of) psychotic disorders: all psychotic disorders, non-affective psychosis, schizophrenia, affective psychosis, bipolar disorder, psychotic depression, substance induced disorders, symptoms, and an 'other' category . These can either be ascertained directly from the paper, or derived manually by dividing the number of cases by the denominator population. The former has preference, and will be used if both are given.

Papers will be categorised on the basis of which class(es) of psychotic disorder(s) investigated insofar as practical. When only very small numbers of citations are found for each category, some might have to be amalgamated. A rule of thumb is a minimum of five studies per class. The method of assessment and classification system will be recorded, and differences in incidence between these will be monitored carefully by narrative synthesis and visual interpretation of data. This will be closely monitored by PBJ, who has extensive clinical and research experience.

Prioritisation

Where more than one citation describing the same study has been identified, the following criteria will be applied in order to prioritise studies for inclusion:

a) Data relevant to the specific outcome and/or exposure under investigation
b) Data presented with a corresponding estimate of the standard error
c) Data most closely relating to the entry criteria for the study
d) Published data
e) Citations published in highest ranked journals

These criteria have been adopted from (Kirkbride, Errazuriz et al. 2012).

2.9 Risk of bias in individual studies

The risk of various types of bias will be assessed via the below criteria, which are a combination of study- and outcome level variables. A cumulative score for each study will be reported, and will be included as a variable in eventual meta-regression (see 2.10 below).

Variables for quality assessment/assessment of risk of bias:

a) Defined catchment area;
b) Accurate denominator data;
c) Population-based case finding;
d) Standardised research diagnosis used;
e) Blinding (of clinician) to demographic variables;
f) Inclusion criteria clearly listed;
g) Leakage study conducted.

The purpose of a) and b) is to assess the accurateness of the reported incidence rate, the purpose of c) and g) is to assess if it is likely that all cases have been found, d) and e) are used to assess diagnostic reliability and f) addresses study quality. These quality criteria have been developed by (Kirkbride, Errazuriz et al. 2012), and are based on epidemiological theory of best-practice.

Publication bias will be assessed by means of visual inspection of funnel plots and a formal Egger's test of publication bias will be conducted if the number of studies included in the meta-analysis exceeds ten (Kirkbride, Errazuriz et al. 2012).

2.10 Data synthesis

First, a narrative synthesis of results will be conducted in order to allow for the identification of broad themes. The exact form of this will depend on the characteristics of the study yield, but careful descriptive accounts of the data will be supplemented by visual representations. Examples can be found in (Kirkbride, Errazuriz et al. 2012).

Secondly, in order for any particular study to be included in the meta-analysis, a standard error of the incidence rate (or its log) will have to be reported or derived. Studies for which no standard error can be calculated will thus not be included in any meta-analyses, though the study would be retained in the systematic review.

To synthesise and quantitatively analyse data, at this stage it will be transformed to the natural logarithmic scale.

Thirdly, heterogeneity of the studies included in the meta-analysis will be tested using the Q-test, and quantified using I^2. A previous meta-analysis showed that heterogeneity was generally very large (Kirkbride, Errazuriz et al. 2012). If this proves to be the case for this meta-analysis, data will be analysed by visual interpretation of forest plots alone, without calculating a pooled estimate. A tentative cut-off point is heterogeneity of over 75%.

If heterogeneity is sufficiently low to meaningfully conduct a meta-analysis, the same method as Kirkbride et al used will be used: a univariate random effects meta-analysis as originally proposed by DerSimonian and Laird (Kirkbride, Errazuriz et al. 2012). A random-effects meta-analysis allows for the incorporation of the estimated in-between study variance in its weighting, which results in wider but more realistic confidence intervals. This is generally more conservative than a fixed-effects meta-analysis, and is more appropriate when heterogeneity is large (Kirkwood 2003).

Because of the anticipated high levels of heterogeneity, an investigation of its possible sources will be carried out via meta-regression (where possible). This technique can be used to examine associations between study characteristics and outcome. The variables to be included in this are degree of urbanicity, latitude and study quality, and the model would test if at least of these is related in a linear manner to incidence of psychotic disorders. This will be tested in a forward stepwise model, where the individually most significant variables are entered first, and the model is considered to have converged when the τ^2 and covariate effects are almost identical to the previous model (Kirkwood 2003).

2.11 Meta-biases

For the description of the assessment of publication bias, please see Section 2.9 above.

2.12 Cumulative evidence

In order to examine the strength of evidence in relation to the outcomes studied, the GRADE criteria will be applied to findings of this systematic review and meta-analysis.

Bibliography

- Bhavsar, V., J. Boydell, R. Murray and P. Power (2014). "Identifying aspects of neighbourhood deprivation associated with increased incidence of schizophrenia." Schizophr Res **156**(1): 115-121.
- Bourque, F., E. van der Ven and A. Malla (2011). "A meta-analysis of the risk for psychotic disorders among first- and second-generation immigrants." Psychol Med **41**(5): 897-910.
- Cheng, F., J. B. Kirkbride, B. R. Lennox, J. Perez, K. Masson, K. Lawrence, K. Hill, L. Feeley, M. Painter, G. K. Murray, O. Gallagher, E. T. Bullmore and P. B. Jones (2011). "Administrative incidence of psychosis assessed in an early intervention service in England: first epidemiological evidence from a diverse, rural and urban setting." Psychol Med **41**(5): 949-958.
- European Network of National Networks studying Gene-Environment Interactions in, S., J. van Os, B. P. Rutten, I. Myin-Germeys, P. Delespaul, W. Viechtbauer, C. van Zelst, R. Bruggeman, U. Reininghaus, C. Morgan, R. M. Murray, M. Di Forti, P. McGuire, L. R. Valmaggia, M. J. Kempton, C. Gayer-Anderson, K. Hubbard, S. Beards, S. A. Stilo, A. Onyejiaka, F. Bourque, G. Modinos, S. Tognin, M. Calem, M. C. O'Donovan, M. J. Owen, P. Holmans, N. Williams, N. Craddock, A. Richards, I. Humphreys, A. Meyer-Lindenberg, F. M. Leweke, H. Tost, C. Akdeniz, C. Rohleder, J. M. Bumb, E. Schwarz, K. Alptekin, A. Ucok, M. C. Saka, E. C. Atbasoglu, S. Guloksuz, G. Gumus-Akay, B. Cihan, H. Karadag, H. Soygur, E. S. Cankurtaran, S. Ulusoy, B. Akdede, T. Binbay, A. Ayer, H. Noyan, G. Karadayi, E. Akturan, H. Ulas, C. Arango, M. Parellada, M. Bernardo, J. Sanjuan, J. Bobes, M. Arrojo, J. L. Santos, P. Cuadrado, J. J. Rodriguez Solano, A. Carracedo, E. Garcia Bernardo, L. Roldan, G. Lopez, B. Cabrera, S. Cruz, E. M. Diaz Mesa, M. Pouso, E. Jimenez, T. Sanchez, M. Rapado, E. Gonzalez, C. Martinez, E. Sanchez, M. S. Olmeda, L. de Haan, E. Velthorst, M. van der Gaag, J. P. Selten, D. van Dam, E. van der Ven, F. van der Meer, E. Messchaert, T. Kraan, N. Burger, M. Leboyer, A. Szoke, F. Schurhoff, P. M. Llorca, S. Jamain, A. Tortelli, F. Frijda, J. Vilain, A. M. Galliot, G. Baudin, A. Ferchiou, J. R. Richard, E. Bulzacka, T. Charpeaud, A. M. Tronche, M. De Hert, R. van Winkel, J. Decoster, C. Derom, E. Thiery, N. C. Stefanis, G. Sachs, H. Aschauer, I. Lasser, B. Winklbaur, M. Schlogelhofer, A. Riecher-Rossler, S. Borgwardt, A. Walter, F. Harrisberger, R. Smieskova, C. Rapp, S. Ittig, F. Soguel-dit-Piquard, E. Studerus, J. Klosterkotter, S. Ruhrmann, J. Paruch, D. Julkowski, D. Hilboll, P. C. Sham, S. S. Cherny, E. Y. Chen, D. D. Campbell, M. Li, C. M. Romeo-Casabona, A. Emaldi Cirion, A. Urruela Mora, P. Jones, J. Kirkbride, M. Cannon, D. Rujescu, I. Tarricone, D. Berardi, E. Bonora, M. Seri, T. Marcacci, L. Chiri, F. Chierzi, V. Storbini, M. Braca, M. G. Minenna, I. Donegani, A. Fioritti, D. La Barbera, C. E. La Cascia, A. Mule, L. Sideli, R. Sartorio, L. Ferraro, G. Tripoli, F. Seminerio, A. M. Marinaro, P. McGorry, B. Nelson, G. P. Amminger, C. Pantelis, P. R. Menezes, C. M. Del-Ben, S. H. Gallo Tenan, R. Shuhama, M. Ruggeri, S. Tosato, A. Lasalvia, C. Bonetto, E. Ira, M. Nordentoft, M. O. Krebs, N. Barrantes-Vidal, P. Cristobal, T. R. Kwapil, E. Brietzke, R. A. Bressan, A. Gadelha, N. P. Maric, S. Andric, M. Mihaljevic and T. Mirjanic

(2014). "Identifying gene-environment interactions in schizophrenia: contemporary challenges for integrated, large-scale investigations." Schizophr Bull **40**(4): 729-736.

- Jablensky, A. S., N; Ernberg, G; Anker, M; Korten, A; Cooper, JE; Day R; Bertelsen, A (1992). "Schizophrenia: manifestations, incidence and course in different cultures. A Wold Health Organization ten-country study." Psychological Medicine. Monograph Supplement **20**: 1-97.
- Kirkbride, J. B., A. Errazuriz, T. J. Croudace, C. Morgan, D. Jackson, J. Boydell, R. M. Murray and P. B. Jones (2012). "Incidence of schizophrenia and other psychoses in England, 1950-2009: a systematic review and meta-analyses." PLoS One **7**(3): e31660.
- Kirkbride, J. F., P; Morgan, C; Dazzan, P; Morkan K; Tarrant, J; Lloyd, T; Holloway, J: Hutchinson, G; Leff, JP; Mallet, RM; Harrison, GL; Murray RM; Jones, PB (2006). "Heterogeneity in incidence rates of schizophrenia and other psychotic syndromes: findings from the 3-center AeSOP study." Arch Gen Psychiatry **63**(3): 250-528.
- Kirkwood, B. R. S., Jonathan A.C. (2003). Essential Medical Statistics. Oxford, whitewell Publishing Ltd.
- McGrath, J. S., S.; Welham, J.; El Saadi, O.; MacCauley, C.; Chant, D. (2004). "A systematic review of the incidence of schizophrenia: the distribution of rates and the influence of sex, urbanicity, migrant status and methodology. ." BMC Medicine: 22.
- van Os, J., G. Kenis and B. P. Rutten (2010). "The environment and schizophrenia." Nature **468** (7321): 203-212.
- Vassos, E., C. B. Pedersen, R. M. Murray, D. A. Collier and C. M. Lewis (2012). "Meta-analysis of the association of urbanicity with schizophrenia." Schizophr Bull **38**(6): 1118-1123.

Appendix 2B: Studies excluded after full-text review

Authors	Year	Title	Journal	Reason excluded
Tarricone I, Berardi D, Mule A, Tosato S, Fiorillo A, Morgan C	2015	Risk factors for psychosis in migrants: preliminary results from the EUGEI and PEP-ITA studies	Conference abstract	Other: EU-GEI data presented
Jansen K, Cardoso TA, Mondin TC, Matos MB, Souza LD, Pinheiro RT, Magalhães PV, Silva RA	2014	Stressful live events and mood disorders: a community sample	Cien Saude Colet	No clinically defined psychosis
Nivoli, AM; Murru A; Pachiarotti I; Valenti M; Rosa AR; Hidalgo D; Virdis V; Strejilevich S; Vieta E; Colom F	2014	Bipolar disorder in the elderly: a cohort study comparing older and younger patients	Acta Psychiatrica Scandinavica	No incidence data
Rolim-Neto ML, Alves Silva E, Teixeira Júnior AG, Sousa-Cartaxo JD, Rolim Lima NN, Nascimento VB, Vieira Dos Santos MD, Lima da Silva CG, Romero de Sousa SI, Silva Costa LD, Nascimento Neto PJ	2014	Bipolar disorder incidence between children and adolescents: a brief communication.	Journal of Affective disorders	No adult incidence data
Salvatore P, Baldessarini RJ, Khalsa HM, Vázquez G, Perez J, Faedda GL, Amore M, Maggini C, Tohen M.	2014	Antecedents of manic versus other first psychotic episodes in 263 bipolar I disorder patients	Acta Psychiatrica Scandinavica	No incidence data
Tarricone I Braca M, Allegri F, Barrasso G, Bellomo A, Berlincioni V, Carpiniello B, Ceregato A, Conforti Donati M, Defilippi S, Del Vecchio V, De Rosa C, Ferrannini L, Ferrari S, Furio MA, Gramaglia C, La Cascia C, Luciano M, Mulè A, Nardini M, Podavini F, Primavera D, Reggianini C, Rigatelli M, Todarello O, Turella E, Ventriglio A, Zeppegno P, Fiorillo A, Berardi D.	2014	First-episode psychosis and migration in Italy (PEP-Ita migration): a study in the Italian mental health services	BMC Psychiatry	No incidence data
Tarricone I, Boydell J, Panigada S, Allegri F, Marcacci T, Minenna MG, Kokona A, Triolo F, Storbini V, Michetti R, Morgan C, Di Forti M, Murray RM, Berardi D.	2014	The impact of substance use at psychosis onset on First Episode Psychosis course: results from 1 year follow-up study in Bologna.	Schizophrenia Research	No incidence data
Allegri F, Belvederi Murri M, Paparelli A, Marcacci T, Braca M, Menchetti M, Michetti R, Berardi D, Tarricone I.	2013	Current cannabis use and age of psychosis onset: a gender-mediated relationship? Results from an 8-year FEP incidence study in Bologna	Psychiatry Research	No incidence data
Bernardo M, Bioque M, Parellada M, Saiz Ruiz J, Cuesta MJ, Llerena A, Sanjuán J, Castro-Fornieles J, Arango C, Cabrera B; PEPs Group	2013	Assessing clinical and functional outcomes in a gene-environment interaction study in first episode of psychosis (PEPs).	Revista de Psiquiatría y Salud Mental	No incidence data
Carta, MG: Angermeyer, MC; Sancassiani, F; Tuligi, F; Pirastu, R: Risano, A; Pintus, E; Mellino, G; Pintus, M; Isanu, E; Moro, MF: Massida, D; Trincas, G; Bhugra, D;	2013	A follow-up on patients with severe mental disorders in Sardinia after two changes in regional policies: poor resources still correlate with poor outcomes.	BMC Psychiatry	No incidence data

Galderisi S, Mucci A, Bitter I, Libiger J, Bucci P, Fleischhacker WW, Kahn RS; Eufest Study Group	2013	Persistent negative symptoms in first episode patients with schizophrenia: results from the European First Episode Schizophrenia Trial	European Neuropsychopharmacology	No incidence data
Menezes PR	2013	Epidemiology of schizophrenia in Brazil.	Conference abstract	Other: data covered elsewhere
Mule A	2013	Lower incidence of schizophrenia in Italy	Conference abstract	Other: EU-GEI data presented
Murray RM, Morgan C, Dazzan P, Jones P, Doody G, Fearon P, Lappin J	2013	The social causes of the high rates of psychosis in Black people in Europe	Conference abstract	Other: no original data presented
Ochoa S, Huerta-Ramos E, Barajas A, Iniesta R, Dolz M, Baños I, Sánchez B, Carlson J, Foix A, Pelaez T, Coromina M, Pardo M; GENIPE group, Usall J.	2013	Cognitive profiles of three clusters of patients with a first-episode psychosis	Schizophrenia Research	No incidence data
Stouten, LH; Veling W; Van der Helm, M; Laan W; Van der Gaag, M	2013	Cognitive deficits and ethnicity: a cohort study of early psychosis patients in The Netherlands	Social Psychiatry and Psychiatric Epidemiology	No incidence data
Tosato S, Lasalvia A, Bonetto C, Mazzoncini R, Cristofalo D, De Santi K, Bertani M, Bissoli S, Lazzarotto L, Marrella G, Lamonaca D, Riolo R, Gardellin F, Urbani A, Tansella M, Ruggeri M; PICOS-VENETO Group.	2013	The impact of cannabis use on age of onset and clinical characteristics in first-episode psychotic patients. Data from the Psychosis Incident Cohort Outcome Study (PICOS).	Journal of Psychiatric Research	No incidence data
Veling, W.	2013	Ethnic minority position and risk for psychotic disorders	Current Opinion in Psychiatry	Other: no original data presented
Bertani M, Lasalvia A, Bonetto C, Tosato S, Cristofalo D, Bissoli S, De Santi K, Mazzoncini R, Lazzarotto L, Santi M, Sale A, Scalabrin D, Abate M, Tansella M, Rugger M	2012	The influence of gender on clinical and social characteristics of patients at psychosis onset: A report from the Psychosis Incident cohort Outcome Study (PICOS)	Psychological Medicine	No incidence data
Carrà G, Johnson S, Bebbington P, Angermeyer MC, Heider D, Brugha T, Azorin JM, Toumi M.	2012	The lifetime and past-year prevalence of dual diagnosis in people with schizophrenia across Europe: findings from the European Schizophrenia Cohort (EuroSC)	European Archives of Psychiatry and Clinical Neuroscience	No incidence data
Lasalvia A, Tosato S, Brambilla P, Bertani M, Bonetto C, Cristofalo D, Bissoli S, De Santi K, Lazzarotto L, Zanatta G, Marrella G, Mazzoncini R, Zanoni M, Garzotto N, Dolce C, Nicolau S, Ramon L, Perlini C, Rambaldelli G, Bellani M, Tansella M, Ruggeri M; PICOS-Veneto Group	2012	Psychosis Incident Cohort Outcome Study (PICOS). A multisite study of clinical, social and biological characteristics, patterns of care and predictors of outcome in first-episode psychosis. Background, methodology and overview of the patient sample	Epidemiology and Psychiatric Sciences	No incidence data
Lora A, Kohn R, Levav I, McBain R, Morris J, Saxena S	2012	Service availability and utilisation and treatment gap for schizophrenic disorders: a survey in 50 low-and middle-income countries	Bulletin of the World Health Organisation	No incidence data
Mule A, Sideli L, La Cascia C, Di Forti M, Murray RMG, La Barbera D	2012	Incidence of psychotic disorders in Palermo: preliminary data	Conference abstract	Other: EU-GEI data presented

Selten, JP; Laan, W; Kupka, R; Smeets, HM; van Os, J;	2012	Risk of psychiatric treatment for mood disorders and psychotic disorders among migrants and Dutch nationals in Utrecht, the Netherlands.	Social Psychiatry and Psychiatric Epidemiology	No incidence data
Bergink V, Lambregtse-van den Berg MP, Koorengevel MP, Kupka R, Kushner SA	2011	First-onset psychosis occurring in the postpartum period: a prospective cohort study	Journal of Clinical Psychiatry	No incidence data
Castro-Fornieles, J; Baeza, I; de la Serna, E; Gonzalez-Pinto, A; Paralleda, M; Grael, M; Moreno, D; Otero, S; Arango, C	2011	Two-year diagnostic stability in early-onset first-episode psychosis.	Journal of Child Psychology and Psychiatry	No incidence data
Galderisi, S; Davidson, M; Kahn, RS: Mucci, A; Boter, H; Gheorghe, MD; Rybakowski, JK; Libiger, J; Dollfus, S; Lopzes-Ibor, JJ; Peuskens, J; Hranov, LG; Fleischhacker, WW; EUFEST Group	2009	Correlates of cognitive impairment in first episode schizophrenia: the EU-FEST study	Schizophrenia Research	No incidence data
Zeppegno, P; Probo, M; Ferrante, D; Lavatelli, L' Airoldi, P; Magnani, C; Torre, E	2009	First admissions for psychoses in Eastern Piedmont-Italy	The European Journal of Psychiatry	Other: paper could not be obtained
Essink-Bot, ML; Stronks, K.	2008	Immigrants in areas of low immigrant density have a greater incidence of psychotic disorders than native Dutch	Evidence Based Mental Health	No incidence data
Alonso, J; Lepine, JP; European Study of the Epidemiology of Mental Disorders	2007	Overview of key data from the European Study of the Epidemiology of Mental Disorders (ESEMed)	Journal of Clinical Psychiatry	Other: paper could not be obtained
Castro-Fornieles J, Paralllada M, Conzales-Pinto A, Moreno D, Graell M, Baeza I, Otero S, Soutullo CA, Crespo-Facorro B, Ruiz-Sancho A, Desco M, Rojas-Corrales O, Patino A, Carrasco-Marin E, Arango C, CAFEPS group	2007	The child and adolescent first episode psychosis study (CAFEPS): design and baseline results	Schizophrenia Research	No incidence data
Selten, JP, Veen, ND, Hoek HW, Laan W, Schols D, van der Tweel I, Feller W, Kahn RS	2007	Early course of schizophrenia in a representative Dutch incidence cohort	Schizophrenia Research	No original incidence data presented
Van Laar M, van Dorsselaer S, Monshouwer K, de Graaf R.	2007	Does cannabis use predict the first incidence of mood and anxiety disorders in the adult population?	Addiction	No clinically defined psychosis
Veling W, Selten JP, Susser E, Laan W, Mackenbach JP, Hoek HW	2007	Discrimination and the incidence of psychotic disorders among ethnic minorities in the Netherlands	International Journal of Epidemiology	No original incidence data presented
Ayuso-Mateos JL, Guttierrez-Recacha P, Haro JM, Chisholm D	2006	Estimating the prevalence of schizophrenia in Spain using a disease model	Schizophrenia Research	No incidence
Boonstra TC, Wunderink A, de Wit, P, Noorthoorn EO, Wiersma, D	2006	Incidence and gender distribution of non-affective psychotic disorders in the Netherlands	Conference abstract	Other: data covered elsewhere
Cougnard A, Parrot M, Grolleau S, Kalmi E, Desage A, Misdrahi D, Brun-Rousseau H, Verdoux H	2006	Patterns of health service utilisation and predictors of readmission after a first admission for psychosis: a 2-year follow-up study.	Acta Psychiatrica Scandinavica	No incidence
Henquet C, Murray R, Linszen D, van Os J	2005	The environment and schizophrenia: the role of cannabis use	Schizophrenia Bulletin	Other: no original data presented

Kovess-Masfety, V; Lecoutour, X, Delavelle, S	2005	Mood disorders and urban/rural settings: comparisons between two French regions.	Social Psychiatry and Psychiatric Epidemiology	No incidence
Ballon, N; Ursulet, G; Merle, S; Eynaud, M; Charles-Nicolas, A; Michalon, M	2004	Excess of psychoses among the French West Indian population	The Canadian Journal of Psychiatry	Other: paper could not be obtained
Messias E, Kirkpatrick B, Bromet E, Ross D, Buchanana RW, Carpenter WT jr, Tek C, Kendler KS, Walsh D, Dollfus S	2004	Summer birth and deficit schizophrenia: a pooled analysis from 6 countries	Archives of General Psychiatry	No original incidence data presented
Kalla, O; Aaltonen, J; Wahlstrom, J; Lehtinen, V; Cabeza, IG; Gonzales de Chavez, M	2002	Duration of untreated psychosis and its correlates in first-episode psychosis in Finland and Spain	Acta Psychiatrica Scandinavica	No incidence data
Kirkpatrick B, Herrera Castanedo S, Vazquez-Barquero JL	2002	Summer birth and deficit schizophrenia: Cantabria, Spain	Journal of Nervous and Mental Disorders	No original incidence data presented
Selten JP, Cantor-Graae E, Slaets J, Kahn RS	2002	Odegaard's selection hypothesis revisited: schizophrenia in Surinamese immigrants to the Netherlands	American Journal of Psychiatry	No incidence data
van Os J, Bak M, Hanssen M, Bijl RV, de Graaf R, Verdoux H	2002	Cannabis use and psychosis: a longitudinal population-based study	American Journal of Epidemiology	No clinically defined psychosis

Appendix 3A: MRC Sociodemographic Schedule

MRC SOCIODEMOGRAPHIC SCHEDULE (Amended) Part 1

STUDIE: EU GEI **Subject number:** |__|EU|__|__| - |__|__|__|__|	**Date of Birth** |__|__|-|__|__|-|1|9|__|__|
Time interval: Present **Interviewer:**	**Period – Replicat** |0|__|-|0|__| **Date** |__|__|-|__|__|-|2|0|__|__|

1a. Subject is:

O1 Case **O2 Sibling** **O3 Control**

1. Gender O1 Male O2 Female

2. Age \|__\|__\|

3. Ethnicity All Sites:

O1 White O2 Black O3 Mixed O4 Asian O5 North African O6 Other

4. Ethnicity Site Specific (UK)

O11 White British	O12 White Irish	O13 White gypsy, traveller	O14 Other White
O15 Mixed (w, bc)	O16 Mixed (w, ba)	O17 Mixed (w, as)	O18 Other Mixed
O19 Indian	O20 Pakistani	O21 Bangladeshi	O22 Chinese
O23 Other Asian	O24 Black Caribbean	O25 Black African	O26 Other Black
O27 Arab	O28 Other		

5. Place of Birth

O1 Austria	O2 Belgium	O3 France	O4 Germany
O5 Ireland	O6 Italy	O7 Spain	O8 Suisse
O9 The Netherlands	O10 Turkey	O11 United Kingdom	O12 Brazil
O13 Australia	O14 other, specify: _____		

6. Age of migration (if applicable) |__|__|

7. Father's place of birth

O1 Austria	O2 Belgium	O3 France	O4 Germany
O5 Ireland	O6 Italy	O7 Spain	O8 Suisse
O9 The Netherlands	O10 Turkey	O11 United Kingdom	O12 Brazil
O13 Australia	O14 other, specify: _____		

8. Mother's place of birth

O1 Austria	O2 Belgium	O3 France	O4 Germany
O5 Ireland	O6 Italy	O7 Spain	O8 Suisse
O9 The Netherlands	O10 Turkey	O11 United Kingdom	O12 Brazil
O13 Australia	O14 other, specify: _____		

9. First language:

O1 English	O2 German	O3 French	O4 Dutch
O5 Spanish	O6 Turkish	O7 Italian	O8 Kurdish
O9 Portuguese	O10 other, specify:_____		

10. Ever employed O0 No O1 Yes

11. Social class (Subject)

(provide descriptions only)	**Current**	**Main**
a. Job Title	_____	_____
b. What do (did) you mainly do?	_____	_____
c. What does/did organization make?	_____	_____

d. Social class subject Main:

O1 Higher grade Professional	O2 Lower grade Professional
O3 Intermediate occupations	O4 Small Employer and self employed occupations
O5 Self employed occupations	O6 Lower supervisory and lower technician occupations
O7 Lower services, sales and clerical occupations	O8 Lower technical occupations
O9 Routine Occupations	O10 Never worked and long-term unemployed

d. Social class subject Current:

O1 Higher grade Professional	O2 Lower grade Professional
O3 Intermediate occupations	O4 Small Employer and self employed occupations
O5 Self employed occupations	O6 Lower supervisory and lower technician occupations
O7 Lower services, sales and clerical occupations	O8 Lower technical occupations
O9 Routine Occupations	O10 Never worked and long-term unemployed

12. Social class Father (other_____)

(provide descriptions only)	**At birth of participant**	**Main**
a. Job Title	_____	_____
b. What do (did) you mainly do?	_____	_____
c. What does/did organization make?	_____	_____

d. Social class at birth of participant

O1 Higher grade Professional	O2 Lower grade Professional
O3 Intermediate occupations	O4 Small Employer and self employed occupations
O5 Self employed occupations	O6 Lower supervisory and lower technician occupations
O7 Lower services, sales and clerical occupations	O8 Lower technical occupations
O9 Routine Occupations	O10 Never worked and long-term unemployed

d. Social class father Main

O1 Higher grade Professional	O2 Lower grade Professional
O3 Intermediate occupations	O4 Small Employer and self employed occupations
O5 Self employed occupations	O6 Lower supervisory and lower technician occupations
O7 Lower services, sales and clerical occupations	O8 Lower technical occupations
O9 Routine Occupations	O10 Never worked and long-term unemployed

13. Mother's age at birth |__|__|

14. Father's age at birth |__|__|

15. Number of brothers and sisters |__|__|

16. Do you consider yourself to have (or ever have had) a hearing impairment? O0 No O1 Yes

17. Was the onset of the hearing impairment before the age of 18 years? O0 No O1 Yes

Note for DATA ENTRY: open **EU_LO, list of Physical Examination**)

(Note: Please use Metric system)

18. Height |_| **M** |_|_| **CM**

19. Weight |_|_| KG

20. Waist circumference |_|_|_| CM

(Note for DATA ENTRY: open **EU_LIVPLA_PREV** and **EU_LIVPLA_CURR**: Living places previously and current)

21. Where have you lived during your life, starting with the place you were born?

No.	Country	City/Town	Street/Postcode	Age					Change of School	
				From			To			
1.									O0 No	O1 Yes
2.									O0 No	O1 Yes
3.									O0 No	O1 Yes
4.									O0 No	O1 Yes
5.									O0 No	O1 Yes
6.									O0 No	O1 Yes
7.									O0 No	O1 Yes
8.									O0 No	O1 Yes
9.									O0 No	O1 Yes
10.									O0 No	O1 Yes
11.									O0 No	O1 Yes
12.									O0 No	O1 Yes
13.									O0 No	O1 Yes
14.									O0 No	O1 Yes
15.									O0 No	O1 Yes

Current

No.	Country	City/Town	Street/Postcode	Age				
				From			To	
1.								

(Note for DATA ENTRY: open **EU_MRC 2_SODEPAT**)

MRC SOCIODEMOGRAPHIC SCHEDULE (Amended) Part 2

1. **Since leaving your parents' home, have you lived with others?** O0 No O1 Yes

2. **Who do you live with…?(For b-d: Rate if age 17 or older at time; otherwise N/A not applicable)**

	Alone	Alone, with children	Partner, Spouse	Partner, Spouse, with children	Parents	Other family	Friends	Other: specify (e.g. hostel, halls of residence)	N/A
a) Now	O1	O2	O3	O4	O5	O6	O7	O8_____	O9
b) At onset	O1	O2	O3	O4	O5	O6	O7	O8_____	O9
c) 1 yr pre-onset	O1	O2	O3	O4	O5	O6	O7	O8_____	O9
d) 5 yrs pre-onset	O1	O2	O3	O4	O5	O6	O7	O8_____	O9

3. **Do you own/ rent your home…? (For b-d: Rate if age 17 or older at time; otherwise N/A not applicable)**

	Privately owned (self)	Privately owned (family)	Rented (Private)	Rented (government)	Other, specify:	N/A
a) Now	O1	O2	O3	O4	O5_____	O6
b) At onset	O1	O2	O3	O4	O5_____	O6
c) 1 yr pre-onset	O1	O2	O3	O4	O5_____	O6
d) 5 yrs pre-onset	O1	O2	O3	O4	O5_____	O6

4. **Overcrowding**

	a) Now	b) At onset	c) 1 yr pre-onset	d) 5 yrs pre-onset
i. How many persons live(d) with you?				
ii. How many rooms in your home (exclude kitchen and bathrooms)				

5. **Have you ever had a long-term relationship (one year or more)** O0 No O1 Yes

6. **How many children do you have?** |__|__|

7. What is your relationship status …? (For b-d: Rate if age 17 or older at time; otherwise N/A not applicable)

		Single	Married, living with someone	In a steady relationship	Divorced, separated	Widowed	N/A
a)	**Now**	O1	O2	O3	O4	O5	O6
b)	**At onset**	O1	O2	O3	O4	O5	O6
c)	**1 yr pre-onset**	O1	O2	O3	O4	O5	O6
d)	**5 yrs pre-onset**	O1	O2	O3	O4	O5	O6

8. What is the highest level of education you have achieved?

O1	School, no qualifications	(to end of compulsory education; passed no exams, tests, etc.)
O2	School, with qualifications	(to end of compulsory education; passed one or more exams, tests, etc.)
O3	Tertiary, Further	(first level of non-compulsory education; e.g. A-levels, Baccalaureate)
O4	Vocational	(job related education, e.g. teacher training, plumber, electrician, etc.)
O5	Higher (undergraduate)	(University; first degree)
O6	Higher (postgraduate)	(University: higher than first degree level, e.g. Masters, PhD)

9. How many years have you been in education? (from beginning of compulsory education) |__|__|

10. Are you employed (paid work) …? (For b-d: Rate if age 17 or older at time; otherwise N/A not applicable)

		Unemployed	Economically inactive (i.e. house person, physical illness/disability, career, retired)	Student	Part-time employee	Full-time employee	Self-employed	N/A
a)	**Now**	O1	O2	O3	O4	O5	O6	O7
b)	**At onset**	O1	O2	O3	O4	O5	O6	O7
c)	**1 yr pre-onset**	O1	O2	O3	O4	O5	O6	O7
d)	**5 yrs pre-onset**	O1	O2	O3	O4	O5	O6	O7

11. What is your monthly income …? (For b-d: Rate if age 17 or older at time; otherwise N/A not applicable)

		a) Now		b) At onset			c) 1 yr pre-onset			d) 5 yrs pre-onset		
i.	Nett monthly income £	_____		_____			_____			_____		
ii.	Income below median	O0 No	O1 Yes	O0 No	O1 Yes	O3 N/A	O0 No	O1 Yes	O3 N/A	O0 No	O1 Yes	O3 N/A
iii.	Income below official poverty line	O0 No	O1 Yes	O0 No	O1 Yes	O3 N/A	O0 No	O1 Yes	O3 N/A	O0 No	O1 Yes	O3 N/A
iv.	Receipt of welfare benefits	O0 No	O1 Yes	O0 No	O1 Yes	O3 N/A	O0 No	O1 Yes	O3 N/A	O0 No	O1 Yes	O3 N/A
	If YES, specify	_____		_____			_____			_____		

12. What is your religious affiliation?

O0 None O1 Christian O2 Jewish
O3Muslim O4 Other, specify _____

13. How often do you attend religious services?

O0 Never O1 Once or twice a year O2 Monthly O3weekly

For first-generation migrants only

14. Where, on a scale from 1 to 10, do you rate your fluency in the majority language? |__|__|
(1=not fluent at all, 10= very fluent)

(Note for DATA ENTRY: open **EU_LTE, list of threatening events**)

Discrimination

In the following questions we are interested in the way other people have treated you or your beliefs about how other people have treated you. Can you tell me if any of the following has ever happened to you? Please indicate number of times, age at first occurrence and note the main reason for this. **Then:**

For any reason, have you ever been unfairly...

	Yes	No	No. of times	Age (first occurred)
1. Fired	O1	O0		

Main reason: O1 Gender O2 Race, ethnicity O3 Religion O4 Mental Illness O5 Sexuality O6 Age O7 Other (specify):

	Yes	No	No. of times	Age
2. Not hired for a job	O1	O0		

Main reason: O1 Gender O2 Race, ethnicity O3 Religion O4 Mental Illness O5 Sexuality O6 Age O7 Other (specify):

	Yes	No	No. of times	Age
3. Denied promotion	O1	O0		

Main reason: O1 Gender O2 Race, ethnicity O3 Religion O4 Mental Illness O5 Sexuality O6 Age O7 Other (specify):

	Yes	No	No. of times	Age
4. Stopped, questioned threatened by police	O1	O0		

Main reason: O1 Gender O2 Race, ethnicity O3 Religion O4 Mental Illness O5 Sexuality O6 Age O7 Other (specify):

	Yes	No	No. of times	Age
5. Treated by court system	O1	O0		

Main reason: O1 Gender O2 Race, ethnicity O3 Religion O4 Mental Illness O5 Sexuality O6 Age O7 Other (specify):

For any reason, have you ever been unfairly...

	Yes	No	No. of times	Age
6. Discouraged from continuing education	O1	O0		

Main reason: O1 Gender O2 Race, ethnicity O3 Religion O4 Mental Illness O5 Sexuality O6 Age O7 Other (specify):

	Yes	No	No. of times	Age
7. Prevented from buying, renting flat or house	O1	O0		

Main reason: O1 Gender O2 Race, ethnicity O3 Religion O4 Mental Illness O5 Sexuality O6 Age O7 Other (specify):

	Yes	No	No. of times	Age
8. Treated by neighbours or your family	O1	O0		

Main reason: O1 Gender O2 Race, ethnicity O3 Religion O4 Mental Illness O5 Sexuality O6 Age O7 Other (specify):

	Yes	No	No. of times	Age
9. Denied a loan or preferable mortgage rate	O1	O0		

Main reason: O1 Gender O2 Race, ethnicity O3 Religion O4 Mental Illness O5 Sexuality O6 Age O7 Other (specify):

	Yes	No	No. of times	Age
10. Received worse service than other people	O1	O0		

Main reason: O1 Gender O2 Race, ethnicity O3 Religion O4 Mental Illness O5 Sexuality O6 Age O7 Other (specify):

	Yes	No	No. of times	Age
11. Treated when getting medical care	O1	O0		

Main reason: O1 Gender O2 Race, ethnicity O3 Religion O4 Mental Illness O5 Sexuality O6 Age O7 Other (specify):

	Yes	No	No. of times	Age
12. Treated when using public transport	O1	O0		

Main reason: O1 Gender O2 Race, ethnicity O3 Religion O4 Mental Illness O5 Sexuality O6 Age O7 Other (specify):