Author:
Turner, Nicholas

Title:
Modifiable Risk Factors for Depression in Adolescence – Understanding the Causal Role of Obesity and Physical Activity

Modifiable Risk Factors for Depression in Adolescence – Understanding the Causal Role of Obesity and Physical Activity

Nicholas Turner

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Health Sciences

Bristol Medical School

Population Health Sciences

October 2017

Word count: 45682

Abstract

Background

Adolescent depression is a significant burden to individuals, families and healthcare systems. Understanding modifiable risk factors, such as obesity and physical activity (PA), is key to informing preventative strategies. The aim of this project was to examine the causal relationships between obesity, PA and depression in adolescents.

Methods

Longitudinal data on obesity, PA and depression in adolescents came from 3 large international cohorts (ALSPAC N=7457, TRAILS N=2230 and NDIT=1294).

Linear regression and generalised estimating equations (longitudinal) were used to model effects of obesity on future depression. Cross-lagged structural equation modelling was used to investigate a bi-directional relationship between obesity and depression. Mendelian Randomization analysis was used to address residual confounding.

The same analytical approaches were used to examine the association between PA and depression. Partial least squares regression was used to identify aspect(s) of PA important in adolescent depression.

SEM was used to investigate the role of biological and psychosocial factors as mediators of the obesity-depression relationship.

Results

There was (inconsistent) evidence of a positive relationship between obesity and depression in females; a 1 SD increase in obesity was associated with a 0.035 SD (95% CI 0.003, 0.067) increase in depression at the next time point. There was evidence (in one cohort) that this relationship may be mediated by body image. There was no consistent evidence of any association between PA and subsequent depression (e.g. a 1 SD increase in PA was associated with a -0.006 SD (SE 0.016) decrease in depression at the next time point).

Conclusion

Reducing obesity may improve the mental health of adolescent females, alongside having physical health benefits. There is little evidence that increased levels of PA are beneficial for depression. Embedding data collection within existing cohorts approaching adolescence will further research in this area and potentially improve outcomes for future generations.

Acknowledgements

I would like to thank my brilliant supervisory team, Nicola Wiles, Kate Tilling and Jon Heron, for their advice, support, and enthusiasm and for all the time and effort they have put into guiding me and this thesis over the past three years. I am also grateful to my colleagues and fellow PhD students who have made working in the School of Social and Community Medicine such a pleasure.

I would like to acknowledge the participants and staff of the ALSPAC, TRAILS and NDIT cohorts, without whom this project would not have been possible. I am also grateful to the National Institute for Health Research for funding my PhD.

Finally I would like to thank my friends and family who have been incredibly supportive over the last three years.

Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: ..... DATE:.....

Contents

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

Depression represents a significant burden to both individuals and the National Health Service (NHS). The increase in depression in adolescence observed in recent years is concerning given the longer term negative consequences on educational attainment and social functioning, and the increased risk of depression in adulthood. However, there is little robust evidence regarding the role of potentially modifiable risk factors such as levels of obesity and physical activity (PA) in adolescent mental health. A small population change in a causal risk factor could lead to reductions in depressive symptoms, thereby improving the mental health of the population and reducing the cost and burden to the NHS. As such the investigation of the role of PA and obesity as potentially modifiable risk factors for depression and the mediators underlying these relationships is an important step in improving preventative strategies. The overall aim of this thesis is to investigate the causal relationship between obesity, physical activity and adolescent depression, in order to inform existing guidance on the prevention of adolescent depression and to identify potential targets for future intervention studies to reduce the burden of adolescent depression to individuals and the NHS.

This introductory chapter provides a brief background on adolescent depression, the public health problem that it represents and the rationale for the need for a preventative approach. Chapter 2 reviews the current literature surrounding the associations between obesity and adolescent depression, physical activity and adolescent depression. An introduction and summary of the cohorts providing data for the thesis and the analytical approaches used in the project will be presented in Chapter 3. The results and discussions of the analyses are presented in Chapters 4 to 7. Chapter 4 presents a descriptive analysis of the cohorts, detailing the characteristics of the three cohorts that have provided data. The

focus of Chapter 5 is the results and a discussion of the analyses examining whether obesity is associated with depression in adolescents. Chapter 6 is concerned with the results and a discussion of the investigation into the association between physical activity and adolescent depression. Chapter 7 investigates mediators of these relationships followed by a discussion of this analysis. Finally, Chapter 8 synthesizes the results, addresses the overall strengths and limitations of the work presented in the thesis and identifies areas that would benefit from future research.

1.1. Depression

Depression is a mood disorder that is characterised by a group of traits, behaviours and associated impairment. There are two main classification systems that aim to identify and measure these depressive symptoms, they are the International Classification of Diseases-10 (ICD-10) [1] and the Diagnostic and Statistical Manual of mental disorders-V (DSM-5) [2]. The symptoms that characterise depression according to these classification systems are [1, 2]:

- Persistent low mood (irritable mood is allowed by the DSM-5)
- Loss of interest or pleasure in activities
- Decreased energy or increased fatigability
- Low self-confidence or self-worth
- Unreasonable feelings of guilt, self-reproach or self-blame*
- Suicidal thoughts or behaviours
- Poor concentration or indecisiveness
- Change in psychomotor activity, either agitation or retardation
- Sleep disturbance

- Change in appetite, either increase or decrease

**This characteristic only appears on the ICD-10 classification system, not in DSM-5*

For an individual to be classified as depressed according to the ICD-10 two of the first three symptoms on the list above must be present, and in addition a further two of the other symptoms listed above must also be present within the same two week period [1]. For an individual to be classified as depressed according to the DSM-5 then one of the first two and a total of five symptoms of the list above must be present [2]. The ICD and DSM classification systems can also provide a grading of severity; severe, moderate or mild depressive episode, which is based upon the number and severity of the depressive symptoms and the level of associated functional impairment. The dichotomy of depressed/not depressed is useful for clinical decision making but depression may be considered as existing as a continuum within the population [3].

Depression has been rated as one of the five leading causes of worldwide disease burden and it is estimated that by the year 2030, in high income countries, depressive disorder will become the main cause of disability, whilst amongst low income countries only HIV/AIDs/perinatal conditions will rank higher [4]. Depression is also associated with many co-morbid conditions (for example cardiovascular disease and diabetes), increased mortality and impairments in many domains such as employment, physical and societal functioning and overall quality of life [5-8].

1.2. Adolescent depression

For the purpose of this thesis the working definition of adolescence is individuals aged 11 to 19 years. Depression in children (under 11) is rare, estimated at <1%,

however this figure rises dramatically during adolescence (11 to 19 years) and there is a particularly acute increase post puberty [9]. The prevalence of adolescent depression is alarmingly high (estimated at ~5%) with evidence that it is increasing [9-12]. This is an important public health problem as adolescence represents a critical period; a major cause of death amongst both adolescents and young adults is suicide, and depression during adolescence drastically increases an individual's risk of suicide [13], depression in adolescence increases rates of smoking and substance abuse [14], impacts negatively on educational attainment and social functioning [15] and leaves individuals more likely to experience depressive episodes (and of greater severity and persistence) later in life [16, 17].

1.3. Importance to the NHS

Depression represents a significant burden to those who suffer with it, their families and to the NHS. In 2007, the total costs to the NHS of treating children and adolescents with mental health problems were estimated at £143 million and it has been projected that this cost will rise to £237 million by 2026 [18]. Depression during adolescence increases risk of depression in later life and depression that persists into adulthood represents a substantial cost to the NHS. In 2007, the average service costs for adults in contact with services and/or receiving treatment was £2,085 [18]. The cost associated with treating depression further highlights the importance of developing effective preventative strategies.

1.4. Population based preventative approach

It is important to promote positive mental health in the general population, as a small improvement in depressive symptoms within the population could result in a large decrease in the percentage of people who are diagnosed with depression, whilst also resulting in a decrease in the percentage of individuals who suffer from sub-clinical levels of depressive symptoms [19]. This would

reduce the significant burden associated with adolescent depression to both individuals and the healthcare system. A population-based preventative approach to improving adolescent depression needs to focus on understanding potentially modifiable risk factors. Two such factors are obesity and physical activity, which are themselves top priorities of the public health agenda [20].

1.5. Adiposity and Obesity

Adiposity and obesity are two terms that have technically different definitions but are often used interchangeably. Adiposity refers to the amount of fat mass of an individual whereas obesity refers to excess fat tissue in an individual's body composition. A person is often classified as "obese" if he or she reaches a certain binary cut-point on a measure of fat mass. From a public health perspective however, obesity is often referred to when speaking about a wide range of levels of adiposity not just a dichotomy. For example the World Health Organisation defines obesity as a condition whereby "abnormal or excessive body fat has been accumulated to an extent that it may have a negative effect on health" [21]. Therefore in this project while using the term obesity this is in reference to a continuous scale of adiposity and not a binary "obese"/"not obese" cut-point.

There are a number of ways in which obesity may be measured, the most frequently reported is Body Mass Index (BMI); calculated as weight (kg) divided by height squared (m^2). Other methods for estimating obesity include scanning methods such as Dual-energy X-ray Absorptiometry (DXA) which measures both fat and lean mass in order to estimate an individual's overall body composition, and other anthropometric measures including waist circumference and skinfold thickness may also be used (see Section 3.3). Levels of obesity have been rising in the population in recent years, including amongst children and adolescents [12, 22]. Whilst it is well known that obesity is associated with deleterious effects on physical health, such as serious co-morbidities like cancer, cardiovascular disease

and diabetes [23], less is known about its relationship with adolescent mental health (see section 2.1).

1.6. Physical Activity

Increasing the level of physical activity in the population as a way to improve both physical and mental health is high on the public health agenda, with the current recommendation being that each day children and adolescents should aim to carry out (at minimum) an hour of moderate/vigorous physical activity [24]. However, the evidence to support the health benefits of physical activity at this recommended level is sadly lacking, and in particular the impact on mental health [25] (see Section 2.2).

1.7. Informing preventative strategies

In order to inform preventative strategies for depression we need to better understand the potentially modifiable causes. If, for example, low levels of physical activity are found to cause higher levels of depression, then a small population change in physical activity could generate large reductions in diagnosed and subclinical depressive symptoms, thereby improving the mental health of the population and reducing the cost and burden to the NHS. However, if in fact depression causes low levels of physical activity (i.e. people with depression become less active), then interventions to increase physical activity will not influence the mental health of the population. Elucidating the mechanisms of action will also help identify novel intervention targets, thus informing the first step towards the development of more successful and cost-effective preventative interventions. Overall a better understanding of the roles of factors such as obesity and physical activity in the aetiology of adolescent depression would inform UK Government guidance [26-28]. This, together with knowledge of the underlying mechanisms involved in these relationships, has

the potential to better inform public health policy and to identify avenues for preventative strategies to reduce the burden of adolescent depression to individuals and the NHS.

CHAPTER 2. LITERATURE REVIEW

This chapter will review the literature relating to the association between obesity or physical activity (PA) and depression in adolescents. There will be a discussion of previous findings and methodological issues relating to the existing literature. The chapter will then conclude with a statement of the aims and objectives of this thesis.

2.1. Obesity and depression in adolescence

One third of UK adolescents are obese (defined as greater than the 95th percentile of BMI for age) and the prevalence is increasing [12]. More than two thirds of obese adolescents will be obese adults, and obesity at this age is a risk factor for chronic pathologies in adulthood [29]. Obesity and depression both have a high prevalence, and are risk factors for many of the same diseases, such as diabetes and cardiovascular disease. Several biological (e.g. dysregulation of the hypothalamic-pituitary-adrenal axis) and psycho-social (e.g. social stigma) pathways between the two conditions have been hypothesised. It is therefore plausible that obesity and depression may be causally related. A potential relationship could be uni-directional (obesity causes depression or vice-versa), or bi-directional. Understanding the nature of the causal relationship is crucial to prevention, as intervening on obesity will clearly not prevent depression if in fact depression causes obesity.

To date much of the scientific literature investigating the relationship between obesity and depression in adolescence has been cross-sectional in nature [30]. Some of the cross-sectional literature suggests evidence of a positive association between obesity and depression in adolescence [31-33], whilst others found no evidence of an association [34-36]. These cross-sectional studies are from different settings, use different measures of obesity and depression but most importantly are limited in terms of determining the direction of causality. Due to this potential for reverse-causality in cross-sectional studies the focus of this literature review is on relevant longitudinal studies.

2.1.1. Is obesity prospectively associated with depression in adolescents?

In order to disentangle the causal nature of a relationship between obesity and depression, analysis of longitudinal data is required. The current literature from longitudinal studies is sparse (only five studies available) and those studies that are available are inconsistent in their findings. Some studies suggest a positive relationship between obesity and later depression in adolescents [37], whilst others have found no evidence of an association [38]. There have been three recent systematic reviews of the literature that have investigated whether obesity is associated with later depression in adolescents; Luppino et al 2010 [39], Korczak et al 2013 [40] and Hoare et al 2014 [41] (Table 2.1). The three reviews differed in the data bases and time frames that they searched; Luppino et al [39] searched three data bases, Korczak et al [40] searched two, one of which was the same as Luppino et al [39] (PubMed), whilst Hoare et al [41] searched four, one of which was the same as Luppino et al [39] (PsychINFO) and another the same as Korczak et al [40] (MEDLINE). The reviews by Luppino [39] and Korczak [40] were very comprehensive in the time periods that they investigated (all studies up to 2008, and studies between 1966 and 2012 respectively), whereas Hoare et al [41] focussed on more recent studies (between 2002 and 2013). All three reviews used similar search terms and exclusion criteria, with the exception that Luppino

et al [39] were reviewing studies of both adolescents and adults. The key thing that all three of the systematic reviews highlight is the lack of longitudinal studies available. Luppino et al [39] identified only two studies (of adolescents) that met the criteria for inclusion in their review, similarly Korczak et al [40] and Hoare et al [41] only identified four studies for inclusion. In total, across the three systematic reviews, only five different prospective longitudinal studies of adolescent obesity and depression were included [37, 38, 42-44].

Table 2.1 – Methodological details of the three recent systematic reviews investigating the association between obesity and depression in adolescence

Review	Databases searched	Time frame	Search terms	Exclusions	Other
Luppino et al 2010 [39]	PubMed, Embase, PsychINFO	Up to March 2008	"depression, depressive disorder, depressive symptoms, major depression, overweight, obesity, adiposity, body mass index, intra-abdominal fat, waist-hip ratio, waist circumference, metabolic syndrome"	"cross-sectional analyses, case reports, comments, letters, reviews, bipolar disorder, highly specific population (i.e. a specific disease), follow-up period less than 1 year, not collected BMI"	English language only
Korczak et al 2013 [40]	PubMed, MEDLINE	Between 1966 and 2012	"depression, attention deficit hyperactivity disorder, conduct disorder, behaviour problems, disruptive behaviour, body mass index, overweight, obesity"	"cross-sectional analyses, retrospective studies, primary outcome not depression/obesity/overweight/ BMI, initial assesment >18 years"	English language only
Hoare et al 2014 [41]	PsychINFO, MEDLINE, Cumulative Index to Nursing and Allied Health Literature, Health Source: Nursing Academic Edition, PsychArticles	January 2002 to April 2013	adolescen*, teen*, youth, depress*, depressed mood, depressive symptom*, BMI, body mass index, weight status, overweight, obes*, waist circumference, skin fold, central adiposity, diet, eating behav*, nutrition, physical activit*, exercise, sport*, sedentary beahv*, screen time, physical inactiv*	pilot/feasibility studies, reviews, highly specific population, not focussed on adolescents (10 - 19 years)	English language only

The five studies identified by the reviews all vary in terms of the populations, study design and length of follow ups investigated (Table 2.2). For example two of the studies investigated only females [38, 44], one of the studies was a secondary analysis of trial data [44] (whilst the others were observational studies) and the length between follow up measurements was highly variable between the five studies.

The methodological approaches used by the studies also differed greatly (Table 2.3). For example, whilst all the studies examined the association between obesity based on BMI and depression, the exposure variables themselves differed. One study derived latent classes of obesity [42], three studies used a three-level categorical variable (normal/overweight/obese) [37, 38, 43] and one study used a binary variable of obese/not-obese [44]. The analytical methods themselves also differed; some studies used generalised estimating equations (GEE) [38, 42, 44], whilst others used logistic regression [37] or cox regression [43].

The results of the five studies are inconsistent (Table 2.4); the overall conclusion that may be drawn from the literature is that there is evidence that obesity is related to later depression but it is not clear whether this relationship applies to all sub-groups within the population. One study found an increased risk of depression in boys belonging to a chronically obese trajectory but not in girls [42], one found that obese adolescents (defined as BMI >95% percentile) had an increased odds of depression and this association was observed in both boys and girls [37], whilst a third found evidence of increased risk of Major Depressive Disorder (MDD) in obese girls but not obese boys [43]. The final two studies analysed only girls (female only cohorts) with one finding no association between obese status and a classification of MDD but an association with increased depressive symptom score [38], whilst the final study found evidence

of an association between obesity and depression but only in white participants [44] (Table 2.4). In the three studies [37, 42, 43] that had data on both males and females all three carried out the analyses stratified by sex but none formally tested for a sex interaction. One of the studies only presented the results for males and not females, even though the analysis was carried out on both [42].

Table 2.2 – Descriptive details of studies identified by systematic reviews examining the association between obesity and depression in adolescence

Study	Sample	Total Sample size	Males	Females	Design	Age (years) at baseline	Follow up
Mustillo 2003 [42]	"Children recruited from 11 counties in western North Carolina (USA). Those in the top 25% of total scores on the Child Behaviour Checklist plus 1-in-10 random sample of the remaining"	991	505	486	Accelerated cohort (using 3 cohorts with different baseline ages)	Cohort 1 = 9 Cohort 2 = 11 Cohort 3 = 13	Annually until child was 16, then every 2 to 3 years
Herva 2006 [37]	"Recruited from live births from two northernmost provinces of Finland"	7512	3524	3988	Prospective birth cohort	14	17 years later
Anderson 2007 [43]	"Random sample of 976 families residing in Upstate New York"	701	350	351	Prospective cohort	14	Four follow ups over 20 years
Boutelle 2010 [38]	"Adolescent girls from 4 public and 4 private middle schools in the Austin, Texas metropolitan area"	496	0	496	Prospective cohort	13	Annually for 4 years
Anderson 2011 [44]	"Adolescent girls from 36 public middle schools six from each of San Diego, Tulane, Arizona, Maryland, Minnesota, South Carolina"	918	0	918	Secondary analysis of of data from a Randomized Controlled Trial	13	2 years after baseline

Table 2.3 – Methodological details of studies identified by systematic reviews examining the association between obesity and depression in adolescence

Study	Measure of obesity	Measure of depression	Exposure variable(s)	Outcome variable(s)	Analytical method	Adjusted for
Mustillo 2003 [42]	BMI was calculated from measured height and weight, obesity was defined as greater than or equal to the age and sex specific 95th percentile. Four distinct obesity trajectories were identified in the data: 1) never obese, 2) childhood obese, 3) adolescent obese, and 4) chronic obese	DSM-IV classification using the Child and Adolescent Psychiatric (CAPA) interview	Membership of one of four obesity trajectories	Binary depressed or not depressed variable	GEE	Age, sex, family income, other psychiatric disorder
Herva 2006 [37]	BMI from self-reported height and weight. Obese defined as equal or greater than the sex specific 95th percentile. Overweight defined as between the 85th and 95th sex specific percentiles	HSCL-25 questionnaire producing a score which a cut-point can be applied to in order to characterise participants as depressed	3-level categorical variable based on BMI percentile: 1) above 95th percentile, 2) 85-95th percentile, and 3) below 85th percentile	3 binary variables using different cut-points to define depression	Logistic regression	Father's social class, family type, chronic somatic diseases, smoking and alcohol use
Anderson 2007 [43]	BMI from self-reported height and weight. Obese defined as equal or greater than the sex specific 95th percentile. Overweight defined as between the 85th and 95th sex specific percentiles	DSM-III classification of major depressive disorder (MDD) using the Diagnostic Interview Schedule for Children	Categorical 3-level variable: obese, overweight, non overweight based on BMI percentiles Continuous - linear and quadratic BMI	Binary MDD classification	Cox regression	Socioeconomic status, race, smoking status, age
Boutelle 2010 [38]	BMI from measured height and weight. BMI scores converted to standardized Z-scores, overweight defined as a Z-score of >1.04 and obesity defined as a Z-score of >1.64	The Schedule for Affective Disorders and Schizophrenia for School-Age Children used as a continuous symptom score and also by classification of Major Depression by the DSM-IV	3-level categorical variable: obese, overweight, non overweight	Binary MDD classification Continuous depressive symptom score	GEE	Age, puberty, previous depression
Anderson 2011 [44]	BMI from measured height and weight. Obesity defined as ≥95th age-specific percentile.	The Centre for Epidemiological Studies Depression Scale was used to measure depressive symptoms and classify participants with Major Depressive Disorder	Binary variable classifying participantsas obese or not obese	Binary MDD classification	GEE	Age, family income, race, time spent home alone

Table 2.4 – Results of studies identified by systematic reviews examining the association between obesity and depression in adolescence

Study	Result: males and females combined	Result: sex interaction	Result: stratified by gender - males	Result: stratified by gender - females
Mustillo 2003 [42]	Results not presented	NA	Increased risk of depression in chronically obese males compared to never obese group (OR 3.7, 95% CI 1.3, 10.2).	Results not presented
Herva 2006 [37]	Results not presented	NA	Association between BMI above the 95th percentile and depression when using the highest cut-off value to define depression (OR 1.55, 95% CI 1.06, 3.68). No association when using other cut-offs	Association between BMI above the 95th percentile and depression when using the middle cut-off value to define depression (OR 1.64, 95% CI 1.16, 2.32). No association when using other cut-offs
Anderson 2007 [43]	Results not presented	NA	Categorical exposure - no evidence of association between BMI category and MDD. Continuous exposure - results not presented.	Categorical exposure - evidence of an association between being obese and MDD (HR 3.9, 95% CI 1.3, 11.8) Continuous exposure - no association with linear term but association with quadratic term (HR 1.3, 95% CI 1.1, 1.6)
Boutelle 2010 [38]	NA	NA	NA	There was no evidence of an association between being overweight (OR 0.61, 95% CI 0.24, 1.57) or obese (OR 1.62, 95% CI 0.77, 3.38) and MDD. There was evidence of an association between being obese and depressive symptom score (β 0.17, SE 0.05, p-value <0.01).
Anderson 2011 [44]	NA	NA	NA	Obese status associated with depression classification in white females (OR 2.09, 95% CI 1.44, 3.02) but not in Black or Hispanic females

The conflicting results found in the five studies may be due in part to the heterogeneity in both design (e.g. population, definition of the exposure variable, length of follow up) and analytical methods of the different studies. A detailed discussion of methodological issues in the existing literature will be given later (section 2.1.3). The authors of all three systematic reviews also raised concerns about lack of adequate adjustment for confounders in the included studies and lack of adjustment for previous level of depression. Other methodological problems such as the use of self-report rather than objective measures of obesity were also highlighted. The reviews called for further high quality studies to be carried out.

In addition to the traditional epidemiological studies, other investigations have employed more novel causal analysis approaches such as Mendelian randomization (MR) to investigate the relationship between obesity and depression in adolescence. MR uses genetic variants as a proxy for an exposure variable in an attempt to address the issues of residual confounding and reverse causation. The methodology of this approach will be discussed in more detail later (see Section 3.7).

Four previous studies have used an MR analysis to examine the relationship between obesity and depression (Table 2.5). The four studies differed greatly in sample size (range of sample sizes: 1731 to 53211) and mean age of the participants (range of mean ages: 33 to 57 years). The evidence from the MR literature is just as inconsistent as that from the observational epidemiological research. Two studies suggested that obesity causes an increased risk of depression [45, 46] whilst the other two reported that obesity causes a reduced risk of depression [47, 48] (Table 2.6). It is difficult to draw clear conclusions from these studies as none of them report the power of the analyses and only one reports an F-statistic (a measure of quality of the genetic instrument used in MR

analysis, an F-statistic of >10 is usually used to define a "good" instrument). All of the studies also utilised different genetic instruments with three out of the four studies using a single genetic variant rather than an allele score which would provide a more powerful analysis (see section 2.1.3). It should also be remembered that these analyses were not carried out specifically on adolescents and we should not extrapolate findings from adults to adolescents.

Table 2.5 - Details of MR studies examining the association between obesity and depression in adults

Study	Sample	Sample Size	Age (years)	Measure of obesity	Measure of Depression	Genetic instrument(s)
Lawlor et al 2011 [47]	Copenhagen General Population Study. Cross-sectional study of the general population, collecting genotypic and phenotypic data on a range of health related problems.	53211	57	BMI from measured height and weight	Three questions eliciting symptoms of anxiety and depression: "Do you feel nervous or stressed?", "Do you have the feeling that you have not accomplished very much recently?" and "Do you feel like giving up on life?". And one question asking if the participant was currently taking antidepressants. Each response was coded 1 for yes and 0 for no.	rs9939609, rs17782313
Kivimaki et al 2011 [46]	Whitehall II Study. London based office staff working in 20 UK government departments.	4145	44	BMI from measured height and weight	Presence of a Common Mental health Disorder (CMD) defined by the self report General Health Questionnaire (GHQ), a 30 item instrument listing symptoms. Respondent is required to state whether or not a symptoms is present. A participant is defined as a "case" if five or more symptoms are present.	rs1421085
Jokela et al 2012 [45]	Cardiovascular Risk in Young Finns study. A population based prospective cohort study.	1731	35	BMI from measured height and weight	Depressive symtpom score generated by summing responses to the 21 items of the Beck Depression Inventory (BDI)	Genetic risk score based on 31 SNPs
Samaan et al 2013 [48]	Data from the EpiDREAM, INTERHEART, DeCC and CoLaus studies	28493	33	BMI from measured height and weight	Major Depressive Disorder defined by DSM-IV	rs9939609

Table 2.6 Results of MR studies examining the association between obesity and depression in adults

Study	Power	F-statistic	Result
Lawlor et al 2011 [47]	Not reported	Not reported	There was evidence when using both SNPs as IVs that BMI was inversely associated with "nervous/stressed" question: OR 0.65, 95% CI 0.46, 0.91 and OR 0.36, 95% CI 0.19, 0.69. The second SNP was also inversely associated with the "not accomplishing" item (OR 0.48, 95% CI 0.24, 0.96).
Kivimaki et al 2011 [46]	Not reported	Not reported	Analysis carried out in men (n=2826) as the genetic instrument was not associated with BMI in females. Increased BMI was associated with CMD (β=0.166, 95% CI 0.025, 0.308)
Jokela et al 2012 [45]	Not reported	19.1	Evidence of an association between BMI and depressive symptom score. A one unit increase in BMI was associated with a 1.96 unit increase in depressive symptom score (95% CI 0.003, 3.90).
Samaan et al 2013 [48]	Not reported	Not reported	Meta-analysis of the results from the four studies showed evidence of an inverse relationship between BMI and depression (OR 0.92, 95% CI 0.89, 0.97)

2.1.2. Is depression prospectively associated with obesity in adolescents?

A recent systematic review of papers published between 1966 and 2012 investigating the association between depression and obesity in adolescence concluded that the current evidence suggests that depressed adolescents are at greater risk of future obesity [40]. Eight longitudinal studies were identified with most studies suggesting around a two to three-fold increased odds of becoming obese in those adolescents who are depressed [49, 50]. To date, there are no MR studies to investigate the association of adolescent depression with future obesity. This is likely due to there being no genetic variants that have found been found to be robustly associated with depression.

2.1.3. Methodological limitations to previous studies examining the association between obesity and depression in adolescence

There are a number of methodological issues relating to the previous research that may (at least partly) explain the conflicting evidence concerning the association between obesity and adolescent depression.

Measurement error

The measurement of obesity in the previous literature has been based on measurement of BMI. Some studies have used BMI based on self-reported height and weight [49, 51, 52] rather than objective measures. Self-report is less precise than objective measurement and may introduce bias into any analyses. Self-report tends to result in an underestimate of participant BMI (particularly at the higher end of the spectrum) as individuals tend to overestimate height and

underestimate weight [53]. It is also possible that people with different levels of depression may judge and/or report their BMI differently, resulting in bias in any analyses. In childhood studies even objectively measured BMI has been criticised as a measurement of obesity, as increased BMI may reflect increased lean mass more than increased fat mass [54, 55]. However, objectively measured BMI and direct measures of adiposity (fat mass) like dual-energy x-ray absorptiometry (DXA) are strongly correlated and the magnitude of associations of the different measures of adiposity with cardiovascular risk factors have been shown to be very similar [56]. To date none of the previous literature has investigated the association between obesity and depression in adolescence using a direct measure of adiposity such as DXA. Future studies should focus on objective measures of BMI and/or direct measures of adiposity.

Confounding and model adjustment

The recent systematic reviews in this field highlighted that, in the current studies, there was inadequate adjustment for important confounders i.e. variables that are associated with both the exposure and outcome variables but which do not lie on the causal pathway between them [39-41]. For example, no studies adjusted for maternal depression, and, in addition, sex, age and socio-economic status were only adjusted for in some studies. Lack of adjustment for these known important confounding factors could introduce bias. As well as this lack of adjustment for known confounders there is also the problem of unknown/residual confounders, as is the case for all observational epidemiological studies.

Some previous studies [57-59] did not measure depression at baseline, hence associations observed may reflect symptom persistence rather than a causal relationship. For example, participants who were obese at baseline and had a high depression score at follow up may have had a high depression score at baseline, which caused their obesity. Without appropriate adjustment for

baseline depression, it is the association between depression at baseline and follow up that is actually being observed rather than the relationship between baseline obesity and later level of depression.

Methodological issues in MR studies

Due to the potential for residual confounding, "classical" observational epidemiology alone cannot prove causation. However, methods such as MR can help strengthen the evidence for causal relationships [60] (see Section 3.7). In such studies genetic variants that are associated with the modifiable risk factor of interest are used as instrumental variables in order to make causal inferences about the relationship between the exposure and health-related outcomes. This approach eliminates the possibility of reverse causality and also ensures that estimates of associations are not subject to confounding [60, 61]. Previous MR studies [46-48] have often used one of the few large-effect genetic loci to serve as an instrumental variable for adiposity. This is problematic as this instrument is likely to be weak, explaining only a small amount of variance in adiposity, resulting in an imprecise estimate. The use of an allele score generated from several genetic variants would be a better instrumental variable [62] but such an approach has not been used in previous studies to investigate the association between obesity and adolescent depression. Another difficulty in the MR approach is that of power, for a MR analysis to have sufficient power it is often necessary to use a very large sample size.

2.1.4. Potential mechanisms underlying the relationship between obesity and depression in adolescence

Identification of factors on the causal pathway between obesity and depression (mediators) may provide novel intervention targets for the prevention of adolescent depression. Several biological and psycho-social mediators have been proposed; for example it has been suggested that inflammatory pathways may be a mechanism that links obesity and depression. Obesity causes chronic low-grade inflammation [63-65] and several studies have provided evidence of an association between inflammation and depression [66-68]. As inflammation has been observed in both depression and obesity then it is possible that it may mediate such an association.

Similarly, it is thought that the hypothalamic-pituitary-adrenal (HPA) axis may mediate an association between obesity and depression. It has been hypothesised that obesity may cause a dysregulation in the HPA axis and that this dysregulation may then lead to depression, or alternatively that depression may induce long-term activation of the HPA axis leading to obesity [69-71]. As well as biological mechanisms, several psycho-social causal pathways have been suggested; obesity may lead to body dissatisfaction, low self-esteem, functional impairment, reduced physical activity, poor rated self-health and social stigma, which may be risk factors for depression [72].

Although a number of potential mediators have been proposed and the mechanisms through which they may act have been suggested, few studies have formally examined the evidence for mediation. Those studies which are available have mostly used samples of adults [73-77] and a clinical population [78]. Only three studies were identified investigating mediation in a sample of children or adolescents [79-81] (Table 2.7). Of the three studies identified, two investigated

body image as a mediator [72, 79, 81] whilst the other investigated cortisol reactivity [80]. The three studies also differed in whether they considered obesity or depression as the outcome variable. All three of the mediation studies however used the Baron and Kenny approach to mediation, which is known to suffer from methodological limitations (see Section 3.7.3) [75, 78, 82] (Table 2.8).

Mond et al [82] found that an increase in depressed mood in obese compared with not obese participants was mediated by body dissatisfaction in males but not in females. Dockray et al [83] found evidence that in females the relationship between depressive symptom score and BMI was mediated by cortisol, whilst the third study [81] found that self-perception of body image fully mediated the observed association between major depressive disorder and obesity (Table 2.9). The small number of studies within an adolescent population highlights the need for future studies to focus on the identification of mediators that could represent viable targets for intervention during adolescence. Such future mediation studies need to utilise repeated data from prospective cohort studies and utilise more advanced statistical methodology in order to ensure robust findings.

Table 2.7 - Descriptive details of studies investigating mediators of the relationship between obesity and depression in adolescence

Study	Sample	Total Sample size	Males	Females	Design	Age (years) at baseline
Mond 2011	High-school students recruited from 31 schools in Minneapolis, Minnesota	806	366	440	Cross-sectional sample taken from a prospective cohort study	13
Dockray 2009	Recruited from the American Student List of children and adolescents in nearby countys to the research team (further details not provided)	111	56	55	Cross-sectional sample taken from a longitudinal study	10
Roberts 2015	Recruited from households in the Houston metropolitan area that were enrolled in two health maintenance organizations	4175	Not given	Not given	Two wave cohort study	Mean age not given. Range of ages stated as 11 to 17

Table 2.8 - Methodological details of studies investigating mediators of the relationship between obesity and depression in adolescence

Study	Measure of obesity	Measure of depression	Measure of mediator	Exposure variable(s)	Outcome variable(s)	Mediator variable(s)	Method	Adjusted for
Mond 2011	BMI was calculated from self-report height and weight. Obesity defined as greater than the 95th percentile.	Depressive mood was assessed using the Kandel Depressive Mood Scale	Body dissatisfaction was measured using the Body Shape Satisfaction Scale questionnaire	Binary obese/not obese variable	Continuous depressed mood score	Continuous body dissatisfaction score	Baron and Kenny	Age, socio-economic status and race
Dockray 2009	BMI from objectively measure height and weight	Depressive symptoms were measured using the Child Behaviour Checklist	Cortisol reactivity was assessd by measuring change in salivary cortisol levels following the Trier Social Stress Test for Children	Log transformed depressive symptom score	Continuous BMI score	Log transformed cortisol reactivity	Baron and Kenny	Age, pubertal stage
Roberts 2015	BMI was calculated from measured height and weight. Weight stauts was characterized as Healthy' if BMI was below the 95th percentile, and Obese' if above the 95th percentile	Depression measured using DISC-IV interviews administered via laptop	Body Image was measured by asking participants if they considered themselves skinny', 'somewhat skinny', considered themselves 'skinny', average weight', 'somewhat overweight', or 'overweight'	Binary variable referring to whether participant has experienced at least one major depressive episode in the last 12 months or not	Binary 'Healthy' or Obese' variable	Binary Body Image variable. Poor body image was defined as responding somewhat overweight' or 'overweight' to the body image question	Baron and Kenny	Age, gender, family income, physical activity and diet behaviour

Table 2.9 - Results of studies investigating mediators of the relationship between obesity and depression in adolescence

Study	Result: males and females combined	Result: sex interaction	Result: stratified by gender - males	Result: stratified by gender - females
Mond 2011	Results not presented	NA	Increase in depressive mood score in obese compared to not obese males was mediated by body dissatisfaction.	No evidence of mediation via body dissatisfaction.
Dockray 2009	Results not presented	NA	No evidence of mediation via cortisol reactivity	Evidence of cortisol reactivity as a mediator between BMI and depressive symptom score
Roberts 2015	Body image is the exposure increasing the risk for obesity, fully mediating the association initially observed between major depression and obesity	NA	NA	NA

2.1.5. Methodological limitations to previous studies investigating the potential mechanisms underlying the relationship between obesity and depression in adolescence

The literature focussed on mediators of the association between adolescent obesity and depression is sparse. Three studies were identified, two of which were cross-sectional in nature and all three utilised the Baron and Kenny approach to mediation [79-81]. The Baron and Kenny approach to mediation has been criticized heavily for several reasons (see Section 3.7.3), including that; confounding of the mediator and outcome is not handled appropriately, the method has been shown to have very low power, the existence of an indirect effect is inferred by the outcome of a set of hypothesis tests rather than based on the actual quantification of an indirect effect, and it requires a "significant" direct effect before testing for mediation whilst it is actually possible for there to be an indirect effect without a "significant" direct effect [84]. Therefore the use of more appropriate mediation methods, such as Structural Equation Modelling (see Section 3.7.3), in studies with longitudinal data on exposure, outcome and mediators are needed to inform the identification of potential targets for intervention.

2.2. Physical activity and depression

In the UK children and adolescents are advised [24] to each day carry out a minimum of one hour of moderate/vigorous physical activity (MVPA). The scientific evidence to support improvements in terms of general health, and even

more so in terms of mental health, from this level of physical activity is however sadly lacking [85].

2.2.1. Is physical activity prospectively associated with depression in adolescents?

The literature surrounding the relationship between physical activity and depression in adolescence is similar to the obesity literature in that it is sparse. A recent systematic review was only able to identify six appropriate longitudinal studies for inclusion in the review [41] (Table 2.10, Table 2.11 and Table 2.12). Of the six studies, five were prospective cohort studies and one was a secondary analysis of data from a Randomised Controlled Trial. All of the studies used a self-report measure of PA. These self-report measures were all different and therefore very different measures of PA were used as exposure variables in the different studies (See Table 2.9). The analytical approach used in the studies also differed, most of the studies took a regression based approach but there was also, for example, latent growth modelling.

The systematic review concluded that low levels of adolescent physical activity were associated with increased symptoms of depression, as five of the six studies seemed to suggest an inverse association between physical activity and depression [86-91]. Some of the studies presented findings for males and females combined and others stratified by sex, however none formally looked at an interaction with sex. The review did also point out that all of the papers used a self-report measure of physical activity (which in some studies consisted only of single question) which is likely to be less precise than an objective measure and may introduce bias to the analyses (see Section 3.4.1). The review also concluded that little is known about the relative importance of the intensity of activity, the frequency and total amount of PA undertaken or the type and context of the PA

(for example team versus individual PA), which is important to understand in order to provide the most useful public health message/policy.

Table 2.10 - Descriptive details of studies examining the relationship between physical activity and depression in adolescence

Study	Sample	Total Sample size	Males	Females	Design	Age (years) at baseline	Follow up
Fredricks 2006	Children recruited from Maryland as part of the Maryland Adolescent Development in Context Study	1060	519	541	Prospective cohort	Not presented	Five follow ups, with aprroximately 2.5 years between waves of data collection
Gore 2001	Recruited from high schools in three areas of Boston	1036	438	598	Prospective cohort	15	One year after baseline
Jerstad 2010	Adolescent girls recruited from eight schools in the southwest of the USA	496	0	496	Prospective cohort	13	Annually for six years
Motl 2004	Recruited as part of the TEENS randomized controlled trial from 7th and 8th grade students from Twin Cities Minnesota	3588	1830	1758	Secondary analysis of data from a Randomized Controlled Trial	13	Follow up two years after baseline
Rothon 2010	Participants were year 7 and 9 pupils from three London boroughs	2789	1367	1422	Prospective cohort	Not given Age range 11 - 14	Two years after baseline
Sund 2011	Participants were 8th and 9th gaders from two counties in Norway	2360	Not Reported	Not Reported	Prospective cohort	14	One year after baseline

Table 2.11 - Methodological details of studies examining the relationship between physical activity and depression in adolescence

Study	Measure of physical activity	Measure of depression	Exposure variable(s)	Outcome variable(s)	Analytical method	Adjusted for
Fredricks 2006	Self-report questionnaire on involvement in a range of extracurricular activites	Depressive symptoms were measured using the Children's Depression Inventory	Number of sports teams that participant played in	Continuous depression score	ANCOVA	Sex, race, parent education,
Gore 2001	Participants were asked what sports they were involved in, this question was then followed up with a likert scale regarding the amount of time spent in each activity	Depressive symptom score was calculated using the Centre for Epidemiological Studies Depression Scale	Continuous scale of amount of time participant spent in physical activity	Continuous depression score	Linear regression	Age, parental education, family structure, standard of living, BMI
Jerstad 2010	Modified version of the Past year Activity Scale. Participants reviewed a list of 26 activities and ticked those they had done more than 10 times over the year	The Schedule for Affective Disorders and Schizophrenia for School-Age Children used as a continuous symptom score and also classification of Major or Minor Depression	Total number of physical activites carried out by participant	Binary MDD classification Binary Minor depression Continuous depression	GLM	BMI percentile, body dissatisfaction, bulimic symptoms
Motl 2004	Single questionnaire item asking if participant takes regular physical activity. Responses were on a three point likert scale	Depressive symptoms were measured using the Centre for Epidemiological Studies Depression Scale	Physical activity latent trait	Depression latent trait	Latent growth model	Smoking and alcohol use, socio-economic status, sex, value placed on health
Rothon 2010	Single questionnaire asking how many hours a week participant exercises in their free time	Depressive symptoms were measured using the Short Mood and Feelings Questionnaire	Continuous score of how many hours per week participant exercises	Binary depressed/ not depressed variable	Logistic regression	Sex, general health, health behaviours, socio-economic status
Sund 2011	Vigorous exercise was assessed by asking the participants the number of hours they exercised vigorously in a week	Depressive symptoms were measured using the Mood and Feelings Questionnaire	Ordinal variable	Binary severe depression variable	Logistic regression	Sex, socio-economic status, BMI, pubertal stage, ethnicity, living situation, life events, physical health

Table 2.12 - Results of studies investigating relationship between physical activity and depression in adolescence

Study	Result: males and females combined	Result: sex interaction	Result: stratified by gender - males	Result: stratified by gender - females
Fredricks 2006	Increased number memberships to sports teams was associated with a lower level of depression	NA	Results not presented	Results not presented
Gore 2004	Results not presented	NA	Larger amounts of time spent in physical activity was associated with reduced depressive symtpom score (-0.72, confidence interval and p-value not provided but stated as "significant")	Larger amounts of time spent in physical activity was associated with reduced depressive symtpom score (-0.75, confidence interval and p-value not provided but stated as "significant"
Jerstad 2010	NA	NA	NA	Evidence that greater number of activites carried a reduced risk of Major depression (RR 0.86, 95% CI 0.77, 0.95, p-value 0.005), Minor Depression (RR 0.90, 95% CI 0.85, 0.96, p-value 0.002)
Motl 2004	A 1 SD increase in physical activity trait was associated with a 2.5 SD decrease in depression trait (p>0.005)	NA	Not presented	Not presented
Rothon 2010	Not presented	NA	No evidence of an association between amount of exercise and depression (OR 0.99, 95% CI 0.89, 1.09, p-value 0.810)	No evidence of an association between amount of exercise and depression (OR 0.95, 95% CI 0.88, 1.03, p-value 0.226)
Sund 2011	Lower amounts of vigorous exercise were associated with increased odds of severe depression (OR 1.23, 95% CI 1.01, 1.49, p<0.05)	NA	Not presented	Not presented

2.2.2. Is depression prospectively associated with physical activity in adolescents?

The literature investigating the relationship between depression and later levels of PA is very sparse as most research is interested in the other causal direction to either prevent or treat depression using physical activity. Studies that are available suggest that increased depression is associated with lower subsequent PA (an inverse relationship) [88, 92].

2.2.3. Methodological limitations of previous studies of physical activity and depression

Measurement error

A key methodological issue is how PA is measured. Self-report questionnaires are commonly used, but children and adolescents may not accurately recall PA [93], additionally, adolescents are often active in short bursts which are difficult to capture by self-report [94]. This is likely to reduce the precision in the measurement of PA and may also potentially introduce bias; when self-report PA is used as an exposure variable this will bias results towards the null. Cohort studies often use different self-report questionnaires to measure PA and as such this makes comparing results across studies problematic. PA can be measured objectively using personally-worn devices such as accelerometers (see Section 3.4). However, an individual's PA may change as a result of wearing such a monitoring device, and these instruments do not capture all PA (e.g. accelerometers must be taken off for swimming).

Confounding and model adjustment

In common with the literature focussed on obesity and depression, adjustment for potential confounders is limited [88, 95-99]. For example none of the previous literature adjusts for maternal levels of depression, one study did not adjust for socio-economic status, both of which may be important confounders. None of the studies adjusted for previous levels of depression, as such findings may be a persistence of symptoms and not a causal relationship between PA and depression.

Lack of genetic instrument for use in MR studies

Again, evidence from MR studies has the potential to strengthen the evidence for a causal relationship between PA and depression. However, no genetic variants have been associated with PA at the level of genome wide significance for use as an instrument in MR analyses [100, 101].

2.2.4. Potential mechanisms underlying the relationship between physical activity and depression

Not much is known about what factors may lie on the causal pathway between PA and depression. There is some evidence that PA may result in a promotion of the release of 5-hydroxytryptamine and cell proliferation, a reduction in level of cortisol and an increase in brain-derived neurotropic factor level [102-104]. PA may also increase social support and self-esteem, provide a distraction from negative thoughts, increase social networks/interaction, give a structure to daily life and help regulate sleeping and eating patterns [104]. It is also thought that depression may have an influence on PA via factors such as lack a of energy, sleep disturbances, low mood, anhedonia and social withdrawal [105].

2.3. Thesis aims

The overall aim of this thesis is to examine the causal relationship between obesity, PA and adolescent depression.

The specific objectives are:

1. To examine whether obesity is prospectively associated with depression in adolescents.

2. To examine whether physical activity is prospectively associated with depression in adolescents.

3. To examine whether biological factors (such as markers of inflammation) and psychosocial factors (e.g. body dissatisfaction, stigma and social support), that could represent novel intervention targets, mediate these relationships.

I will use sophisticated statistical methods to conduct cross-cohort analyses using data from three international cohorts to model the causal relationships between these factors. This will help address the inconsistencies and gaps in the current evidence base in order to inform preventive strategies to improve the mental health of adolescents and potentially reduce the associated burden to the NHS.

CHAPTER 3. METHODS

3.1. Datasets

The data for this thesis comes from a cohort collaboration that I established specifically for this project. The cohort collaboration comprises three cohorts all with repeated measurements of adolescent depression, obesity and PA. These three cohorts are the UK based Avon Longitudinal Study of Parents and Children (ALSPAC), the Dutch Tracking Adolescents' Individual Lives Survey (TRAILS) cohort and the Canadian Nicotine Dependence in Teens (NDIT) cohort.

3.1.1. ALSPAC

ALSPAC is a population based prospective birth cohort designed to examine the influences of genetic, biological, psychological, social and other environmental factors on development, health and behaviour [106, 107]. Pregnant women who were living in the UK (in what was at the time) the county of Avon were invited to take part in the study if their estimated delivery date was between 1st April 1991 and 31st December 1992. In total 20,248 eligible pregnancies were identified, of these 15,247 were successfully enrolled and a total of 14,701 study children were alive 1 year post-natally. The study children have been followed up regularly; 68 waves of follow up measurements have taken place between the study children being born and age 18. These assessment waves have included child self-report questionnaires, questionnaires relating to the study child but filled out by the mother (or other primary care giver) and clinical assessments.

3.1.2. TRAILS

TRAILS is a cohort based in the Netherlands that aims to better understand the causes and mechanisms involved in mental health disorders and social development of adolescents and young adults [108, 109]. The municipalities of five regions in the North of the Netherlands were asked to provide basic information (such as name, date of birth, gender, address etc.) of all those born between 1st October 1989 and 30th September 1990 (in the first two municipalities) and 1st October 1990 and 30th September 1991 (for the remaining three municipalities). Following this, all primary schools (135 schools, 3483 children) in the regions were contacted and provided with information about the proposed TRAILS study. Of these 135 primary schools, 123 agreed to participate (potential number of participants = 3145 children). A total of 210 children were excluded due to being unable to participate because of severe physical illness, handicap or mental retardation. The parents/guardians of the remaining eligible children (n=2930) were contacted and invited to take part in the study (a total of 705 did not respond). A total of 2230 children were included in the first wave of measurements between March 2001 and July 2002 (at a mean age of 11.1 years). Since this baseline measurement a further four follow up waves have taken place.

3.1.3. NDIT

NDIT is a prospective cohort study based in the Montreal area of Canada [110]. The primary aim of the NDIT cohort was to investigate the determinants and course of cigarette smoking and nicotine addiction in adolescents. The study began in 1999. High schools in the Montreal region of Canada were selected after consultation with head teachers and school boards to consist of a mixture of both English and French language schools, rural, suburban and urban schools and also schools from a mixture of high, middle and lower socio-economic neighbourhoods (private schools were excluded). A total of 13 high schools were identified to participate. These 13 schools were then reduced to 10 due to very poor response to parental consent

forms. All grade 7 students (mean age 12.8 years) were given take-home information about the study, a total of 1294 of the 2325 eligible students (56%) provided data at the baseline data collection. Data was collected via in class self-report questionnaires every 3 months from grade 7 to 11 (1999-2005). In addition to this, participants had anthropometric measures collected in selected survey cycles.

3.2. Measurement of depression

3.2.1. Observer rated vs. self-report measures of depression

Depression may be measured using self-report questionnaires or observer-rated instruments [111]. The measurement of depression by a clinician or clinicians in a large scale epidemiological study is likely to be expensive, time consuming, and impractical logistically. As such it has been suggested that in the assessment of child/adolescent mental health that other individuals, parents and teachers, may provide observer rated measures [112]. However, this also presents a problem; teachers are likely to be useful in identifying externalising behaviours, such as disobedience and fighting, but may not be useful for children/adolescents as they may not identify a child's internalising behaviours. Whilst parents may pick up on some emotional problems they are still likely to identify less problems than child self-report [113] – symptoms of depression in children and adolescence may be misclassified by parents and teachers as "a difficult child", or "typical moody teenager". Another potential problem with observer rated measures is the issue of inter-rater reliability; different observers may rate the same individuals differently. Due to these problems with clinician and other observer rated measures of depression, many epidemiological studies utilise self-report depression instruments. Most self-report measures consist of questionnaires asking if participants have experienced various symptoms of depression (and sometimes how often and to what degree) within a specified time period. A potential issue with self-reported depression measures is the under-reporting of depressive symptoms compared to

clinical assessment. There is evidence within community samples that self-report prevalence of depression is lower than what is identified when using clinical interviews [114].

3.2.2. Measures of depression available in the cohorts

Only a small proportion of adolescents are likely to fulfil diagnostic (ICD/DSM) criteria for depression and given that depression exists as a continuum within the population, depression symptom score will be considered in a continuous form. Different self-report measures of depression have been used in the three Cohorts; ALSPAC - Short Mood and Feelings Questionnaire (SMFQ) [115], TRAILS - Youth Self Report (YSR) [116] and NDIT - the Kandel Depressive Symptom Score (KDSS) [117].

The SMFQ (ALSPAC) was developed in the 1990's by Angold, Costello and Messer who aimed to produce "a short questionnaire for use in epidemiological studies of depression in children and adolescents" by reducing the longer Moods and Feelings Questionnaire (MFQ) [115]. The questionnaire consists of 13 statements relating to low mood (e.g. "I felt miserable or unhappy") and other related psychological correlates (such as low self-esteem and self-worth) and asks respondents to rate each of these statements as "not true", "sometimes true" or "true" (scoring 0-2 respectively) for the past two weeks. The score from each individual item can then be summed producing a total score within the range of 0 to 26. The total score may be dichotomised to classify individuals as depressed or not-depressed, with a cut-point of 11 having previously been shown to have a high sensitivity and specificity when judged against the Clinical Interview Schedule – Revised Form [111].

The YSR (TRAILS) is a self-report questionnaire consisting of 112 items covering emotional and behavioural problems in the past 6 months [116]. The 13 item

Affective Problems Scale (APS) of the YSR covers depressive symptoms and may be used to classify individuals as depressed/not-depressed according to DSM-IV criteria. The 13 items of the APS scale are a list of problems which are scored on a 3-point Likert scale ("not true or never true" = 0, "sometimes or a bit true" = 1, "often or very true" = 2). The scores of the items are then averaged to create a depressive symptom scale ranging from 0 to 2.

The KDDS (NDIT) is a self-report questionnaire that asks participants to report on a four point Likert scale the frequency (never, rarely, sometimes or often) with which they have experienced six symptoms of depression in the past three months [117]. These six symptoms are: "felt too tired to do things", "had trouble going to sleep or staying asleep", "felt unhappy, sad or depressed", "felt hopeless about the future", "felt nervous or tense" and "worried too much about things". A depression score is calculated by summing the scores from each of the individual items and then dividing this total by the number of items which were responded to. This produces a depressive symptom score ranging from one to four.

As the different cohorts have each used a different instrument to measure depression, standardised Z-scores (raw score – mean score / standard deviation) will be used in the analyses. This will transform the scores from the different depression measures onto one scale (with a mean of zero and SD of one) aiding in comparison of results across the cohorts.

3.2.3. Genetic instrument for depression

At the time of conducting this project there is no reliable genetic instrument for depression, although research in this area is ongoing.

3.3. Measurement of obesity

3.3.1. Self-report vs. objective measurement of obesity

There are a number of different methods that may be used to measure obesity in adolescence. When considering which measurement method is appropriate to use the difference between self-report and objective measures must first be discussed. Self-report measures are likely to be less precise than objective measures. Consider for example BMI calculated from height and weight of a participant (weight (kg) divided by height squared (m^2)) - actually measuring the height and weight of a participant will likely be far more precise than asking the participant to estimate these measures [54, 118, 119]. As well as a lack of precision, the use of self-report measures of obesity has the potential to introduce bias to a study. Studies have shown that, in both adults and children, when self-report measures are used, erroneously low prevalence of overweight and obesity are produced due to systematic under-reporting of body weight and overestimation of height [120-124]. For example in one study of 418 adolescents, 13.9% were classified as "overweight" whilst 2.8% were classified as "obese" when BMI was calculated from self-report of height and weight. However, when this was compared to concurrently measured BMI from objectively measured height and weight the prevalence of "overweight" and "obese" was far greater; 18.7% and 4.4% respectively [121]. When considering an analysis of obesity and depression it is also possible that those with greater depressive symptoms may judge and/or report their level of obesity differently to those with lower levels of depression, another potential source of bias that exists in a self-report measure that can be avoided with the use of an objective measure. Due to issues of precision and bias in self-report measures, an objective measure of obesity is preferred.

A commonly used objective measure is BMI from measured height and weight. However, even objectively measured BMI has well recognised limitations that

should be acknowledged. One major criticism is that BMI is not a direct measure of body fat and is not able to distinguish between lean mass and fat mass [54, 119]. Although BMI has been shown to correlate with fat mass it is also correlated with non-fat mass e.g. two individuals of the same height and weight may differ in body composition in terms of percentage fat mass and percentage lean mass but would still have the same BMI [125]. Body composition also changes with age and level of sexual maturity, however these changes in body composition are not necessarily reflected in someone's height and weight and therefore may not be captured by BMI [54, 119, 125]. Alternative measures to BMI that directly measure body fat are Dual energy x-ray absorptiometry (DXA) fat mass and subscapular skinfold thickness. A DXA scan directly measures fat, lean and bone mass, from which body fat percentage can be calculated. During a DXA scan an individual is x-rayed, the level of attenuation of the x-ray radiation through body tissue varies dependent on the composition of the tissue; fat, lean and bone tissue all cause attenuation of differing amounts. As such the level of attenuation can be used to determine body composition [126]. Subscapular skinfold thickness is a measure of subcutaneous fat, callipers are used to measure the thickness of a fold of skin below the point of the shoulder blade [127].

3.3.2. Measures of obesity available in the cohorts

An objective measure of obesity – BMI, calculated as participants' measured weight (in kg) divided by their measured height squared (in m^2), was collected in all three cohorts. Although BMI has been criticized the use of BMI does however have the advantage that this measure has been collected in all of the cohorts used in the analyses, allowing cross-cohort comparison. Use of BMI also allows comparison with the existing literature, as BMI is the measure most frequently cited, and is useful in clinical practice as it is the measure of obesity most commonly used by clinicians e.g. in growth monitoring. Some of the cohorts have collected additional measures of adiposity e.g. DXA fat mass (ALSPAC), subscapular skinfold thickness (NDIT and

TRAILS) and waist circumference (ALSPAC, NDIT and TRAILS). As a sensitivity analysis, the analyses were repeated (where possible) using these additional measures of obesity.

3.3.3. Genetic instrument for obesity

To date 97 genetic variants have been robustly associated with BMI in a recent meta-analysis of genome wide association studies [128]. These 97 genetic variants can be summarized into a single variable, an allele score, which can be used as an instrumental variable in MR analyses (see Section 3.7.1) [129]. Data on the 97 genetic variants associated with BMI were available in the ALSPAC and TRAILS cohort, but not in the NDIT cohort.

3.4. Measurement of physical activity

The term 'physical activity' is imprecisely defined, as there are a variety of different aspects to physical activity. For example, total amount, frequency, intensity, whether the activity is part of a group sport or individual activity etc., these different aspects may all be measured, and may be important for different outcomes. As it is not well understood what aspect of PA may be important in relation to depression in adolescence, there is no "gold standard" measure of PA in this situation. The difference between PA and sedentary behaviour should also be noted. Sedentary behaviour is a distinct behaviour which is different from lack of physical activity; sedentary behaviour is a group of behaviours that are carried out whilst sitting or lying down that require only a very low energy expenditure (e.g. watching television). It is possible for someone to do enough physical activity to reach the levels recommended by national guidelines but to also be considered sedentary if they spend a lot of time sitting down (perhaps at a computer for work/school).

3.4.1. Self-report vs. objective measurement of physical activity

As was the case discussed above regarding measurement of obesity, PA can be measured using self-report and objective measures. Self-report measures such as questionnaires and activity diaries are cheap and easy to administer making them a popular choice for many epidemiological studies. However, asking a study participant to recall their PA over a set time period is likely to be far less accurate and precise than an objective measurement. Some studies have suggested that children and adolescents are poor at accurately recalling their PA [93]. Children are also often active in short bursts which are hard to capture on self-report questionnaires [94]. As well as this problem of recall there is also the potential problem of reporting bias due to factors such as social desirability (i.e. participants may over report their level of activity as being active may be more socially desirable). Due to these limitations in self-report measures, an objective measure of PA is preferred. Two commonly used objective measures of PA are pedometers and accelerometers. A pedometer measures activity by recording the number of steps taken by an individual based on the movement of a pendulum within the device, whereas an accelerometer measures the amount and force of activity based on microelectromechanical movement. A limitation of both pedometers and accelerometry is that swimming and cycling cannot be captured by either device. A further limitation is that it is possible that an individual may change their activity levels because they are aware that they are wearing the device and that their activity is being recorded. An advantage of accelerometers over pedometers is that a measure of activity intensity can be calculated using an accelerometer and this is not possible with a pedometer.

3.4.2. Measures of physical activity available in the cohorts

Objective data on PA based on accelerometery was only available in one cohort - the ALSPAC cohort. In the TRAILS and NDIT cohort only self-report PA data was available.

Participants in the ALSPAC cohort were asked to wear around their waist, on their right hip, the Actigraph AM7164 2.2 accelerometer (Actigraph LLC, Fort Walton Beach, FL, USA) for a week (seven consecutive days) [130-132]. These accelerometers were to be worn by the participants at all times (when awake) except for if the device was likely to get significantly wet (for example when the participant showered, had a bath or was taking part in water sports such as swimming or water polo). Only participants who wore the accelerometer for at least 10 hours per day for a minimum of three days were included in the analyses, this criterion has been shown to produce the greatest power and a high level of reliability [130-132]. The accelerometer records PA data in movement counts, these counts are then averaged over a specified time period (1 minute). The daily mean number of minutes spent sedentary, in light, moderate and vigorous activity was defined by the cut-points ≤199, 200 – 3599, 3600 – 6199 and ≥6200 counts per minute (cpm) respectively. These PA intensity cut-points were defined by a calibration study using a sub-sample of the ALSPAC participants [130-132]. The sum of the average daily number of minutes spent in light, moderate and vigorous activity was used in the analyses to investigate the relationship between depression and amount of PA. Both total time spent in moderate to vigorous PA (MVPA) (i.e. ≥3600 cpm) and proportion of time spent in MVPA were also used in the analyses to investigate the relationship between PA intensity and depression. A binary variable was also created from the accelerometer data to identify those participants who met the UK guidelines [24] of achieving at least one hour of MVPA per day.

All three cohorts collected self-reported PA by asking about time spent in various activities. In the ALSPAC cohort a single question was asked; "In the past year how often did you carry out physical activity/exercise" – "never", "less than once a month", "1-3 times a month", "1-4 times a week", or "5+ times a week". Similarly in the TRAILS cohort, a single question asked: "How many days in an average week do you take part in physical activities?", with the possible responses being on an eight point Likert scale which ranged from 0="never" to 7="seven days a week". In the NDIT cohort, physical activity was measured using an adapted Time Spending Pattern questionnaire [133]. This asks participants on how many days in the past week they participated (for at least 5 minutes) in any of a list of 29 physical activities. Certain activities on the 29 item list have previously been defined as representing moderate or vigorous activity [134]: bicycling, swimming/diving, basketball, baseball/softball, footfall, soccer, racket sports, ice/ball hockey, jumping rope, downhill skiing, cross-country skiing, ice skating, rollerblading/skateboarding, exercise/physical conditioning, ball-playing, track and field, playing games, jazz/classical ballet, outdoor play, karate/judo/tai chi, boxing/wrestling, mixed walking, and running/jogging. The number of days on which participants took part in these activities were summed to produce a score representing the number of sessions of MVPA a participant engaged in per week.

Analyses were carried out across all of the cohorts using the (unstandardized) questionnaire data and also using the accelerometry data within ALSPAC – allowing us to learn from the one cohort where both measures are available about how we interpret the findings from studies using self-reported PA.

3.4.3. Genetic instrument for physical activity

To date no single nucleotide polymorphisms (SNPs) have shown an association with PA at the level of genome wide significance for use as an instrument in MR analyses [100, 101]. A genome wide aggregation score has the potential to recover information

that is lost by the dismissal of false negative findings in GWAS by assessing the combined contribution of variation across the genome in a specific trait (in this case PA). Physical activity GWAS data was only available in the ALSPAC cohort, therefore a genome wide aggregation score was used as a genetic instrument in an MR analysis of physical activity and depression within ALSPAC (see Section 3.7.2).

3.5. Potential mediators

Certain variables may lie on the causal pathway between obesity and depression in adolescents (i.e. obesity may lower self-esteem which causes an increase in depressive symptoms). These intermediate variables (known as mediators) may explain the effect of an exposure variable (obesity) on an outcome (depression) (see Figure 3.1). A number of variables were investigated as potential mediators in order to attempt to identify novel prevention targets. Data on potential psychosocial mediators: body image and self-esteem were available in the ALSPAC and TRAILS cohort (Table 3.1). Data on potential biological mediators: cortisol, C-reactive protein and IgE were available in the TRAILS cohort. Data on CRP was also available in the ALSPAC cohort. No data on potential mediators were available in the NDIT cohort.

Figure 3.1 - Simple diagrammatic representation of mediation

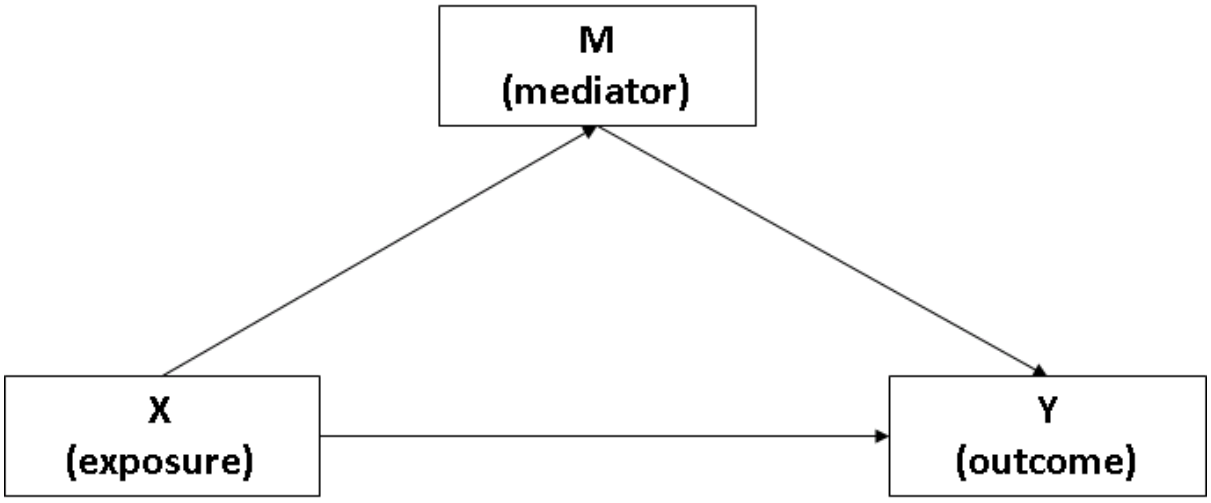


Table 3.1 – Data on potential mediators available in the three cohorts

Potential Mediator	ALSPAC	TRAILS	NDIT
Cortisol		X	
C-Reactive Protein	X	X	
IgE		X	
Body Image	X	X	
Self-Esteem		X	

In the TRAILS cohort cortisol was measured from saliva samples that the participants took themselves using a swab device that was provided by the research team. Participants were asked to collect two saliva samples, the first immediately upon waking in the morning and the second 30 minutes later. Since all schools participating in the TRAILS studies started at the same time in the morning the saliva sampling time variation is likely to be fairly small and has been estimated at approximately 7.00 am for the first sample and 7.30 am for the second.

Participants were asked to collect the samples when both the sampling day and the day previously were normal school days without any special occasions or stressful events. Participants were also asked not to take a sample on a day if they felt unwell or, in the case of girls, if they were menstruating. Cortisol levels within the samples were assessed using solid phase time-resolved fluorescence immunoassay with fluorometric end point detection [135].

In the ALSPAC cohort cortisol was similarly measured from saliva samples. Participants were asked to collect saliva samples using provided swab devices immediately upon waking and then again 30 minutes later on three consecutive "normal" school days. Level of cortisol in the salivary samples was determined using enzyme immune assay [136].

C-reactive protein was measured in the ALSPAC cohort through blood samples provided by participants during a clinical assessment visit. Participants were asked to fast overnight if their clinic appointment was in the morning and for a minimum of six hours if their appointment was in the afternoon [137, 138]. Similarly in the TRAILS cohort, CRP was measured from blood samples given by participants during a clinical assessment after a period of fasting [139]. In the TRAILS cohort serum IgE levels were also measured from blood samples collected at a clinical assessment.

3.6. Potential confounders

A confounding variable is a variable that is associated with both the exposure and outcome variable but is not on the causal pathway between them (see Figure 3.2). Lack of adjustment for confounders may result in bias, and this is one of the main drawbacks to observational epidemiology. Potential confounders considered included: sex; age; socio-economic position (SEP); maternal depression; and participant substance use (cigarette smoking and alcohol use) (where available, see Table 3.2).

Figure 3.2 - Simple diagrammatic representation of confounding

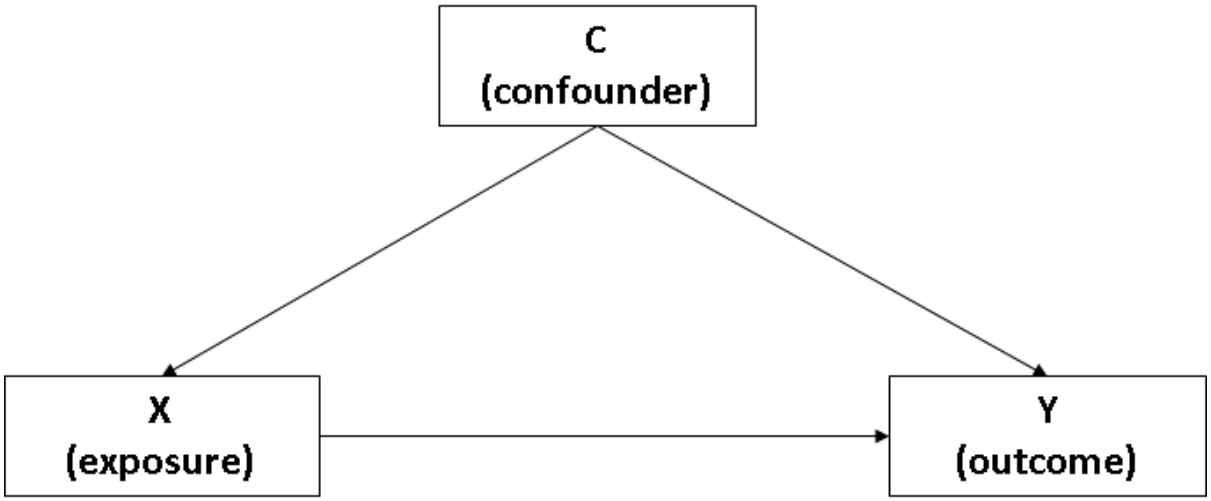


Table 3.2 – Data on confounders available in the three cohorts

Confounders	ALSPAC	TRAILS	NDIT
Age	X	X	X
Gender	X	X	X
Socio-economic position	X	X	X
Maternal depression	X	X	
Alcohol use	X	X	X
Smoking	X	X	

In each cohort, age was measured at each wave of data collection and gender was recorded at the time of enrolment to the cohort. In the ALSPAC cohort SEP was measured using two variables; level of maternal education and maternal social class. These two variables were completed by the participant's mother when she was at 32 weeks gestation stage with the study child. The maternal education variable asked the mother what her highest level of education was: "none", "vocational", "O-level", "A-level", or "University Degree". The social class variable was based on the mother's profession: "unskilled", "partly skilled", "skilled (manual)", "skilled (non-manual)", "managerial and technical", or "professional". In TRAILS, SEP was measured at study enrolment and is a variable that splits individuals into lowest 25%, middle 50% and highest 25% of SEP based on parental education, profession and income. In NDIT, SEP was measured using two variables collected at enrolment; maternal education and maternal profession. The maternal education question asked the mother of the study participant what her highest level of education was: "High School – attended", "High School – graduated", "CEGEP – attended", "CEGEP – graduated", "University – attended", "University – graduated BSc", "University – graduated MSc", or "University – graduated PhD". The maternal profession variable asked the mother how she would describe her job status; "employed full time", "employed part time", "full time student", "part time student", "homemaker", "not working due to health reasons", "unemployed", or "on welfare".

Data on participant alcohol consumption was available in the ALSPAC, TRAILS and NDIT cohorts. However, in the ALSPAC cohort the inclusion of alcohol consumption as a confounder dramatically reduced the number of participants that could be included in the analyses, whilst inclusion of alcohol did not alter the relationship between BMI and depressive symptoms, as such it was excluded (see Appendix 2). In the first wave of data collection for the TRAILS cohort participants were asked if they had ever drunk alcohol, with the possible response categories: "never"; "once"; "2-3 times"; "4-6 times"; or "7 times or more". In the second, third and fourth waves of data collection participants were asked to write down the number of alcoholic drinks they had consumed in the past week. In the NDIT cohort, the participants were asked at each wave of data collection if they had ever drunk alcohol with the possible response categories being: "never"; "a bit to try"; "once or a couple of times a month"; "once or a couple of times a week"; or "every day".

Data on cigarette smoking was available in the ALSPAC and TRAILS cohorts. However, in the ALSPAC cohort the inclusion of smoking as a confounder reduced the number with data available for analysis and the instrument used to measure smoking varied between time points and as such was excluded (see Appendix 3). Cigarette smoking data from the TRAILS cohort was included as a potential confounder. In the TRAILS cohort at each time point participants were asked how often they smoked cigarettes. At the first data collection point the response categories were: "not at all"; "sometimes"; or "often". At the second and third data collection points the response categories were: "I've never smoked"; "not during the last month"; "less than one a week"; "less than one a day"; "1-5 a day"; "6-10 a day"; "10-20 a day"; or "more than 20 a day". At the fourth data collection point the response categories were: "I've never smoked"; "not during the last month"; "less than one a week"; "less than one a day"; "1-5 a day"; "6-10 a day"; "10-20 a day"; "21-30 a day"; or "more than 30 a day".

Data on maternal depression was available in the ALSPAC and TRAILS cohorts. Within the ALSPAC cohort maternal depression was measured at 32 weeks gestation using the Edinburgh Postnatal Depression Scale (EPDS) [140]. The EPDS is a 10 item self-report depression questionnaire that is often used in the perinatal period because it does not contain any items relating to physical symptoms. In the TRAILS cohort maternal depression was measured using the Depression Anxiety Stress Scales (DASS) [141]. The DASS is a self-report questionnaire that consists of 42 items relating to negative emotional symptoms. The 42 items can be split into three groups of 14 items, with the three subgroups representing subscales for depression, anxiety and stress. The depression subscale of the DASS was used as the measure of maternal depression at the first wave of data collection in the TRAILS cohort. In the NDIT cohort no measure of maternal depression was collected.

It has been suggested that, in females particularly, pubertal stage may be associated with depressive symptoms[142] [143]. To investigate the potential impact of puberty a sensitivity analysis was carried out using the ALSPAC data (data only available in ALSPAC); a measure of puberty were investigated as a potential confounder/covariate. The measure of puberty investigated was whether or not the participant had experienced their first period (data collected at ages 10years 8months, 12years 10months, 13years 10months and 17years 10months).

3.7. Statistical analyses

The statistical software package STATA version 14 was used for all analyses described below with the exception of structural equation modelling which was carried out in MPlus and partial least squares regression which was carried out in R. Descriptive statistics were produced for each of the three cohorts to provide a brief summary and overview of the different data sets. These included how certain variables changed over time and their associations with depression.

3.7.1. Objective 1 – obesity and depression

The analytical approach employed in this project moves from simple linear regression through to more complex statistical modelling techniques utilising the longitudinal repeated measures nature of the data available. This enables a comparison with findings from other studies whilst also providing the most robust evidence of the association between obesity and depression.

Linear Regression

Linear regression was used to examine the effect of BMI (exposure) on depressive symptoms score (Z-score) at the next follow up occasion (outcome), for each occasion, for each of the three cohorts, adjusting for relevant confounders (see section 3.6). The linear regression model was also fitted including an interaction term between BMI and sex in order to formally test whether the association between obesity and depression was different in males and females. Linear regression models were also fitted separately for males and females, this stratified analysis was carried out in addition to the formal test for an interaction as this test was likely to be underpowered [144]. Analyses were repeated using alternative objective measures of obesity (DXA fat mass, waist circumference and subscapular skinfold thickness), where available.

As a sensitivity analysis a quadratic BMI term was included into the linear regression model in the ALSPAC cohort in order to test for a potential "U" shaped relationship between obesity and depression.

ALSPAC

In the ALSPAC cohort the time points at which measures of depression and obesity were collected are outlined in Table 3.3 below.

Table 3.3 - Obesity and Depression data collected in ALSPAC

		Data Collection Wave							
		F10 10y 7m	F11 11y 6m	TF1 12y 10m	TF2 13y 10m	TF3 15y 6m	CCS 16y 8m	TF4 17y 10m	CCT 18y 6m
Depression:	SMFQ	X		X	X		X	X	X
Obesity:	BMI	X	X	X	X	X		X	
	DXA		X		X	X		X	
	Waist Circumference	X	X	X	X	X			

From Table 3.3 it can be seen that regression models using BMI as the exposure and SMFQ as the outcome (adjusting for depression measured at the same time as the exposure) could be carried out across four time points:

(1) exposure measured at 10y 7m and outcome at 12y 10m

(2) exposure at 12y 10m and outcome at 13y 10m

(3) exposure at 13y 10m and outcome at 16y 8m

(4) exposure at 17y 10m and outcome at 18y 6m

For each of these analyses four regression models were carried out:

- Model 1 – BMI as exposure on SMFQ depression Z-score outcome, adjusted for age (at measurement of outcome), sex, previous depression score, maternal education, maternal social class and maternal depression.
- Model 2 – model 1 plus BMI*sex interaction term.
- Model 3 – model 1 restricted to males and without sex as a confounder.
- Model 4 – model 1 restricted to females and without sex as a confounder.

TRAILS

In the TRAILS cohort the time points at which measures of depression and obesity were collected are outlined in Table 3.4 below.

Table 3.4 - Obesity and Depression data collected in TRAILS

		Data Collection Wave			
		T1 10y 7m	T2 13y 1m	T3 15y 9m	T4 18y 7m
Depression:	YSR	X	X	X	X
Obesity:	BMI	X	X	X	X
	Waist Circumference			X	X
	Subscapular Skinfold Thickness	X		X	

From Table 3.4 it can be seen that regression models investigating obesity (as BMI) on depression at the next follow up could be analysed for three occasions:

(1) exposure measured at age 10y 7m and outcome at 13y 1m

(2) exposure at 13y 1m and outcome at 15y 9m

(3) exposure at 15y 9m and outcome at 18y 7m.

For each of these regression analyses, four models were examined:

- Model 1 – BMI as exposure on YSR APS depression Z-score outcome, adjusted for age, sex, previous depression score, social class, maternal depression, cigarette smoking and alcohol use.
- Model 2 – model 1 plus BMI*sex interaction term.
- Model 3 – model 1 restricted to males and without sex as a confounder.
- Model 4 – model 1 restricted to females and without sex as a confounder.

NDIT

In the NDIT cohort the time points at which measures of depression and obesity were collected are outlined in Table 3.5 below. From Table 3.5 it can be seen that regression models investigating the association between obesity and depression at the next follow up could be analysed for three occasions:

(1) exposure at age 12y 9m and outcome at 13y 0m

(2) exposure at 15y 7m and outcome at 15y 10m

(3) exposure at 17y 0m and outcome at 17y 1m

In each of these regression analyses four models were examined:

- Model 1 – BMI as exposure on Kandell depression Z-score outcome, adjusted for age, sex, previous depression score, maternal education, maternal social class, maternal depression and alcohol use.
- Model 2 – model 1 plus BMI*sex interaction term.
- Model 3 – model 1 restricted to males and without sex as a confounder.
- Model 4 – model 1 restricted to females and without sex as a confounder.

Regression coefficients and associated 95% confidence intervals and p-values will be reported for each of the above models.

Table 3.5 - Obesity and Depression data collected in NDIT

		Data Collection Wave									
		1	2	3	4	5	6	7	8	9	10
		12y 9m	13y 0m	13y 2m	13y 8m	13y 10m	14y 1m	14y 2m	14y 7m	14y 10m	15y 0m
Depression:	Kandel	X	X	X	X	X	X	X	X	X	X
Obesity:	BMI	X									
	Waist circumference	X									
	Subscapular skinfold thickness	X									

		Data Collection Wave									
		11	12	13	14	15	16	17	18	19	20
		15y 2m	15y 7m	15y 10m	16y 0m	16y 2m	16y 6m	16y 9m	17y 0m	17y 1m	20y 1m
Depression:	Kandel	X	X	X	X	X	X	X	X	X	X
Obesity:	BMI		X						X		
	Waist circumference		X						X		
	Subscapular skinfold thickness		X						X		

Generalized Estimating Equations

The linear regression analyses described above ignore the fact that repeated measurements on individuals are available and hence full use is not being made of the available data. Therefore to utilise all the data effectively, it is necessary to move towards models that incorporate repeated measurements. In standard generalised linear models (GLM) a key assumption is independence amongst the response data. In longitudinal studies where repeated measurements of the same outcome variable are taken on the same individuals this assumption of independence will clearly not hold. This lack of independence means that an analysis simply combining repeated measurements would result in standard errors that are too small and as such confidence intervals that are too narrow. Therefore, appropriate techniques to account for such non-independence are needed.

Generalized Estimating Equations (GEE) were used to model the repeated exposure-outcome association [145]. GEE models produce 'population-averaged' parameter estimates, hence the coefficient represents an average association over the population not the association for a particular individual within the population. So, for the models of interest, they give the average association between obesity and future depressive symptoms.

GEEs are an extension of GLM to correlated data using quasi-likelihood estimation. To take account of the correlation between repeated measurements on the same individual over time, a working correlation structure is applied to the set of response data. There are various choices of different correlation structures available to apply to the model depending on the data set [145]. Commonly used correlation structures include independent – no correlation within clusters; exchangeable – correlations are the same for all observations within a cluster; auto-regressive – correlation depends on the amount of time between measurements; unstructured – no assumptions made about the correlations, the correlations are estimated from the data; and fixed – the

user must specify a correlation matrix. The choice of which working correlation structure to use can be narrowed depending on three conditions relating to the spacing of the observations within the data set, as certain structures cannot handle certain data formats. These three conditions are: 1) balance – the data are balanced if each measurement occasion has the same number of observations, 2) equal spacing – the data are equally spaced if the time interval between observations is constant, and 3) gaps – the data have gaps if some observations are missing [145] (see Table 3.6)

Table 3.6 - Available correlation structures for GEE models

	Characteristics Allowed		
Correlation Structure	**Unbalanced**	**Unequal Spacing**	**Gaps**
independent	yes	yes	yes
exchangeable	yes	yes	yes
autoregressive	yes	no	no
stationary	yes	no	no
unstationary	yes	no	no
unstructured	yes	yes	yes
fixed	yes	yes	yes

Once the choice of correlation structure has been narrowed by the characteristics of the data set being analysed then final selection of a correlation structure can be achieved by fitting the GEE model using the different possible correlation structures and calculating and comparing a model fit statistic for each model. As GEE uses quasi-likelihood estimation rather than maximum likelihood, some of the statistics based on likelihood theory that are often used in model selection cannot be applied to GEE. A commonly used model fit statistic, the Akaike information criterion (AIC) [146], is not applicable to GEE, however a modification of the AIC termed the quasi-information criterion (QIC), may be used to assess model fit in GEE [147]. The model producing the lowest QIC value will be selected as the preferred correlation structure.

If the correlation structure is misspecified the parameter estimates produced by GEE are still robust, however, the standard errors produced may not be accurate. To provide a valid estimate of standard error even in the case of a misspecified correlation structure then the Huber-White sandwich estimator can be used [148].

The data being used in this project were unbalanced, unequally spaced and contained gaps. As such the correlation structures that may be used were the independent, exchangeable, unstructured or fixed structures. As mentioned above, in this analysis repeated measurements of the same outcome variable were taken on the same individuals and as such the independent correlation structure should not be used. The fixed structure may also not be used as this requires the specification of a known "assumed" correlation matrix. Therefore in this analysis either the exchangeable or unstructured correlation structures may be used. The GEE analysis was carried out using both possible correlation structures and the value of the QIC statistic produced from the two models was calculated using STATA 14's [149] *qic* command. The QIC values were compared and the correlation structure that produced the lowest QIC value was selected for the final models. The *vce(robust)* option of STATA 14's [149] *xtgee* command was used. This option specifies that the Huber-White sandwich estimator be used in the calculation of standard error.

GEE models were fitted within each cohort using lagged BMI as the exposure variable and standardised depressive symptom score as the outcome with regression coefficients, 95% confidence intervals and p values being reported for each model. All models were adjusted for baseline BMI, baseline depression, lagged depression, age, sex, maternal depression and socio-economic status. A lagged variable means the value of that variable at the previous time point. In the ALSPAC cohort four GEE models were fitted:

- Model 1 – Lagged BMI as exposure on SMFQ depression Z-score outcome, adjusted for baseline BMI, baseline depression score, lagged depression score, age, sex, maternal education, maternal social class and maternal depression.
- Model 2 – model 1 plus lagged BMI*sex interaction term.
- Model 3 – model 1 restricted to males and without sex as a confounder.
- Model 4 – model 1 restricted to females and without sex as a confounder.

In the TRAILS GEE analysis there was additional adjustment for alcohol and smoking use. Four GEE models were fitted:

- Model 1 – lagged BMI as exposure on YSR APS depression Z-score outcome, adjusted for baseline BMI, baseline depression score, lagged depression score, age, sex, social class, maternal depression, cigarette smoking and alcohol use.
- Model 2 – model 1 plus lagged BMI*sex interaction term.
- Model 3 – model 1 restricted to males and without sex as a confounder.
- Model 4 – model 1 restricted to females and without sex as a confounder.

In the NDIT cohort the following four GEE models were fitted:

- Model 1 – Lagged BMI as exposure on Kandell depression Z-score outcome, adjusted for baseline BMI, baseline depression score, lagged depression score, age, sex, maternal education, maternal social class, maternal depression and alcohol use.
- Model 2 – model 1 plus lagged BMI*sex interaction term.
- Model 3 – model 1 restricted to males and without sex as a confounder.
- Model 4 – model 1 restricted to females and without sex as a confounder.

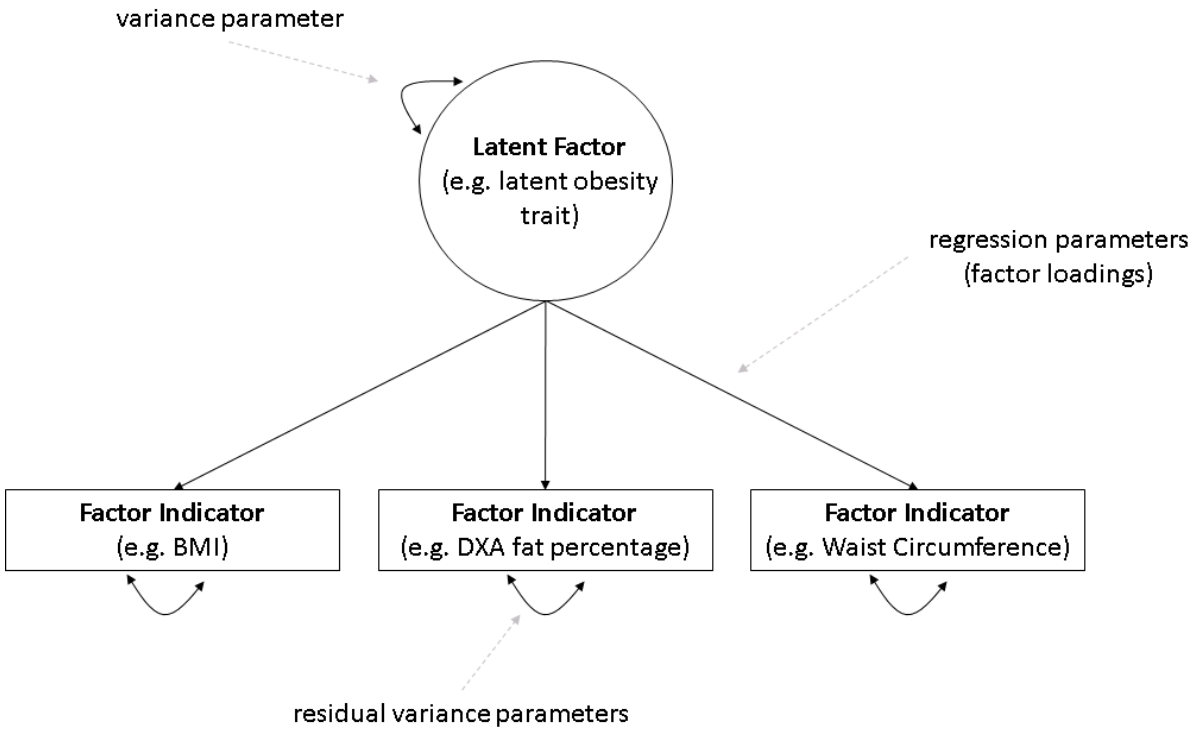
The results of the GEE analyses from each of the three cohorts were pooled to produce a single result using inverse variance weighted meta-analysis (STATA 14's *metan* command) [149].

Structural Equation Models

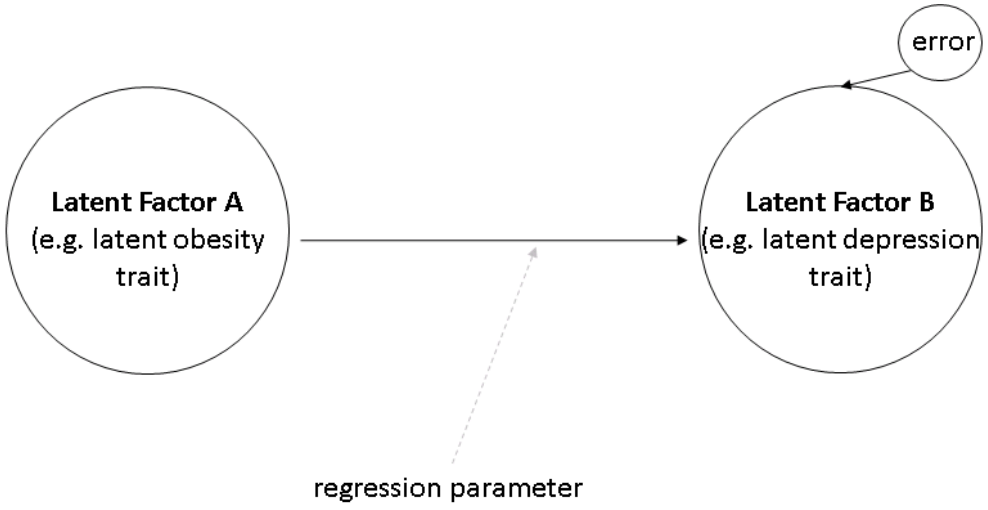
A cross-lagged structural equation modelling (SEM) approach was then used to investigate the potential bi-directional nature of the relationship. Cross-lagged models are an approach which can be used to examine the interplay between two processes which develop and influence each other in parallel. The "cross-lag" refers to a variable being regressed on a different variable from the previous time point (i.e. depression at time point 2 is regressed on obesity at time point 1 and vice-versa). SEM is a modelling framework where regression models with latent variables are fitted – a latent variable refers to a variable that is not directly observed but is inferred from other directly observed variables. SEM can be used to test how groups of observed variables may define unobserved constructs, how these unobserved constructs may be related to one another, and crucially permits the estimation of directional associations between multiple dependent variables (unlike standard or multivariate regression). SEM consists of two parts: 1) a measurement model part and, 2) a structural model part. The measurement model part of SEM consists of a set of regression equations that describe the relationship between observed variables (known as factor indicators) and unobserved latent factors (e.g. the relationship between BMI, DXA fat mass and waist circumference and a latent obesity trait). The structural model, in a single set of multivariate regression equations, can describe associations between latent factors, the relationship between observed variables, and finally, the relationship between observed variables (which are not factor indicators) and latent factors, depending on which of these associations are included in the SEM (see Figure 3.3).

Figure 3.3 - Graphical representation of a simple SEM, showing the measurement and structural models

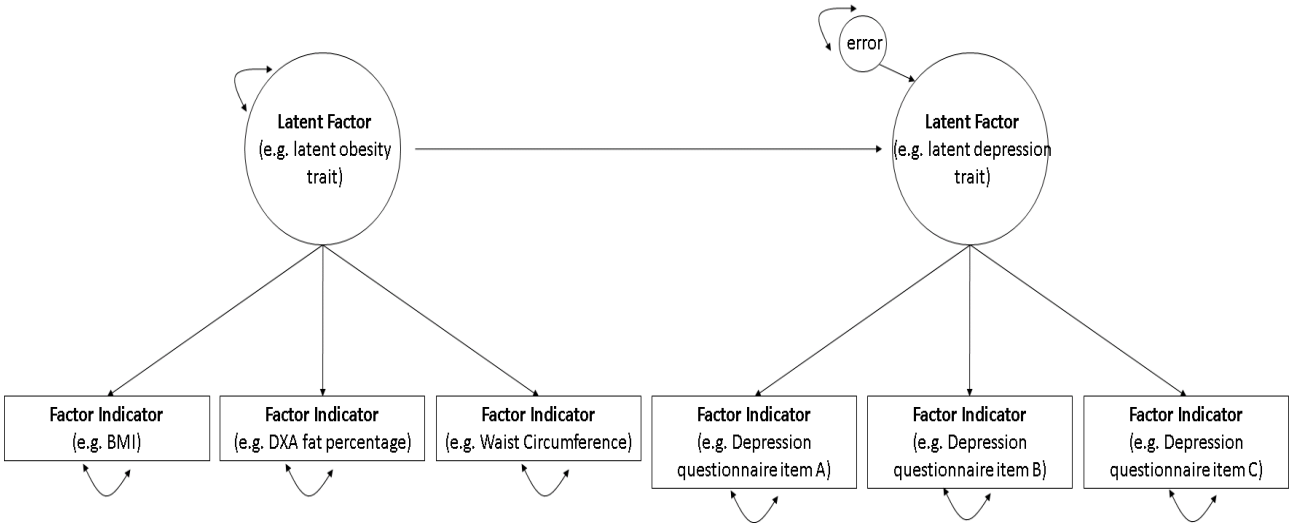
(A) Measurement model part of SEM



(B) Structural model part of SEM



(C) Combined measurement and structural models in an example of a simple SEM



In many cohort studies the data is largely comprised of responses to questionnaires. This is particularly true when collecting observational data on mental health. Depression for example, is often measured by participants responding to questionnaire items that rank the frequency of certain depressive symptoms. When using SEM with such measurements it is common to use groups of items (referred to as item parcels) rather than each individual item of a questionnaire as factor indicators. Item parcelling refers to taking two or more items and grouping them together (i.e. by summing or averaging the scores of the individual items) and then using these as factor indicators. For example in a depression questionnaire consisting of 12 items you could group these into 3 item parcels each consisting of 4 summed items. These 3 item parcels would then be used as factor indicators to define the latent depression variable. In this analysis the use of item parcels allowed the use of Maximum Likelihood estimation without incurring multiple dimensions of integration, as would have been the case if individual depression questionnaire items had been used. There are a number of other advantages to the use of item parcels in SEM compared to the use of individual items as factor indicators, these advantages are outlined in Table 3.7 [150].

Table 3.7 - Advantages of using item parcels rather than items as factor indicators

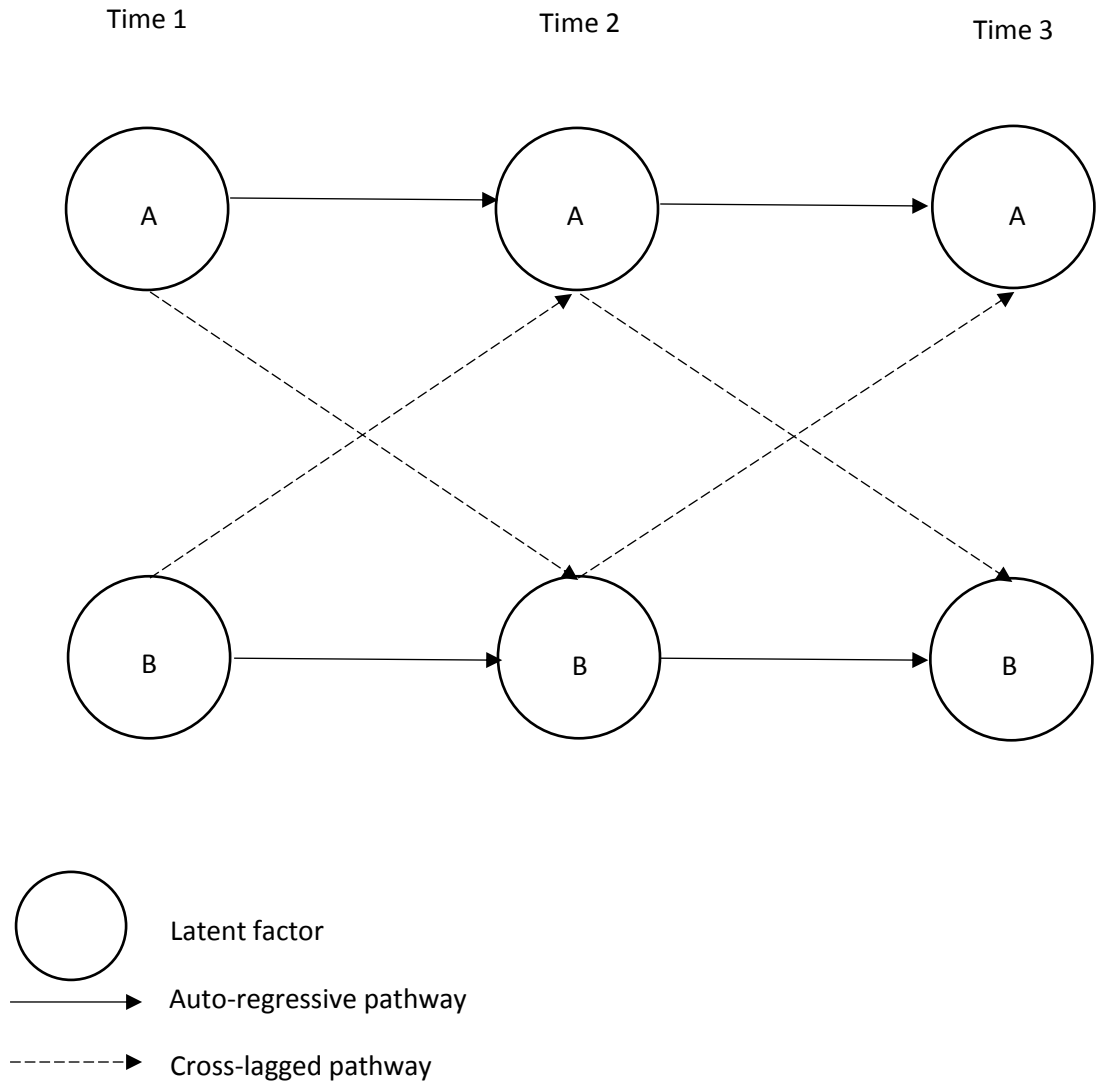
Psychometric Characteristics
Parcels, compared to items, have:
Higher reliability
Greater communality (amount of variance in observed variable accounted for by the factor)
Higher ratio of common to unique factor variance
Lower likelihood of distributional violations
More, tighter, and more equal intervals
Model estimation and fit characteristics
Models with parcels, compared to items have:
Fewer parameter estimates
Lower indicator to subject ratio
Lower likelihood of correlated residuals and dual factor loadings
Reduced sources of sampling error
Easier estimation

[150]

Once the nature of the measurement model has been decided then the form of the structural model can be considered. In order to test for a potential bi-directional relationship between two constructs measured repeatedly over time, auto-regressive cross-lagged SEM can be used. The auto-regressive part of the model refers to a variable being regressed on its previous value (i.e. depression at time point 2 is regressed on depression at time point 1), the cross-lagged part of the model refers to a variable being regressed on a different variable from the previous time point (i.e. depression at time point 2 is regressed on obesity at time point 1). An auto-regressive cross-lagged SEM makes use of the intrinsic time ordered nature of the panel data available in cohort studies to address questions of the causal ordering of variables [151]. In this type of model the two latent variables of interest are defined

by indicators at each measurement occasion and these latent variables are then both regressed on their own lagged variable and the lagged score of the other variable at the previous measurement occasion (Figure 3.4). The model then provides an estimate of the effect of each variable of interest on the other. This allows investigation into whether there is a bi-directional relationship between the two latent factors and if cross-lagged effects in one direction are of a different magnitude to those in the other direction (i.e. if depression at time point 1 has a greater influence on obesity at time point 2, than obesity at time point 1 has on depression at time point 2).

Figure 3.4 – Simplified diagrammatic representation of an auto-regressive cross-lagged structural equation model



In this analysis to create a latent factor for obesity at each follow-up occasion all measures of obesity (e.g. BMI, DXA, waist circumference) were used (where available) as factor indicators to define the latent obesity trait. The items from the depression questionnaires were grouped into item parcels and these parcels were used as factor indicators to define the latent depression construct. In ALSPAC the 13-item SMFQ instrument was used to measure depression[115]. The 13 items were randomly split into three parcels (as three factor indicators are required for model identification [152, 153]); one parcel of five items and two parcels of four items. Once this first random assignment of items to parcels was carried out at the first measurement occasion, the same items were used to produce the item parcels at the other follow up occasions. Similarly the 13 items of the YSR APS [116] from the TRAILS cohort were separated into three parcels; one parcel of five items and two parcels of four items. In the NDIT cohort the six item KDDS [117] was separated into three parcels of two items each. The obesity and depression latent traits at the different waves of follow up were then analysed in an auto-regressive cross-lagged SEM, allowing investigation into whether there was a bi-directional relationship between obesity and depression; i.e. whether those who have a greater level of obesity have an increase in subsequent symptoms of depression and/or whether those with more depressive symptoms are more likely to have a subsequent increase in the level of obesity. Where appropriate results were meta-analysed to produce a single pooled result.

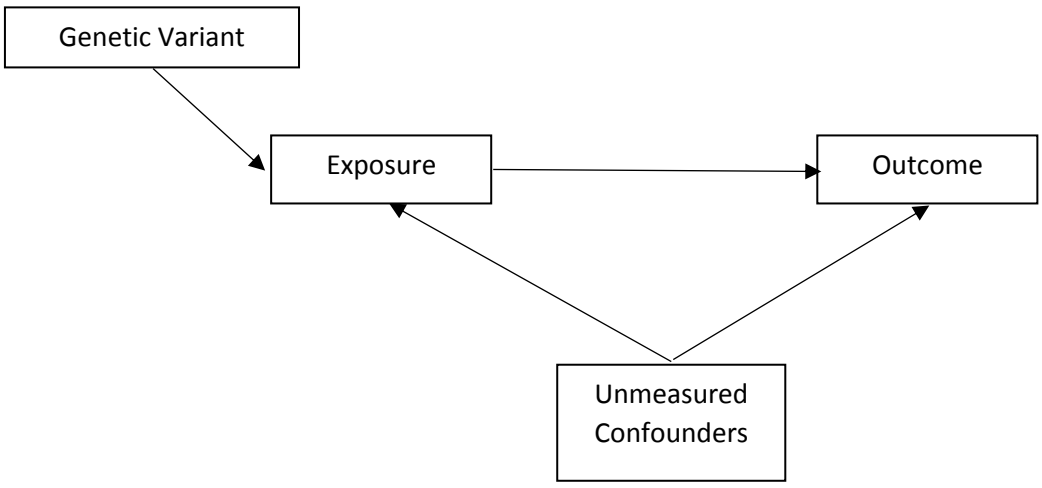
Mendelian Randomisation

In observational studies an association between an outcome and an exposure cannot definitively be shown to be causal due to the problems of reverse causation (where the outcome variable may temporally precede and have an effect on the exposure) and residual confounding (bias that remains after controlling for confounders due to additional confounding factors that were not controlled for and/or poor measurement of confounding factors that were adjusted for – measured variable does not well represent the confounder or is collected with a large amount of error).

Mendelian Randomization is a method which utilises an instrumental variable approach, whereby genetic variants are used as proxies for measured exposure variables when investigating the association between an exposure and outcome [60]. This method of MR analysis, can be thought of as analogous to a randomized trial where instead of participants being randomized to different interventions, the randomization is to different genotypes. As this randomization happens at the time of conception, it cannot be susceptible to the problems of reverse causation or confounding (as standard epidemiological techniques are) and therefore is useful in the investigation of observational associations to strengthen the evidence for a causal association. Hence this approach was also used to investigate the relationship between obesity and depression.

However, for an MR analysis to provide an unbiased estimate of a causal effect, there are three key assumptions [60]; 1) the genetic variant is associated with the exposure variable, 2) the genetic variant is associated with the outcome variable only through its association with the exposure, and 3) the genetic variant is not associated with (unmeasured) confounders. It is only possible to test assumption 1 using observed data.

Figure 3.5 - Diagrammatic representation of Mendelian Randomization approach



A weighted allelic score was generated for use as a genetic instrumental variable, using independent genetic variants that have been shown (in a large GWAS meta-analysis study) to be robustly associated with BMI [129]. At each genetic locus the number of risk alleles was multiplied by the size of the effect of that risk allele, then divided by the mean effect size of all the variants and summed to produce a weighted genetic risk score [129]. The use of a genetic score variable rather than a single genetic variant is a superior approach [129] as it explains a larger proportion of the variation in BMI, captures a greater range of adiposity and increases specificity to the adiposity trait, hence increasing statistical power.

For the MR analyses, two-stage least squares (2SLS) regression was performed utilising the genetic risk score (generated as described above) as an instrumental variable for obesity utilising the "*ivreg2*" command in STATA 14 [149]. 2SLS regression is named so because it consists of two consecutive ordinary least squares (OLS) regressions. The first OLS regression is the regression of the exposure (in this case BMI) on the instrumental variable (here the weighted allele score) to produce an estimator of the exposure. The second OLS regression is the regression of the outcome (here depression score) on the estimator of the exposure. The F-statistic was examined from the first-stage regression (i.e. the regression between the weighted allele score and BMI) to investigate the MR assumption of association between the instrumental variable and the exposure variable, with the aim of assessing potential weak instrument bias (bias toward the observational association)[154]. The F-statistic can be defined as the ratio of the mean square of the model to the mean square of the error and is a measure of the "strength" of an instrumental variable. It is generally accepted within the MR literature that a F-statistic value of less than 10 is considered to be a "weak" instrument [155]. A potential issue with MR analysis is that of pleiotropy; the genetic instrument may affect the outcome through a different biological pathway than the one being investigated. Pleiotropy may be investigated using Egger regression [156].

3.7.2. Objective 2 – physical activity and depression

The analytical approach employed in Objective 2 of the project is largely similar to that used in Objective 1; moving from simple linear regression through to more complex statistical modelling techniques utilising the longitudinal repeated measures nature of the data available (GEE and SEM) and MR to try and deal with the potential problem of residual confounding. An additional analytical approach used in Objective 2 that was not used in Objective 1 is partial least squares regression (PLS-R). PLS-R was used to try to identify what aspects of physical activity may be important in the association with adolescent depression.

Linear Regression

Linear regression was used to examine the effect of PA (exposure) on depressive symptoms score (Z-score) (outcome) at the next follow up occasion for each of the three cohorts, adjusting for relevant confounders. The linear regression model was also carried out including an interaction term between PA and sex and separately for males and females.

ALSPAC

The time points at which measures of depression and PA were collected in the ALSPAC cohort are outlined in Table 3.8 below.

Table 3.8 - Physical Activity and Depression data collected in ALSPAC

		Data Collection Wave												
		F10	F11	TF1	TF2	CCQ	PUB6	PUB7	TF3	PUB8	CCS	PUB9	TF4	CCT
		10y 7m	11y 6m	12y 10m	13y 10m	14y 0m	14y 7m	15y 5m	15y 6m	16y 0m	16y 6m	17y 0m	17y 10m	18y 6m
Depression:	SMFQ	X		X	X						X		X	X
Physical Activity:	Accelerometer		X		X				X					
	Self Report Questionnaire					X	X	X		X	X	X		X

From Table 3.8 it can be seen that regression models with PA measured using accelerometery as the exposure and depression measured on the SMFQ as the outcome can be carried out at with the exposure measured at 13y 10m and outcome at 16y 6m. Regression models using the self-report measure of PA as the exposure were fitted with:

(1) exposure measured at 14y 0m and outcome at 16y 6m (adjusted for depression measured at 13y 10m)

(2) exposure was measured at 16y 6m and the outcome at 17y 10m (adjusted for depression measured at 16y 6m)

For each of these analyses, four regression models were fitted:

- Model 1 – PA as exposure on SMFQ depression Z-score outcome, adjusted for age, sex, previous depression score, maternal education, maternal social class and maternal depression.
- Model 2 – model 1 plus PA*sex interaction term.
- Model 3 – model 1 restricted to males and without sex as a confounder.
- Model 4 – model 1 restricted to females and without sex as a confounder.

TRAILS

The time points at which measures of depression and PA were collected in the TRAILS cohort are outlined in Table 3.9 below.

Table 3.9 - Physical Activity and Depression data collected in TRAILS

		Data Collection Wave			
		T1	T2	T3	T4
		10y 7m	13y 1m	15y 9m	18y 7m
Depression:	YSR	X	X	X	X
Physical Activity:	Self Report Questionnaire	X	X	X	X

From Table 3.9 it can be seen that regression models investigating level of PA on level of depression at next follow up can be analysed from age 10y 7m to 13y 1m, from 13y 1m to 15y 9m and from 15y 9m to 18y 7m. For each of these regression analyses, four models were fitted:

- Model 1 – PA as exposure on YSR APS depression Z-score outcome, adjusted for age, sex, previous depression score, social class, maternal depression, cigarette smoking and alcohol use.
- Model 2 – model 1 plus PA*sex interaction term.
- Model 3 – model 1 restricted to males and without sex as a confounder.
- Model 4 – model 1 restricted to females and without sex as a confounder.

NDIT

The time points at which measures of depression and PA were collected in the NDIT cohort are outlined in Table 3.10 below. For each of these regression analyses, four models were fitted:

- Model 1 – PA as exposure on Kandel depression Z-score outcome, adjusted for age, sex, previous depression score, social class, maternal education, maternal profession and participant alcohol consumption.
- Model 2 – model 1 plus PA*sex interaction term.
- Model 3 – model 1 restricted to males and without sex as a confounder.
- Model 4 – model 1 restricted to females and without sex as a confounder.

Table 3.10 - Physical Activity and Depression data collected in NDIT

		Data Collection Wave									
		1	2	3	4	5	6	7	8	9	10
		12y 9m	13y 0m	13y 2m	13y 8m	13y 10m	14y 1m	14y 2m	14y 7m	14y 10m	15y 0m
Depression:	Kandel	X	X	X	X	X	X	X	X	X	X
Physical Activity:	Time Spending Patterns Questionnaire	X	X	X	X	X	X	X	X	X	X

Table continued:

		Data Collection Wave									
		11	12	13	14	15	16	17	18	19	20
		15y 2m	15y 7m	15y 10m	16y 0m	16y 2m	16y 6m	16y 9m	17y 0m	17y 1m	20y 1m
Depression:	Kandel	X	X	X	X	X	X	X	X	X	X
Physical Activity:	Time Spending Patterns Questionnaire	X	X	X	X	X	X	X	X	X	X

Generalised Estimating Equations

GEE analyses were carried out as described above in section 3.7.1 but using lagged PA as the exposure (rather than lagged BMI) and adjusting for baseline PA (rather than baseline BMI). The GEE models were carried out within each cohort using lagged PA as the exposure variable with regression coefficients, 95% confidence intervals and p values being reported for each model. All models were adjusted for baseline PA, baseline depression, lagged depression, age, gender, maternal depression and socio-economic status. The GEE analysis within the NDIT cohort was additionally adjusted for concurrent alcohol use and the analysis within the TRAILS cohort was additionally adjusted for concurrent alcohol use and smoking. The analyses were repeated including an interaction term between lagged PA and sex to formally test for a difference in the association between PA and depression in boys and girls, the analysis was then stratified by sex.

In the ALSPAC cohort four GEE models were fitted:

- Model 1 – Lagged self-report PA as exposure on SMFQ depression Z-score outcome, adjusted for baseline PA, baseline depression score, lagged depression score, age, sex, maternal education, maternal social class and maternal depression.
- Model 2 – model 1 plus lagged self-report PA*sex interaction term.
- Model 3 – model 1 restricted to males and without sex as a confounder.
- Model 4 – model 1 restricted to females and without sex as a confounder.

Only the ALSPAC self-report PA data could be used in the GEE analysis, the ALSPAC accelerometer data was not collected at enough appropriate time points for use in GEE.

In the TRAILS GEE analysis there was additional adjustment for alcohol and smoking use. Four GEE models were fitted:

- Model 1 – lagged PA as exposure on YSR APS depression Z-score outcome, adjusted for baseline PA, baseline depression score, lagged depression score, age, sex, social class, maternal depression, cigarette smoking and alcohol use.
- Model 2 – model 1 plus lagged PA*sex interaction term.
- Model 3 – model 1 restricted to males and without sex as a confounder.
- Model 4 – model 1 restricted to females and without sex as a confounder.

In the NDIT cohort the following four GEE models were fitted:

- Model 1 – Lagged self-report PA as exposure on Kandell depression Z-score outcome, adjusted for baseline PA, baseline depression score, lagged depression score, age, sex, maternal education, maternal social class, maternal depression and alcohol use.
- Model 2 – model 1 plus lagged PA*sex interaction term.
- Model 3 – model 1 restricted to males and without sex as a confounder.
- Model 4 – model 1 restricted to females and without sex as a confounder.

The results of the GEE analyses from each of the three cohorts were pooled to produce a single result using fixed effects inverse variance weighted meta-analysis (STATA 14's *metan* command) [149].

Partial Least Squares Regression

Partial least squares regression (PLS-R) [157] was used to identify which aspects of PA (e.g. frequency vs. intensity) or types (e.g. team vs. individual exercise) may be important in relation to adolescent depression. Self-report physical activity questionnaire items that ask participants about a large range of activities, that normally would not be able to be entered into the same standard regression model due to collinearity, can all be included in one PLS-R model. We can then observe if

these items are grouped into components that can be seen to represent different aspects of physical activity and see how much of the variance in depression is explained by these aspects. Hence this approach may enable us to disentangle what aspects of PA may be important in adolescent depression.

PLS-R is a method that combines aspects of principal components analysis and multiple regression [157]. The aim of PLS-R, as in standard multiple regression, is to predict an outcome from a set of exposure variables. However if the number of exposure variables is large and the exposure variables are highly collinear then trying to carry out a standard regression model is not possible due to problems of multicollinearity and over-fitting. A way to solve this problem is to perform a dimension reduction technique to reduce the number of predictors in the model.

One of the most common dimension reduction methods is principal components analysis (PCA). In PCA, factors are extracted from the exposure variables (hence avoiding the problem of collinearity) in order of the proportion of variance in the exposure that they explain. These factors may then be used as regressors on the outcome variable. The problem with this technique is that the extracted components have been selected to explain the exposure not the outcome and as such may be completely irrelevant for the outcome variable; in PCA regression a component that only explains a small amount of variance in the exposure but is highly related to the outcome would not be retained in the final model. In PLS-R components are extracted that model the exposure and simultaneously predict the outcome. Orthogonal (uncorrelated) factors are extracted (from the exposures) in order of the proportion of the variance they explain in both the exposure and in the outcome, thereby producing components that explain as much of the covariance between the exposures and outcome as possible. This is followed by a regression step where the components are used to predict the outcome.

The first aspect of PLS-R was to choose the appropriate number of components to retain in the model. The PLS-R is carried out containing a large number of components (e.g. 10). The actual number of components that are sufficient to retain in the model can then be judged based on the percentage of variance in the exposure and outcome explained by the components and by the size of the root mean squared error of prediction (RMSEP).

Ideally the model would explain a large amount of the variance in both the exposures and the outcome; the more components that are retained the greater the amount of this variance that will be explained. Components should be retained until the addition of further components only increases the amount of variance explained by a small amount. Using the amount of variance explained by the components on its own to decide the total number of components to retain may prove difficult (after all, what is "a small increase" in explained variance?) and is likely to suggest retaining a range of components (for example retaining 6 to 8 components). Hence the additional use of the RMSEP. The RMSEP represents the difference between values predicted by a model and observed in the data, as such we are looking for the model with the number of components that produces the smallest RMSEP value. Once the number of components to retain has been decided the PLS-R model can then be re-run retaining only that number of components in the model.

The ALSPAC and TRAILS PA self-report questionnaire data both comprise only one question relating to total time spent in PA, and as such are not suitable for PLS-R modelling. However, in the NDIT cohort, the self-report PA data was collected using the time spending patterns questionnaire which asks participants about time spent in a large variety of different activities. Therefore the PLS-R modelling approach was carried out in the NDIT cohort.

Structural Equation Modelling

Auto-regressive cross-lagged SEM was used as described above in section 3.7.1 to examine the association between PA and obesity, with PA variables used as indicators to define a PA latent factor in the place of an obesity latent factor. This allowed for the investigation into a possible bi-directional relationship between PA and depression in adolescence.

Mendelian Randomization

To date, there is no robust evidence of association between physical activity and any genetic variants. A genome wide prediction score may be used as a genetic instrument in MR analysis in place of single genetic variant or allele score [158]. When a GWAS is used to try to identify genetic variants that are associated with a trait a very stringent p-value is applied to assess these potential associations (due to large amounts of multiple testing) and as such "true" associations may be dismissed if they do not meet the criteria of this very stringent p-value. In an attempt to recover some of this potentially lost information, a genome wide prediction score uses the genome wide variation of a trait aggregated into a single score as a genetic instrument for MR analysis. This approach was used to explore the relationship between PA and depression in the ALSPAC cohort, the only cohort with PA genetic data available.

A split sample genome wide prediction score MR analysis was carried out using the ALSPAC data. First randomization of the ALSPAC participants into two sub-groups was carried out. Next, using the genetic and PA data from sub-group one, genetic variants were extracted if in a GWAS study for PA, they produced a p-value ≤ 0.1 [159]. The number of risk alleles across the genetic variants were then summed, with each one weighted by multiplying by the effect size from the PA GWAS to produce a prediction score. This prediction score was then applied to the individuals in the other sub-group and used as the instrumental variable in 2-stage least squares

regression. The analysis was then repeated using prediction scores produced from sub-group two applied to the first group as an instrumental variable. The results from the MR analyses of the two sub-groups were then pooled using a fixed effects inverse variance weighted meta-analysis using STATA's *metan* command.

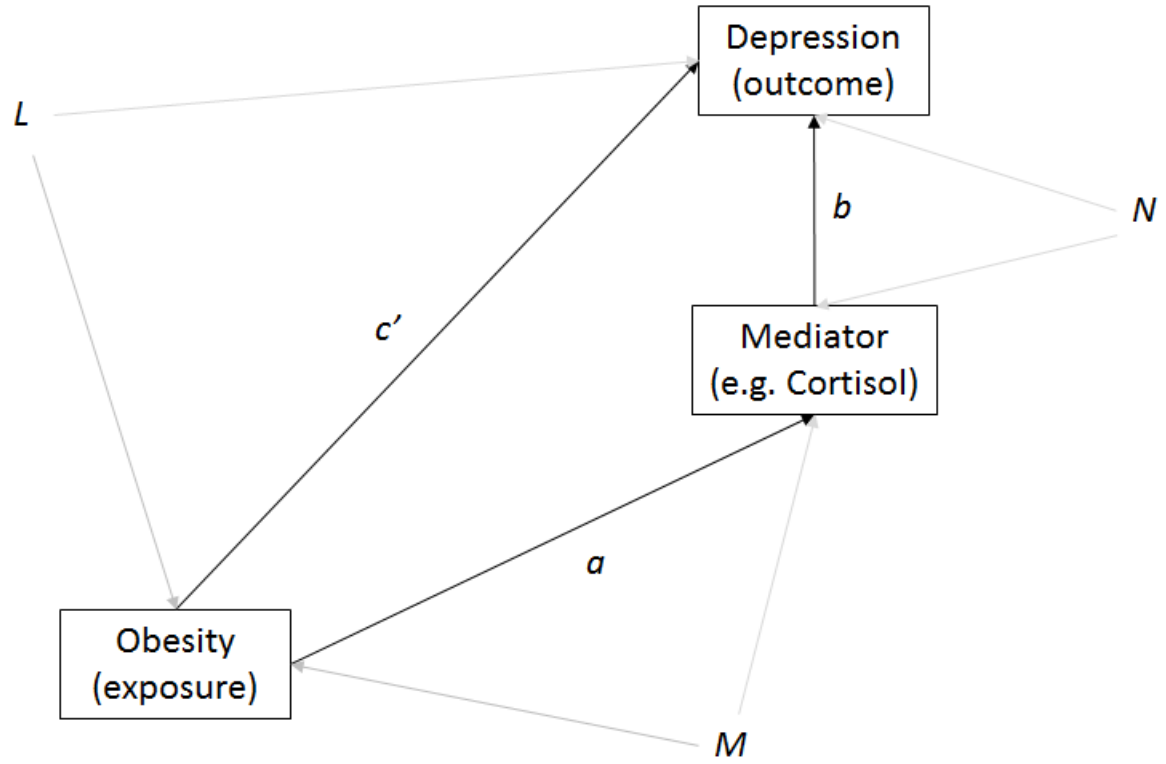
3.7.3. Objective 3 – mediation analyses

In order to identify novel intervention targets on the causal pathway between obesity and depression in adolescence the cross-lagged SEM analyses as outlined in section 3.7.1 were extended into mediation analyses. A mediation analysis is concerned with the extent to which an intermediate variable (known as a mediator) explains the effect of an exposure variable on an outcome. The mediators that were investigated were body image (ALSPAC and TRAILS cohorts), self-esteem (TRAILS cohort), C-reactive protein (ALSPAC and TRAILS cohorts), cortisol (TRAILS cohort) and IgE (TRAILS cohort) (See section 3.5 for details regarding the mediators investigated).

In order to examine the role of each of the potential mediators, an intermediate variable (the mediator) was included in the SEM model (at the time point at which the mediator was measured) between the obesity (exposure) and depression (outcome) latent traits. A diagrammatic example of this can be seen in Figure 3.6; in this model c' represents the direct effect of obesity on depression (adjusted for relevant confounders L), the coefficient a represents the effect of obesity on the mediator of interest (adjusted for confounders M), and b the effect of the mediator on depression (adjusted for confounders N). The indirect effect of obesity on depression through the mediator was calculated as the product ab. As such $ab + c'$ can be thought of as the total effect of obesity on depression. In other words the total effect represents the combination of the direct effect - representing the effect of obesity on depression independent of the pathway through the mediator, and the indirect effect - representing the difference in depression caused by the effect that a one unit

increase in obesity has on a mediator, which then in turn affects the depression outcome (assuming that both the outcome variable and the mediator variable are continuous – as they are in this investigation). The direct and indirect effects are conditional on the model, using each different mediator will result in a different direct and indirect effect. The total effect of the exposure variable (obesity) on the outcome (depression) will however remain the same.

Figure 3.6 Diagrammatic representation of mediation analysis



In the present analysis bootstrapping was used to estimate the confidence intervals for the indirect effect (*ab*). Using bootstrapping a distribution for the indirect effect is generated by treating the obtained result as a representation sample of size *n* of the population as a whole [84]. The data is re-sampled with replacement and the coefficients *a* and *b* are estimated from the re-sampled data set and the product *ab* is

recorded. This process is repeated *k* number of times (in this analysis *k* was 1000) resulting in *k* estimates of the indirect effect, these *k* estimates approximate a distribution of the indirect effect when taking a sample of size *n* from the original population. The distribution of the indirect effects can be used to estimate a confidence interval; the k indirect effects are ordered from smallest to largest, the lower bound of a 95% confidence interval is defined as $k(0.5 - 95/200)$, the upper bound of a 95% confidence interval is defined as $1 + k(0.5 + 95/200)$. So for example if k=1000, then the lower bound would be the value in the 25^{th} position as $1000(0.5 - 95/200)=25$, and the upper bound would be the value in the 976^{th} position as $1 + 1000(0.5 + 95/200)=976$. The mediation analyses were carried out on the entire cohort with available data (males and females combined) and separately by sex. The analysis will provide both estimates of the direct effect of obesity on depression, and an estimate of the indirect effect of obesity on depression via the tested mediator.

CHAPTER 4. RESULTS – COHORT DESCRIPTION

4.1. ALSPAC

As discussed in Section 3.1.1, ALSPAC is a population based prospective birth cohort designed to investigate the influences of genetic, biological, psychological, social and other environmental factors on physical and mental health, behaviour and development. At the first measurement occasion that was used in this study there were a total of 7457 participants (Figure 4.1); 3680 (49%) males and 3777 (51%) females (Table 4.1). The mean age of the participants at this time point was 10 years 8 months, participants were then followed approximately annually for the next three years (mean age at next three follow up occasions; 11 years 9 months, 12 years 10 months and 13 years 10 months). Subsequent to these annual follow-ups, the next follow up occasion was almost two years later (approximately 20 months) when participants were a mean age of 15 years 6 months. After this, participants were once again followed up approximately annually for the next three years (mean age at next three follow up occasions; 16 years 8 months, 17 years 10 months and 18 years 8 months) (Table 4.2). Data on maternal depression, maternal level of education and social class were also collected (at 32 weeks gestation) (Table 4.1).

Figure 4.1 Flow chart of participant retention in ALSPAC cohort

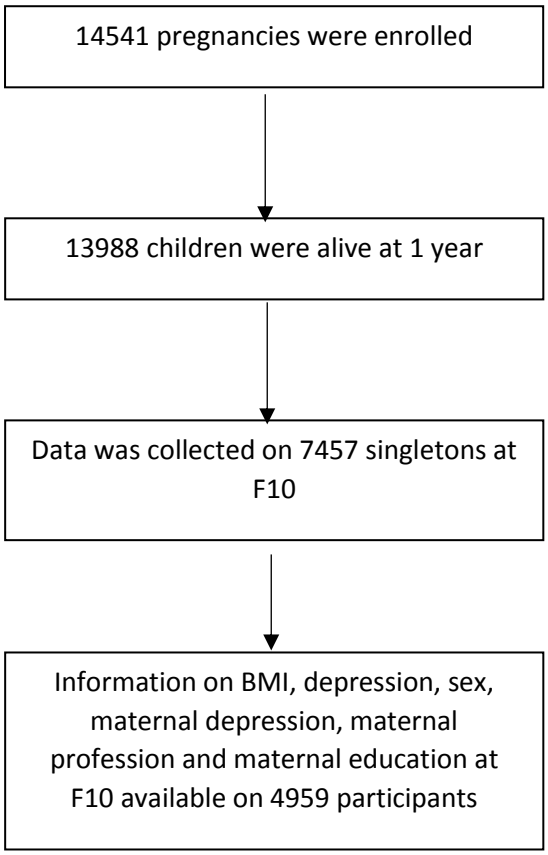


Table 4.1 – Time invariant sociodemographic characteristics of ALSPAC participants

Variable	n	%	Mean (SD)	Median (IQR)
Sex:	7457			
Male	3680	49%		
Female	3777	51%		
Maternal depression:				
EPDS	6018		6 (5.0)	5 (2, 9)
Maternal education:	6811			
CSE	922	14%		
Vocational	584	9%		
O level	2412	35%		
A level	1812	27%		
Degree	1081	16%		
Maternal social class:	5878			
I	410	7%		
II	2047	35%		
III (non-manual)	2474	42%		
III (manual)	398	7%		
IV	463	8%		
V	85	1%		

Table 4.2 – Age of ALSPAC participants at each wave of follow up

Time point	n	Age	
		Mean (SD)	Median (IQR)
F10	7457	10y8m (3.2m)	10y7m (10y6m, 10y9m)
F@11	7060	11y9m (2.9m)	11y9m (11y7m, 11y10m)
TF1	6745	12y10m (2.8m)	12y10m (12y8m, 12y11m)
TF2	6062	13y10m (2.5m)	13y10m (13y9m, 13y11m)
TF3	5441	15y6m (4.2m)	15y5m (15y3m, 15y7m)
CCS	5079	16y8m (2.8m)	16y7m (16y6m, 16y10m)
TF4	5164	17y10m (5.4m)	17y9m (17y7m, 17y11m)
CCT	3343	18y8m (5.9m)	18y8m (18y3m, 19y1m)

Various anthropometric variables were collected sporadically across the eight waves of follow up used in this project: height and weight (and therefore BMI) were measured at six of the eight follow up occasions, waist circumference at five time points and DXA fat percentage at four (Table 4.3). Mean height and weight increased steadily across the time points and were very similar between males and females for the first four measurement occasions. At the last two follow ups males were taller and heavier compared with females (Table 4.3). There was little difference in waist circumference between males and females at any time point, however, females consistently had a greater body fat percentage than males (Table 4.3).

Table 4.3 - Descriptive statistics for anthropometric measurements in the ALSPAC cohort

Variable	F10 (10y8m)			F@11 (11y9m)			TF1 (12y10m)		
	n	Mean (SD)	Median (IQR)	n	Mean (SD)	Median (IQR)	n	Mean (SD)	Median (IQR)
Height (cm)	7392	144 (6.7)	144 (140, 148)	7019	151 (7.3)	151 (146, 156)	6693	157 (7.7)	157 (152, 162)
Weight (kg)	7418	38 (8.6)	36 (32, 43)	7022	44 (10.2)	42 (36, 49)	6622	49 (11.0)	48 (42, 55)
BMI	7374	18 (3.2)	18 (16, 20)	7014	19 (3.5)	18 (17, 21)	6622	20 (3.5)	19 (17, 22)
DXA fat percentage	*NA*	*NA*	*NA*	6912	26 (9.5)	25 (18, 33)	*NA*	*NA*	*NA*
Waist Cirumference (cm)	7418	66 (8.7)	64 (59, 70)	7017	68 (9.5)	66 (62, 73)	6638	71 (9.6)	69 (64, 75)
Males									
Height (cm)	3649	144 (6.4)	144 (140, 148)	3450	150 (7.1)	152 (146, 156)	3278	157 (8.3)	157 (151, 162)
Weight (kg)	3663	38 (8.2)	36 (32, 42)	3454	43 (9.8)	41 (36, 48)	3257	49 (11.1)	47 (41, 55)
BMI	3644	18 (3.0)	17 (16, 20)	3450	19 (3.3)	18 (16, 21)	3257	20 (3.5)	19 (17, 21)
DXA fat percentage	*NA*	*NA*	*NA*	3392	23 (9.6)	21 (16, 30)	*NA*	*NA*	*NA*
Waist Cirumference (cm)	3657	66 (8.7)	64 (60, 70)	3450	69 (9.7)	66 (62, 73)	3259	71 (9.9)	69 (64, 76)
Females									
Height (cm)	3743	144 (7.0)	144 (139, 149)	3569	151 (7.3)	152 (146, 156)	3415	158 (6.9)	158 (153, 162)
Weight (kg)	3755	39 (9.0)	37 (32, 43)	3568	45 (10.5)	43 (37, 51)	3365	50 (10.8)	49 (43, 56)
BMI	3730	18 (3.3)	18 (16, 20)	3564	19 (3.6)	19 (17, 21)	3365	20 (3.6)	19 (18, 22)
DXA fat percentage	*NA*	*NA*	*NA*	3520	28 (8.7)	27 (21, 35)	*NA*	*NA*	*NA*
Waist Cirumference (cm)	3761	65 (8.8)	64 (59, 70)	3567	68 (9.3)	66 (62, 73	3379	70 (9.3)	69 (64, 75)

Table 4.3 continued

Variable	TF2 (13y10m) n	TF2 (13y10m) Mean (SD)	TF2 (13y10m) Median (IQR)	TF3 (15y6m) n	TF3 (15y6m) Mean (SD)	TF3 (15y6m) Median (IQR)	CCS (16y8m) n	CCS (16y8m) Mean (SD)	CCS (16y8m) Median (IQR)
Height (cm)	6053	163 (7.8)	163 (158, 168)	5366	169 (8.4)	169 (163, 175)	*NA*	*NA*	*NA*
Weight (kg)	6045	55 (11.3)	53 (47, 61)	5352	62 (11.9)	60 (54, 67)	*NA*	*NA*	*NA*
BMI	6045	20 (3.5)	20 (18, 22)	5352	21 (3.6)	21 (19, 23)	*NA*	*NA*	*NA*
DXA fat percentage	5955	24 (10.3)	24 (16, 32)	5090	24 (11.2)	25 (15, 33)	*NA*	*NA*	*NA*
Waist Cirumference (cm)	6040	72 (9.3)	70 (66, 76)	4414	77 (9.0)	75 (71, 81)	*NA*	*NA*	*NA*
Males									
Height (cm)	2975	165 (8.8)	165 (159, 171)	2543	174 (7.5)	175 (170, 179)	*NA*	*NA*	*NA*
Weight (kg)	2975	55 (11.9)	53 (46, 61)	2539	64 (12.3)	62 (56, 70)	*NA*	*NA*	*NA*
BMI	2975	20 (3.4)	19 (18, 22)	2539	21 (3.4)	20 (19, 22)	*NA*	*NA*	*NA*
DXA fat percentage	2922	19 (9.8)	17 (12, 26)	2426	17 (9.1)	14 (10, 21)	*NA*	*NA*	*NA*
Waist Cirumference (cm)	2969	73 (9.6)	70 (66, 76)	2033	77 (9.0)	75 (71, 81)	*NA*	*NA*	*NA*
Females									
Height (cm)	3078	162 (6.3)	162 (158, 166)	2823	165 (6.1)	165 (161, 169)	*NA*	*NA*	*NA*
Weight (kg)	3070	55 (10.8)	53 (48, 60)	2813	59 (11.1)	57 (52, 64)	*NA*	*NA*	*NA*
BMI	3070	21 (3.6)	20 (18, 22)	2813	22 (3.7)	21 (19, 23)	*NA*	*NA*	*NA*
DXA fat percentage	3033	29 (8.3)	29 (23, 35)	2664	31 (8.1)	31 (26, 37)	*NA*	*NA*	*NA*
Waist Cirumference (cm)	3071	72 (9.0)	70 (65, 76)	2381	77 (9.0)	75 (70, 82)	*NA*	*NA*	*NA*

Table 4.3 continued

Variable	TF4 (17y10m) n	TF4 (17y10m) Mean (SD)	TF4 (17y10m) Median (IQR)	CCT (18y8m) n	CCT (18y8m) Mean (SD)	CCT (18y8m) Median (IQR)
Height (cm)	5018	171 (9.3)	170 (164, 178)			
Weight (kg)	5017	67 (14)	65 (57, 74)	*NA*	*NA*	*NA*
BMI	5012	23 (4.2)	22 (20, 25)	*NA*	*NA*	*NA*
DXA fat percentage	4806	27 (11.7)	27 (17, 35)	*NA*	*NA*	*NA*
Waist Cirumference (cm)	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
Males						
Height (cm)	2197	179 (6.7)	178 (174, 183)	*NA*	*NA*	*NA*
Weight (kg)	2198	72 (13.4)	70 (63, 79)	*NA*	*NA*	*NA*
BMI	2195	23 (3.9)	22 (20, 24)	*NA*	*NA*	*NA*
DXA fat percentage	2123	18 (9.4)	16 (11, 23)	*NA*	*NA*	*NA*
Waist Cirumference (cm)	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
Females						
Height (cm)	2818	165 (6.2)	165 (161, 169)	*NA*	*NA*	*NA*
Weight (kg)	2816	63 (13.0)	61 (55, 69)	*NA*	*NA*	*NA*
BMI	2814	23 (4.5)	22 (20, 24)	*NA*	*NA*	*NA*
DXA fat percentage	2681	34 (8.4)	33 (28, 29)	*NA*	*NA*	*NA*
Waist Cirumference (cm)	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*

Data on PA were also collected sporadically in the ALSPAC cohort. Accelerometer data were collected at three of the eight time points whilst self-report PA questionnaire data were collected at four of the time points. There was a general decrease in total amount of PA carried out as participants got older (Table 4.4), with males (Table 4.5) consistently carrying out more PA than females (Table 4.6).

Depression data were collected at six of the eight time points. There was a steady increase in mean depressive symptom score over time and females had a higher mean score than males (Table 4.7).

Table 4.4 – Descriptive statistics for physical activity measures collected in the ALSPAC cohort

Variable	F@11 (11y9m)				TF2 (13y10m)			
	n	%	Mean (SD)	Median (IQR)	n	%	Mean (SD)	Median (IQR)
Total daily minutes of PA	5529		351 (65)	348 (305, 393)	3759		304 (65)	299 (258, 345)
Accelerometer counts per minute	5529		604 (178)	580 (474, 710)	3759		539 (182)	511 (407, 644)
Daily minutes of MVPA	5529		23 (15)	20 (12, 31)	3759		24 (17)	21 (12, 32)
Percentage of time spent in MVPA	5529		3 (1.9)	3 (2, 4)	3759		3 (2.1)	3 (1, 4)
At least 1 hour of MVPA a day	144	3%			158	4%		
Self report frequency of PA in past year:					6055			
Never	NA	NA	NA	NA	139	2%		
Less than once a month	NA	NA	NA	NA	105	2%		
1 - 3 times a month	NA	NA	NA	NA	418	7%		
1 - 4 times a week	NA	NA	NA	NA	3168	52%		
5 or more times a week	NA	NA	NA	NA	2225	37%		

Table 4.4 continued

Variable	TF3 (15y6m)				CCS (16y8m)			
	n	%	Mean (SD)	Median (IQR)	n	%	Mean (SD)	Median (IQR)
Total daily minutes of PA	2017		270 (62)	263 (226, 308)	NA		NA	NA
Accelerometer counts per minute	2017		479 (165)	447 (361, 573)	NA		NA	NA
Daily minutes of MVPA	2017		24 (18)	19 (10, 34)	NA		NA	NA
Percentage of time spent in MVPA	2017		3 (2.2)	2 (1, 4)	NA		NA	NA
At least 1 hour of MVPA a day	89	4%			NA	NA		
Self report frequency of PA in past year:	4751				4996			
Never	112	2%			245	5%		
Less than once a month	397	8%			288	6%		
1 - 3 times a month	2261	48%			682	14%		
1 - 4 times a week	1238	26%			2528	51%		
5 or more times a week	743	16%			1253	25%		

Table 4.4 continued

Variable	CCT (18y8m)			
	n	%	Mean (SD)	Median (IQR)
Total daily minutes of PA	NA		NA	NA
Accelerometer counts per minute	NA		NA	NA
Daily minutes of MVPA	NA		NA	NA
Percentage of time spent in MVPA	NA		NA	NA
At least 1 hour of MVPA a day	NA	NA		
Self report frequency of PA in past year:	3260			
Never	222	7%		
Less than once a month	268	8%		
1 - 3 times a month	561	17%		
1 - 4 times a week	1670	51%		
5 or more times a week	539	17%		

Table 4.5 – Descriptive statistics for physical activity measures collected in males in the ALSPAC cohort

Variable	F@11 (11y9m)				TF2 (13y10m)			
	n	%	Mean (SD)	Median (IQR)	n	%	Mean (SD)	Median (IQR)
Total daily minutes of PA	2630		364 (66)	361 (318, 408)	1776		322 (67)	319 (277, 365)
Accelerometer counts per minute	2630		662 (186)	644 (528, 772)	1776		596 (192)	571 (456, 704)
Daily minutes of MVPA	2630		28 (17)	25 (16, 38)	1776		29 (18)	25 (15, 38)
Percentage of time spent in MVPA	2630		4 (2.1)	3 (2, 5)	1776		4 (2.3)	3 (2, 5)
At least 1 hour of MVPA a day	131	5%			122	7%		
Self report frequency of PA in past year:					2687			
Never	NA	NA	NA	NA	92	3%		
Less than once a month	NA	NA	NA	NA	50	2%		
1 - 3 times a month	NA	NA	NA	NA	177	7%		
1 - 4 times a week	NA	NA	NA	NA	1171	44%		
5 or more times a week	NA	NA	NA	NA	1197	45%		

Table 4.5 continued

Variable	TF3 (15y6m)				CCS (16y8m)			
	n	%	Mean (SD)	Median (IQR)	n	%	Mean (SD)	Median (IQR)
Total daily minutes of PA	897		286 (66)	282 (238, 326)	NA		NA	NA
Accelerometer counts per minute	897		533 (178)	509 (405, 643)	NA		NA	NA
Daily minutes of MVPA	897		30 (19)	25 (16, 41)	NA		NA	NA
Percentage of time spent in MVPA	897		4 (2.4)	3 (2, 5)	NA		NA	NA
At least 1 hour of MVPA a day	64	7%			NA	NA		
Self report frequency of PA in past year:	2236				2045			
Never	27	1%			84	4%		
Less than once a month	116	5%			73	4%		
1 - 3 times a month	871	39%			185	9%		
1 - 4 times a week	682	31%			969	47%		
5 or more times a week	540	24%			734	36%		

Table 4.5 continued

Variable	CCT (18y8m)			
	n	%	Mean (SD)	Median (IQR)
Total daily minutes of PA	NA		NA	NA
Accelerometer counts per minute	NA		NA	NA
Daily minutes of MVPA	NA		NA	NA
Percentage of time spent in MVPA	NA		NA	NA
At least 1 hour of MVPA a day	NA	NA		
Self report frequency of PA in past year:	1153			
Never	63	5%		
Less than once a month	56	5%		
1 - 3 times a month	139	12%		
1 - 4 times a week	607	53%		
5 or more times a week	288	25%		

Table 4.6 - Descriptive statistics for physical activity measures collected in females in the ALSPAC cohort

Variable	F@11 (11y9m)				TF2 (13y10m)			
	n	%	Mean (SD)	Median (IQR)	n	%	Mean (SD)	Median (IQR)
Total daily minutes of PA	2899		334 (61)	335 (296, 379)	1983		288 (57)	283 (247, 323)
Accelerometer counts per minute	2899		551 (152)	528 (443, 638)	1983		487 (155)	463 (376, 570)
Daily minutes of MVPA	2899		18 (12)	16 (10, 25)	1983		20 (14)	17 (10, 27)
Percentage of time spent in MVPA	2899		2 (1.5)	2 (1, 3)	1983		3 (1.8)	2 (1, 3)
At least 1 hour of MVPA a day	13	<1%			36	2%		
Self report frequency of PA in past year:					3368			
Never	NA	NA	NA	NA	47	1%		
Less than once a month	NA	NA	NA	NA	55	2%		
1 - 3 times a month	NA	NA	NA	NA	241	7%		
1 - 4 times a week	NA	NA	NA	NA	1997	59%		
5 or more times a week	NA	NA	NA	NA	1028	31%		

Table 4.6 continued

Variable	TF3 (15y6m)				CCS (16y8m)			
	n	%	Mean (SD)	Median (IQR)	n	%	Mean (SD)	Median (IQR)
Total daily minutes of PA	1120		258 (55)	252 (219, 292)	NA		NA	NA
Accelerometer counts per minute	1120		435 (139)	415 (336, 505)	NA		NA	NA
Daily minutes of MVPA	1120		19 (15)	15 (8, 26)	NA		NA	NA
Percentage of time spent in MVPA	1120		2 (1.9)	2 (1, 3)	NA		NA	NA
At least 1 hour of MVPA a day	25	2%			NA	NA		
Self report frequency of PA in past year:	2515				2951			
Never	85	3%			161	5%		
Less than once a month	281	11%			215	7%		
1 - 3 times a month	1390	55%			497	17%		
1 - 4 times a week	556	22%			1559	53%		
5 or more times a week	203	8%			519	18%		

Table 4.6 continued

Variable	CCT (18y8m)			
	n	%	Mean (SD)	Median (IQR)
Total daily minutes of PA	NA		NA	NA
Accelerometer counts per minute	NA		NA	NA
Daily minutes of MVPA	NA		NA	NA
Percentage of time spent in MVPA	NA		NA	NA
At least 1 hour of MVPA a day	NA	NA		
Self report frequency of PA in past year:	2107			
Never	159	8%		
Less than once a month	212	10%		
1 - 3 times a month	422	20%		
1 - 4 times a week	1063	50%		
5 or more times a week	251	12%		

Table 4.7 – Descriptive statistics for depression measures collected in the ALSPAC cohort

Variable	F10 (10y 8m)				F@11 (11y 9m)				TF1 (12y 10m)				TF2 (13y 10m)			
	n	%	Mean (SD)	Median (IQR)	n	%	Mean (SD)	Median (IQR)	n	%	Mean (SD)	Median (IQR)	n	%	Mean (SD)	Median (IQR)
Depressive symptom score	7272		4 (3.5)	3 (1, 6)	*NA*		*NA*	*NA*	6632		4 (3.8)	3 (1, 5)	5944		5 (4.5)	4 (2, 7)
Classified as depressed*	430	6%			*NA*	*NA*			467	7%			691	12%		
Males																
Depressive symptom score	3581		4 (3.4)	3 (2, 6)	*NA*		*NA*	*NA*	3248		4 (3.5)	3 (1, 5)	2908		4 (3.8)	3 (1, 6)
Classified as depressed*	208	6%			*NA*	*NA*			167	5%			217	8%		
Females																
Depressive symptom score	3691		4 (3.6)	3 (1, 6)	*NA*		*NA*	*NA*	3384		4 (4.2)	3 (1, 6)	3036		6 (4.9)	4 (2, 8)
Classified as depressed*	222	6%			*NA*	*NA*			300	9%			474	16%		

**Classified as depressed if SMFQ score ≥ 11*

Table 4.7 continued

Variable	TF3 (15y 6m) n	%	Mean (SD)	Median (IQR)	CCS (16y 8m) n	%	Mean (SD)	Median (IQR)	TF4 (17y 10m) n	%	Mean (SD)	Median (IQR)	CCT (18y 8m) n	%	Mean (SD)	Median (IQR)
Depressive symptom score	*NA*		*NA*	*NA*	4954		6 (5.6)	4 (2, 8)	4457		7 (5.2)	5 (3, 10)	3307		7 (5.9)	5 (2, 9)
Classified as depressed*	*NA*	*NA*			892	18%			963	22%			721	22%		
Males																
Depressive symptom score	*NA*		*NA*	*NA*	2012		4 (4.6)	3 (1, 6)	1894		6 (4.8)	4 (2, 8)	1174		5 (5.0)	4 (2, 7)
Classified as depressed*	*NA*	*NA*			216	11%			316	17%			157	13%		
Females																
Depressive symptom score	*NA*		*NA*	*NA*	2942		7 (6.0)	5 (2, 10)	2560		7 (5.5)	6 (3, 11)	2133		8 (6.2)	6 (3, 11)
Classified as depressed*	*NA*	*NA*			676	23%			647	25%			564	26%		

**Classified as depressed if SMFQ score ≥ 11*

There was strong evidence that the time invariant confounders were associated with depressive symptom score. At the first time point being female was associated with a lower depression score than males, however at all other time points being female was associated with a higher depression score. There was evidence that higher maternal levels of depression and lower social class were associated with an increased depressive symptom score at all time points (Table 4.8). The relationship between cross-sectional age and depression score and maternal education and depression was inconsistent across time points (Table 4.9)

Table 4.8 – Univariable association between depressive symptom score and time invariant confounders/covariates in the ALSPAC cohort

Variable	F10 (10y 8m)			F@11 (11y 9m)			TF1 (12y 10m)			TF2 (13y 10)		
	n	Mean (SD)	p-value	n	Mean (SD)	p-value	n	Mean (SD)	p-value	n	Mean (SD)	p-value
Sex:	7272			*NA*			6632			5944		
Male	3581	4.15 (3.43)		*NA*	*NA*		3248	3.57 (3.46)		2908	4.09 (3.80)	
Female	3691	3.92 (3.56)	0.004	*NA*	*NA*	*NA*	3384	4.34 (4.16)	<0.001	3036	5.71 (4.92)	<0.001
Maternal depression	5875			*NA*	*NA*	*NA*	5409			4914		
High	3057	4.26 (3.62)					2787	4.23 (3.94)		2534	5.24 (4.58)	
Low	2818	3.71 (3.30)	<0.001				2622	3.61 (3.65)	<0.001	2380	4.49 (4.25)	<0.001
Maternal education:	6645			*NA*			6072			5485		
CSE	896	4.43 (3.81)		*NA*	*NA*		778	3.84 (3.72)		670	5.10 (4.68)	
Vocational	568	4.09 (3.73)		*NA*	*NA*		512	3.90 (3.94)		445	4.64 (4.36)	
O level	2356	4.01 (3.55)		*NA*	*NA*		2146	3.86 (3.87)		1933	4.88 (4.49)	
A level	1766	3.99 (3.40)		*NA*	*NA*		1637	4.06 (3.85)		1501	4.89 (4.43)	
Degree	1059	3.69 (3.23)	<0.001	*NA*	*NA*	*NA*	999	4.10 (3.75)	0.311	936	5.07 (4.43)	0.405
Maternal social class:	5739			*NA*			5275			4804		
I	407	3.53 (3.07)		*NA*	*NA*		373	4.28 (3.79)		358	4.61 (4.19)	
II	1994	3.92 (3.36)		*NA*	*NA*		1851	4.07 (3.80)		1704	5.04 (4.51)	
III (non-manual)	2413	4.04 (3.58)		*NA*	*NA*		2230	3.84 (3.90)		2004	4.69 (4.34)	
III (manual)	387	3.95 (3.54)		*NA*	*NA*		340	3.61 (3.50)		317	4.65 (4.30)	
IV	453	4.41 (3.79)		*NA*	*NA*		409	4.18 (4.01)		361	5.63 (5.31)	
V	84	4.50 (3.92)	0.010	*NA*	*NA*	*NA*	71	3.58 (2.81)	0.061	58	4.29 (3.88)	0.003

Table 4.8 continued

Variable	TF3 (15y 6m) n	TF3 (15y 6m) Mean (SD)	TF3 (15y 6m) p-value	CCS (16y 8m) n	CCS (16y 8m) Mean (SD)	CCS (16y 8m) p-value	TF4 (17y 10m) n	TF4 (17y 10m) Mean (SD)	TF4 (17y 10m) p-value	CCT (18y 8m) n	CCT (18y 8m) Mean (SD)	CCT (18y 8m) p-value
Sex:	*NA*			4954			4454			3307		
Male	*NA*	*NA*		2012	4.30 (4.58)		1894	5.64 (4.77)		1174	5.30 (4.96)	
Female	*NA*	*NA*	*NA*	2942	7.00 (6.02)	<0.001	2560	7.28 (5.47)	<0.001	2133	7.63 (6.21)	<0.001
Maternal depression	*NA*	*NA*	*NA*	4173			3647			2806		
High				2124	6.32 (5.68)		1863	7.05 (5.31)		1422	7.38 (6.10)	
Low				2049	5.23 (5.30)	<0.001	1784	5.92 (5.04)	<0.001	1384	6.07 (5.54)	<0.001
Maternal education:	*NA*			4614			4082			3091		
CSE	*NA*	*NA*		522	6.49 (6.31)		453	7.36 (6.14)		321	7.90 (6.84)	
Vocational	*NA*	*NA*		333	5.94 (5.68)		284	6.69 (5.52)		214	6.82 (5.92)	
O level	*NA*	*NA*		1531	6.01 (5.82)		1372	6.76 (5.33)		996	7.25 (6.10)	
A level	*NA*	*NA*		1296	5.68 (5.37)		1159	6.44 (5.08)		865	6.45 (5.65)	
Degree	*NA*	*NA*	*NA*	932	5.52 (4.99)	0.013	814	5.90 (4.58)	<0.001	695	5.95 (5.11)	<0.001
Maternal social class:	*NA*			4027			3580			2725		
I	*NA*	*NA*		352	5.40 (4.85)		315	5.82 (4.56)		254	6.00 (5.28)	
II	*NA*	*NA*		1466	5.76 (5.34)		1315	6.38 (5.07)		1012	6.50 (5.72)	
III (non-manual)	*NA*	*NA*		1622	5.73 (5.58)		1412	6.52 (5.30)		1097	6.92 (5.95)	
III (manual)	*NA*	*NA*		230	5.97 (5.79)		215	7.35 (5.50)		156	6.44 (6.08)	
IV	*NA*	*NA*		304	6.85 (6.46)		273	7.18 (5.30)		181	7.24 (5.52)	
V	*NA*	*NA*	*NA*	52	5.02 (5.22)	<0.001	48	8.00 (7.18)	<0.001	24	7.79 (6.07)	0.052

Table 4.9 - Association between age and depressive symptom score at each time point in the ALSPAC cohort

Time point	Depressive symptom score			
	n	Mean (SD)	Correlation	p-value
F10 (10y8m)	7272	4.03 (3.50)	0.005	0.668
F@11 (11y9m)	*NA*	*NA*	*NA*	*NA*
TF1 (12y10m)	6632	3.96 (3.85)	0.038	0.002
TF2 (13y10m)	5944	4.92 (4.48)	0.043	0.001
TF3 (15y6m)	*NA*	*NA*	*NA*	*NA*
CCS (16y8m)	4954	5.90 (5.64)	-0.019	0.190
TF4 (17y10m)	4313	6.55 (5.24)	0.003	0.840
CCT (18y8m)	3307	6.80 (5.90)	-0.032	0.062

4.2. TRAILS

As described in Section 3.1.2, TRAILS is a population cohort based in the Netherlands that aims to better understand the causes and mechanisms involved in mental health disorders and social development of adolescents and young adults of various ages (Table 4.10). The TRAILS cohort consists of 2230 participants; 1132 (51%) females and 1098 (49%) males (Table 4.11).

Figure 4.2 - Flow chart of participant retention in TRAILS cohort

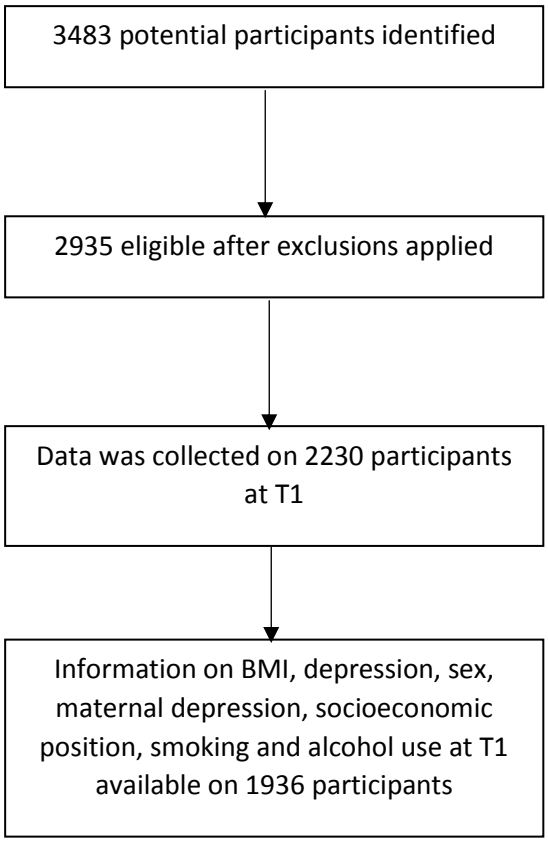


Table 4.10 - Ages of TRAILS participants at the different waves of follow up

		Age	
Time point	**n**	**Mean (SD)**	**Median (IQR)**
T1	2230	10y7m (7.8m)	11y (10y, 11y)
T2	2149	13y1m (7.2m)	13y (13y, 13y)
T3	1819	15y10m (9.1m)	16y (15y, 16y)
T4	1881	18y7m (7.4m)	19y (18y, 19y)

Table 4.11 - Time invariant sociodemographic characteristics of TRAILS participants

Variable	n	%	Mean (SD)	Median (IQR)
Sex:	2230			
Male	1098	49%		
Female	1132	51%		
Maternal Depression (DASS)	2039		0.257 (0.345)	0.143 (0, 0.429)
Socio-Economic Status:	2188			
Lowest 25% SES	553	25%		
Middle 50% SES	1084	50%		
Highest 25% SES	551	25%		

The mean age of participants at the first measurement wave was 10.6 years (approximately 10years 7months), 13.1 years (approximately 13years 1month) at the second measurement wave, 15.8 years (approximately 15years 10months) at the third measurement wave and 18.6 years (approximately 18years 7months) at the fourth wave of data collection (Table 4.10). Data on participant cigarette smoking was collected at all four waves of follow up. At the first wave of data collection the vast majority of participants (98%) responded that they had never smoked. Across the four waves of data collection the distribution of responses to frequency of smoking changed from highly positively skewed to almost bimodal in nature (distribution of responses at final time point: "not at all" – 62%; "sometimes" – 7%; "often" – 31%) (Table 4.12). Data on the number of alcoholic drinks consumed per week was also collected at all four waves of data collection. From follow up occasion one the distribution of responses changes from highly positively skewed (Number of alcoholic drinks per week: "0" – 69%; "1" – 16%; "2-3" – 8%; "4-6" – 3%; "7 or more" – 4%) to negatively skewed (Number of alcoholic drinks per week: "0" – 11%; "1" – 4%; "2-3" – 14%; "4-6" – 23%; "7 or more" – 49%) (Table 4.12).

Table 4.12 – Time varying characteristics of participants of the TRAILS cohort

Variable	T1 (10y7m) n	%	T2 (13y1m) n	%	T3 (15y10m) n	%	T4 (18y7m) n	%
Smoking in past month:	2179		2084		1657		1627	
Not at all	2126	98%	1903	91%	1208	73%	1008	62%
Sometimes (less than 1 a day)	47	2%	51	2%	87	5%	113	7%
Often (at least 1 a day)	6	0%	130	6%	362	22%	506	31%
Alcohol drinks per week:	2199		2060		1625		1618	
0	1518	69%	1206	59%	351	22%	173	11%
1	341	16%	381	19%	141	9%	58	4%
2 to 3	177	8%	215	10%	268	16%	225	14%
4 to 6	71	3%	138	7%	343	21%	366	23%
7 or more	92	4%	120	6%	522	32%	796	49%

Various anthropometric measures were collected at the four follow up occasions: height, weight and BMI at all four follow ups, subscapular skinfold thickness at follow up one and three, and waist circumference at follow up three and four. Mean height and weight increased across the four waves of data collection (in both males and females) (Table 4.13). Mean height and weight were very similar in males and females at the first two follow up occasions, however at follow ups three and four males were taller and heavier than females (Table 4.13). Mean BMI steadily increased across the measurement waves and was similar in both males and females (Table 4.13). In males mean subscapular skinfold thickness was very similar, where measured, at time points 1 and 3 (41mm and 40mm respectively). In females however mean subscapular skinfold thickness increased from 48mm at time point 1 to 62mm at time point 3 (Table 4.13). Waist circumference was measured at follow up occasions 3 and 4, in males there was increase in mean waist circumference from 76cm to 82cm, in females however there was no change in mean waist circumference (75cm at both time points) (Table 4.13).

Table 4.13 – Descriptive statistics for anthropometric measures collected in the TRAILS cohort

Variable	T1 (10y7m)			T2 (13y1m)		
	n	Mean (SD)	Median (IQR)	n	Mean (SD)	Median (IQR)
Height (cm)	2166	151 (7.8)	152 (147, 157)	2041	165 (8.2)	165 (159, 170)
Weight (kg)	2161	42 (9.3)	40 (35, 47)	2030	53 (11.1)	51 (45, 59)
BMI	2161	18 (3.1)	17 (16, 19)	2028	19 (3.2)	18 (17, 20)
Subscapular Skinfold Thickness (mm)	1569	44 (23.8)	37 (27, 56)	*NA*	*NA*	*NA*
Waist Cirumference (cm)	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
<u>Males</u>						
Height (cm)	1064	151 (7.8)	151 (146, 156)	1003	166 (9.4)	166 (159, 172)
Weight (kg)	1062	41 (9.0)	39 (35, 45)	1001	53 (12.1)	51 (44, 59)
BMI	1062	18 (2.9)	17 (16, 19)	1000	19 (3.1)	18 (17, 20)
Subscapular Skinfold Thickness (mm)	796	41 (24.1)	33 (24, 51)	*NA*	*NA*	*NA*
Waist Cirumference (cm)	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*
<u>Females</u>						
Height (cm)	1102	152 (7.8)	152 (148, 157)	1038	164 (6.9)	164 (160, 169)
Weight (kg)	1099	43 (9.5)	42 (36, 48)	1029	53 (10.0)	52 (46, 58)
BMI	1099	18 (3.2)	18 (16, 20)	1028	19 (3.3)	19 (17, 21)
Subscapular Skinfold Thickness (mm)	773	48 (23.0)	42 (31, 61)	*NA*	*NA*	*NA*
Waist Cirumference (cm)	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*

Table 4.13 continued

Variable	T3 (15y10m) n	T3 (15y10m) Mean (SD)	T3 (15y10m) Median (IQR)	T4 (18y7m) n	T4 (18y7m) Mean (SD)	T4 (18y7m) Median (IQR)
Height (cm)	1597	174 (9.1)	173 (168, 180)	1576	178 (9.9)	175 (169, 183)
Weight (kg)	1593	65 (11.9)	63 (57, 70)	1574	71 (13.6)	69 (62, 77)
BMI	1593	21 (3.3)	21 (19, 23)	1574	23 (3.9)	22 (20, 24)
Subscapular Skinfold Thickness (mm)	1580	51 (24.5)	47 (32, 65)	*NA*	*NA*	*NA*
Waist Cirumference (cm)	1589	75 (8.9)	74 (70, 79)	1563	78 (10.0)	77 (72, 83)
Males						
Height (cm)	759	180 (7.8)	180 (174, 185)	728	183 (7.4)	184 (178, 188)
Weight (kg)	759	68 (13.2)	68 (60, 74)	728	76 (13.9)	73 (67, 82)
BMI	759	21 (3.4)	20 (19, 22)	728	23 (3.8)	22 (20, 24)
Subscapular Skinfold Thickness (mm)	755	40 (22.2)	33 (26, 45)	*NA*	*NA*	*NA*
Waist Cirumference (cm)	757	76 (9.3)	74 (70, 78)	726	82 (9.7)	79 (75, 86)
Females						
Height (cm)	838	169 (6.7)	169 (160, 173)	848	170 (6.9)	170 (165, 174)
Weight (kg)	834	62 (9.6)	61 (55, 66)	846	66 (11.6)	65 (59, 72)
BMI	834	22 (3.2)	21 (20, 23)	846	23 (4.0)	22 (21, 25)
Subscapular Skinfold Thickness (mm)	825	62 (21.5)	59 (47, 73)	*NA*	*NA*	*NA*
Waist Cirumference (cm)	832	75 (8.4)	74 (70, 79)	837	75 (9.3)	74 (70, 79)

Self-report data on frequency of PA per week ("none", "once a week", "two or three days a week", "four or five days a week", or "six or seven days a week") was collected at each of the four measurement occasions. The distribution of responses across the four measurement occasions changed from an approximately normal to a normal distribution with a slight negative skew. When stratified by gender the distribution of responses was very similar between males and females at each follow up occasion (Table 4.14).

Table 4.14 – Descriptive statistics for physical activity measures collected in the TRAILS cohort

	Time 1		Time 2		Time 3		Time 4	
Variable	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
Frequency of PA:	2191		2070		1644		1639	
None	271	12%	177	9%	158	10%	236	14%
Once a week	528	24%	274	13%	225	14%	228	14%
2 or 3 days a week	768	35%	826	40%	596	36%	494	30%
4 or 5 days a week	333	15%	519	25%	365	22%	409	35%
6 or 7 days a week	291	13%	274	13%	300	18%	272	17%
Males								
Frequency of PA:	1076		1008		769		738	
None	124	12%	78	8%	66	9%	94	13%
Once a week	183	17%	120	12%	78	10%	87	12%
2 or 3 days a week	368	34%	390	39%	284	37%	212	29%
4 or 5 days a week	190	18%	273	27%	181	24%	208	28%
6 or 7 days a week	211	20%	147	15%	160	21%	137	19%
Females								
Frequency of PA:	1115		1062		875		901	
None	147	13%	99	9%	92	11%	142	16%
Once a week	345	31%	154	15%	147	17%	141	16%
2 or 3 days a week	400	36%	436	41%	312	36%	282	31%
4 or 5 days a week	143	13%	246	23%	184	21%	201	22%
6 or 7 days a week	80	7%	127	12%	140	16%	135	15%

Depression was measured using the YSR APS at each of the four waves of data collection. Depressive symptom score was highly positively skewed at each time point. The mean depressive symptom score and the percentage of participants classified as depressed remained fairly constant across the four waves of data collection (Table 4.15). However, when stratified by gender, in males there was a decrease in mean depressive symptom score (0.28 to 0.22) and percentage classified as depressed (25% to 17%) between time points 1 (mean age 10y 7m) and 2 (mean age 13y 1m) but, at subsequent time points, both the mean symptom score and percentage classified as depressed remained fairly constant. In females, there was an increase in the mean depressive symptom score and percentage classified as depressed from time point 1 (mean age 10y 7m) to 3 (mean age 15y 10m) and then mean depressive symptom score was fairly stable between time points 3 (mean age 15y 10m) and 4 (mean age 18y 7m) although there was a slight decrease in percentage classified as depressed (Table 4.15).

Table 4.15 – Descriptive statistics for the depression measure collected in the TRAILS cohort

Variable	T1 (10y7m) n	%	Mean (SD)	Median (IQR)	T2 (13y1m) n	%	Mean (SD)	Median (IQR)
Depressive symtom score	2191		0.29 (0.25)	0.23 (0.08, 0.46)	2092		0.27 (0.26)	0.23 (0.08, 0.38)
Classified as depressed	578	26%			495	24%		
<u>Males</u>								
Depressive symtom score	1074		0.28 (0.25)	0.23 (0.08, 0.38)	1019		0.22 (0.22)	0.15 (0.08, 0.31)
Classified as depressed	266	25%			169	17%		
<u>Females</u>								
Depressive symtom score	1117		0.30 (0.25)	0.23 (0.08, 0.46)	1073		0.32 (0.29)	0.23 (0.08, 0.46)
Classified as depressed	312	28%			326	30%		

Table 4.15 continued

Variable	T3 (15y10m) n	%	Mean (SD)	Median (IQR)	T4 (18y7m) n	%	Mean (SD)	Median (IQR)
Depressive symtom score	1661		0.30 (0.27)	0.23 (0.08, 0.46)	1696		0.30 (0.30)	0.21 (0.07, 0.43)
Classified as depressed	442	27%			402	24%		
Males								
Depressive symtom score	777		0.22 (0.22)	0.15 (0.08, 0.31)	768		0.23 (0.26)	0.14 (0.07, 0.32)
Classified as depressed	126	16%			126	16%		
Females								
Depressive symtom score	884		0.36 (0.30)	0.31 (0.15, 0.54)	928		0.36 (0.33)	0.29 (0.14, 0.57)
Classified as depressed	316	36%			276	30%		

When associations between depression and potential confounders/covariates were investigated, there was evidence that the mean depression score was higher in females than in males (Table 4.16), and that maternal and offspring depressive symptom scores were positively correlated (Table 4.16). However, there was no evidence that depression symptom scores differed by socio-economic status (Table 4.16). There was evidence of a positive association between depressive symptom score and smoking and alcohol use (Table 4.17). There was however no evidence of correlation between cross-sectional age and depression symptom score (Table 4.17).

Table 4.16 – Univariable association between depressive symptom score and time invariant confounders/covariates in the TRAILS cohort

Variable	Time 1 (10y7m)			Time 2 (13y1m)			Time 3 (15y10m)			Time 4 (18y7m)		
	n	Mean (SD)	p-value*	n	Mean (SD)	p-value*	n	Mean (SD)	p-value*	n	Mean (SD)	p-value*
Sex:	2191			2092			1661			1696		
Male	1074	0.28 (0.25)		1019	0.22 (0.22)		777	0.22 (0.22)		768	0.23 (0.26)	
Female	1117	0.30 (0.25)	0.216	1073	0.32 (0.29)	<0.001	884	0.36 (0.30)	<0.001	928	0.36 (0.33)	<0.001
Maternal Depression	2007			1926			1556			1594		
High	1206	0.30 (0.25)		1158	0.29 (0.26)		943	0.31 (0.27)		964	0.32 (0.31)	
Low	801	0.27 (0.25)	0.004	768	0.25 (0.25)	<0.001	613	0.27 (0.27)	0.004	630	0.26 (0.27)	<0.001
Socio-Economic Status:	2151			2061			1637			1673		
Lowest 25% SES	535	0.30 (0.26)		501	0.28 (0.26)		352	0.30 (0.27)		339	0.31 (0.31)	
Middle 50% SES	1069	0.29 (0.25)		1027	0.28 (0.26)		807	0.30 (0.28)		847	0.30 (0.30)	
Highest 25% SES	547	0.29 (0.24)	0.763	533	0.26 (0.25)	0.594	478	0.27 (0.25)	0.113	487	0.29 (0.29)	0.468

**p-value from t-test if non time-varying variable is binary, from pairwise correlation if continuous, and from ANOVA if categorical*

Table 4.17 – Univariable association between depressive symptom score and time-varying confounders/covariates in the TRAILS cohort

Variable	Time 1 (10y7m)				Time 2 (13y1m)				Time 3 (15y10m)				Time 4 (18y7m)			
	n	Mean (SD)	Correlation	p-value*	n	Mean (SD)	Correlation	p-value*	n	Mean (SD)	Correlation	p-value*	n	Mean (SD)	Correlation	p-value*
Age	2191	0.29 (0.25)	-0.024	0.270	2092	0.27 (0.26)	-0.016	0.477	1661	0.30 (0.27)	0.024	0.339	1696	0.30 (0.30)	0.040	0.102
Smoking in past month:	2178				2080				1652				1621			
Not at all	2125	0.29 (0.24)			1873	0.26 (0.25)			1183	0.27 (0.25)			972	0.28 (0.29)		
Sometimes (less than 1 a day)	47	0.57 (0.31)			78	0.36 (0.32)			108	0.38 (0.33)			146	0.33 (0.32)		
Often (at least 1 a day)	6	0.56 (0.28)		<0.001	129	0.41 (0.31)		<0.001	361	0.36 (0.29)		<0.001	503	0.33 (0.32)		0.008
Number of alcoholic drinks per week:	2184				2056				1620				1612			
0	1507	0.27 (0.23)			1203	0.25 (0.24)			351	0.30 (0.29)			172	0.36 (0.35)		
1	339	0.33 (0.26)			381	0.27 (0.26)			140	0.27 (0.26)			58	0.25 (0.30)		
2 to 3	176	0.35 (0.26)			215	0.32 (0.29)			267	0.27 (0.25)			225	0.31 (0.31)		
4 to 6	71	0.35 (0.25)			138	0.31 (0.26)			340	0.29 (0.27)			364	0.28 (0.27)		
7 or more	91	0.40 (0.30)		<0.001	119	0.37 (0.33)		<0.001	522	0.31 (0.27)		0.184	793	0.30 (0.30)		0.039

**p-value from pairwise correlation if continuous and from ANOVA if categorical*

4.3. NDIT

As described in Section 3.1.3, NDIT is a population cohort based in Montreal in Canada consisting of participants from 11 schools in the area. The NDIT cohort consisted of 1294 participants; 671 (52%) females and 623 (48%) males (Table 4.19). Data on maternal history of depression, maternal education and maternal job status were collected at baseline (Table 4.19).

Figure 4.3 - Flow chart of participant retention in NDIT cohort

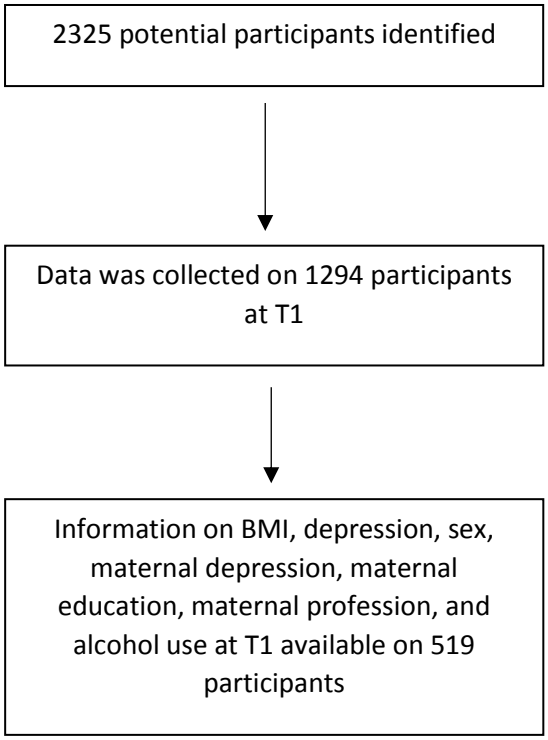


Table 4.18 - Age of NDIT participants at the different waves of follow up

Time point	Age		
	n	Mean (SD)	Median (IQR)
T1	1267	12y9m (6.6m)	12y8m (12y5m, 13y0m)
T2	1198	13y0m (6.2m)	12y11m (12y8m, 13y3m)
T3	1191	13y2m (5.8m)	13y1m (12y10m, 13y5m)
T4	545	13y2m (5.2m)	13y1m (12y10m, 13y5m)
T5	1104	13y8m (5.7m)	13y7m (13y3m, 13y11m)
T6	1101	13y10m (5.7m)	13y10m (13y6m, 14y1m)
T7	960	14y1m (5.8m)	14y0m (13y8m, 14y4m)
T8	982	14y2m (5.3m)	14y1m (13y10m, 14y5m)
T9	1022	14y7m (5.4m)	14y7m (14y3m, 14y10m)
T10	995	14y10m (5.2m)	14y10m (14y6m, 15y1m)
T11	972	15y0m (5.2m)	15y0m (14y8m, 15y3m)
T12	987	15y2m (5.0m)	15y1m (14y10m, 15y5m)
T13	914	15y7m (4.8m)	15y6m (15y3m, 15y10m)
T14	906	15y10m (4.7m)	15y9m (15y6m, 16y1m)
T15	904	16y0m (4.7m)	16y0m (15y8m, 16y3m)
T16	887	16y2m (4.7m)	16y1m (15y10m, 16y5m)
T17	871	16y6m (4.5m)	16y6m (16y3m, 16y9m)
T18	852	16y9m (4.4m)	16y9m (16y6m, 17y0m)
T19	844	17y0m (4.5m)	16y11m (16y8m, 17y3m)
T20	840	17y1m (4.4m)	17y1m (16y10m, 17y4m)

Table 4.19 – Descriptive statistics for time invariant sociodemographic characteristics of participants in the NDIT cohort

Variable	n	%
Sex:	1294	
Male	623	48.1%
Female	671	51.9%
Maternal history of depression:	578	
No	463	80.1%
Yes	115	19.9%
Maternal highest level of education:	591	
High Schoool - attended	48	8.1%
High Schoool - graduated	108	18.3%
CEGEP - attended	50	8.5%
CEGEP - graduated	92	15.6%
University - attended	57	9.6%
University - graduated BSc	141	23.9%
University - graduated MSc	45	7.6%
University - graduated PhD	4	0.7%
Other	46	7.8%
Maternal job status:	589	
Full-time job	345	58.6%
Part-time job	112	19.0%
Full-time student	2	0.3%
Part-time student	1	0.2%
Homemaker	52	8.8%
Not working for health reasons	10	1.7%
Unemployed	17	2.9%
On welfare	3	0.5%
Other	47	8.0%

The mean age of participants at the first measurement wave was 12.8 years (approximately 12years 10months), follow-up measurements were taken approximately every three months. Information on participant's level of alcohol consumption was collected at every follow up occasion. At the first measurement occasion the responses to the alcohol consumption question were highly positively skewed, 50% of respondents answered "never" and 25% responded "a bit to try" (the two lowest categories), throughout the later follow up occasions the distribution of responses became less strongly positively skewed (Table 4.20).

Table 4.20 – Descriptive statistics for time-varying confounders/covariates in the NDIT cohort

Variable	T1 (12y9m)		T2 (13y0m)		T3 (13y2m)		T4 (13y2m)		T5 (13y8m)		T6 (13y10m)		T7 (14y1m)	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Alcohol	1214		1191		1177		539		1094		1089		944	
Never	692	57%	628	53%	710	60%	323	60%	550	50%	497	46%	485	51%
A bit to try	373	31%	342	29%	255	22%	121	22%	270	25%	231	21%	187	20%
One or a couple of times a month	115	9%	176	15%	161	14%	70	13%	200	18%	249	23%	173	18%
One or a couple of times a weeek	27	2%	34	3%	40	3%	20	4%	62	6%	92	8%	85	9%
Everyday	7	1%	11	1%	11	1%	5	1%	12	1%	20	2%	14	1%

Table 4.20 continued

Variable	T8 (14y2m)		T9 (14y7m)		T10 (14y10m)		T11 (15y0m)		T12 (15y2m)		T13 (15y7m)		T14 (15y10m)	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Alcohol	966		1007		982		949		971		904		895	
Never	484	50%	460	46%	429	44%	438	46%	459	47%	338	37%	301	34%
A bit to try	182	19%	180	18%	175	18%	149	16%	152	16%	141	16%	134	15%
One or a couple of times a month	209	22%	241	24%	267	27%	253	27%	260	27%	294	33%	320	36%
One or a couple of times a weeek	67	7%	107	11%	95	10%	87	9%	90	9%	117	13%	126	14%
Everyday	24	2%	19	2%	16	2%	22	2%	10	1%	14	2%	14	2%

Table 4.20 continued

Variable	T15 (16y0m)		T16 (16y2m)		T17 (16y6m)		T18 (16y9m)		T19 (17y0m)		T20 (17y1m)	
	n	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
Alcohol	894		874		851		846		832		827	
Never	333	37%	317	36%	254	30%	263	31%	224	27%	258	31%
A bit to try	115	13%	124	14%	85	10%	78	9%	84	10%	76	9%
One or a couple of times a month	323	36%	293	34%	327	38%	326	39%	354	43%	312	38%
One or a couple of times a weeek	108	12%	131	15%	170	20%	168	20%	156	19%	167	20%
Everyday	15	2%	9	1%	15	2%	11	1%	14	2%	14	2%

Measures of height, weight, BMI, subscapular skinfold thickness and waist circumference were collected at follow up occasions one (12y 9m), twelve (15y 2m) and nineteen (17y 0m). Mean height, weight, BMI, subscapular skinfold thickness and waist circumference increased across the waves of data collection (Table 4.21). Males were taller and heavier than females at each follow up occasion but the mean BMI was very similar in males and females (Table 4.21). In males, mean subscapular skinfold thickness was very similar at time points 1 and 12 but then showed an increase by time point 19 (Table 4.21). In females, mean subscapular skinfold thickness increased at each time point and also was greater than those of males (Table 4.21). Waist circumference increased at each time point in both males and females and was greater in males than in females (Table 4.21).

Table 4.21 – Descriptive statistics for anthropometric data collect in the NDIT cohort

Variable	Time 1			Time 12			Time 19		
	n	Mean (SD)	Median (IQR)	n	Mean (SD)	Median (IQR)	n	Mean (SD)	Median (IQR)
Height (cm)	1195	156 (7.8)	156 (151, 161)	951	167 (8.2)	166 (161, 172)	801	169 (8.8)	169 (163, 176)
Weight (kg)	1195	49 (11.8)	48 (41, 56)	951	60 (12.2)	58 (52, 67)	801	65 (13.0)	63 (55, 70)
BMI	1195	20 (3.8)	19 (17, 22)	951	22 (3.7)	21 (19, 23)	801	22 (3.7)	22 (20,24)
Subscapular skinfold thickness (mm)	1194	10 (5.6)	8 (6, 13)	949	12 (5.6)	11 (8, 15)	800	15 (6.9)	13 (11, 18)
Waist circumference (cm)	1195	71 (10.1)	69 (64, 76)	951	76 (9.4)	74 (69, 80)	801	78 (9.5)	76 (72, 82)
Males									
Height (cm)	577	157 (8.8)	156 (150, 163)	462	171 (7.4)	171 (167, 176)	389	176 (6.5)	176 (171, 180)
Weight (kg)	577	50 (12.0)	48 (41, 58)	462	64 (12.2)	62 (56, 70)	389	70 (12.0)	67 (62, 76)
BMI	577	20 (3.8)	19 (17, 22)	462	22 (3.6)	21 (91, 23)	389	23 (3.7)	22 (20, 24)
Subscapular skinfold thickness (mm)	577	10 (5.9)	7 (6, 11)	460	10 (4.7)	9 (7, 12)	389	13 (6.9)	12 (9, 15)
Waist circumference (cm)	577	73 (10.3)	70 (65, 78)	462	77 (9.3)	75 (71, 80)	389	80 (9.3)	78 (74, 84)
Females									
Height (cm)	618	156 (6.7)	156 (152, 161)	489	162 (5.9)	162 (158, 166)	412	163 (6.0)	163 (159, 167)
Weight (kg)	618	49 (11.5)	48 (41, 55)	489	57 (11.3)	55 (50, 62)	412	59 (11.4)	57 (53, 64)
BMI	618	20 (3.9)	19 (17, 22)	489	22 (3.8)	21 (19, 23)	412	22 (3.8)	21 (20, 24)
Subscapular skinfold thickness (mm)	617	11 (5.3)	9 (7, 13)	489	14 (5.8)	13 (10, 17)	411	17 (6.6)	16 (12, 20)
Waist circumference (cm)	618	70 (9.8)	68 (63, 74)	489	74 (9.2)	72 (68, 79)	412	76 (9.2)	74 (70, 80)

Information on participants' level of PA was collected at each follow up occasion in NDIT. The number of bouts of MVPA a week reported by participants remained fairly stable at each follow up occasion with perhaps a slight increase over time (overall and in both males and females) (Table 4.22). Females consistently reported more bouts of MVPA a week than males did (Table 4.22). A bout of MVPA consists of a period of at least 5 minutes in an activity that has been previously defined as moderate or vigorous.

Table 4.22 – Descriptive statistics for the physical activity data collected in the NDIT cohort

Variable	T1 (12y9m) n	Mean (SD)	Median (IQR)	T2 (13y0m) n	Mean (SD)	Median (IQR)	T3 (13y2m) n	Mean (SD)	Median (IQR)	T4 (13y2m) n	Mean (SD)	Median (IQR)
No. of bouts of MVPA	1227	149 (12.7)	152 (145, 158)	1198	149 (12.5)	153 (145, 157)	1185	145 (15.1)	149 (139, 155)	540	144 (16.0)	148 (137, 155)
Males												
No. of bouts of MVPA	591	146 (14.4)	150 (141, 156)	575	146 (14.5)	150 (141, 156)	579	142 (16.6)	146 (135, 153)	263	140 (18.0)	144 (131, 153)
Females												
No. of bouts of MVPA	636	152 (10.0)	154 (148, 158)	623	151 (9.6)	154 (148, 158)	606	148 (12.7)	151 (143, 157)	277	147 (13.1)	151 (141, 156)

Table continued

Variable	T5 (13y8m) n	Mean (SD)	Median (IQR)	T6 (13y10m) n	Mean (SD)	Median (IQR)	T7 (14y1m) n	Mean (SD)	Median (IQR)	T8 (14y2m) n	Mean (SD)	Median (IQR)
No. of bouts of MVPA	1104	146 (14.7)	150 (140, 157)	1097	149 (12.5)	152 (145, 157)	954	148 (12.5)	151 (144, 157)	977	145 (14.6)	148 (138, 155)
Males												
No. of bouts of MVPA	527	143 (16.3)	148 (134, 154)	529	146 (14.2)	150 (141, 156)	463	146 (13.5)	149 (140, 156)	483	141 (16.2)	145 (135, 153)
Females												
No. of bouts of MVPA	577	149 (12.4)	153 (145, 158)	568	151 (10.0)	154 (148, 159)	491	150 (11.2)	153 (147, 158)	494	148 (12.0)	152 (142, 157)

Table continued

Variable	T9 (14y7m) n	Mean (SD)	Median (IQR)	T10 (14y10m) n	Mean (SD)	Median (IQR)	T11 (15y0m) n	Mean (SD)	Median (IQR)	T12 (15y2m) n	Mean (SD)	Median (IQR)
No. of bouts of MVPA	1015	148 (12.7)	151 (143, 157)	993	150 (11.7)	154 (145, 158)	971	150 (10.9)	153 (146, 159)	987	148 (12.5)	152 (144, 157)
Males												
No. of bouts of MVPA	482	144 (14.5)	147 (138, 154)	471	147 (13.6)	151 (142, 157)	461	148 (11.8)	151 (143, 156)	473	146 (14.2)	149 (141, 155)
Females												
No. of bouts of MVPA	533	151 (9.5)	154 (147, 158)	522	152 (9.0)	155 (149, 159)	510	153 (9.5)	154 (149, 160)	514	151 (10.2)	154 (147, 158)

Table continued

Variable	T13 (15y7m) n	Mean (SD)	Median (IQR)	T14 (15y10m) n	Mean (SD)	Median (IQR)	T15 (16y0m) n	Mean (SD)	Median (IQR)	T16 (16y2m) n	Mean (SD)	Median (IQR)
No. of bouts of MVPA	911	149 (11.4)	152 (145, 158)	902	151 (10.2)	154 (147, 159)	900	152 (10.2)	154 (147, 159)	881	150 (11.0)	153 (146, 158)
Males												
No. of bouts of MVPA	435	146 (12.6)	149 (140, 155)	430	149 (12.0)	152 (144, 157)	430	149 (12.0)	153 (145, 157)	425	147 (12.2)	150 (142, 156)
Females												
No. of bouts of MVPA	476	152 (9.1)	155 (149, 159)	472	154 (7.3)	156 (151, 160)	470	154 (7.7)	156 (151, 160)	456	153 (9.1)	155 (149, 159)

Table continued

Variable	T17 (16y6m) n	Mean (SD)	Median (IQR)	T18 (16y9m) n	Mean (SD)	Median (IQR)	T19 (17y0m) n	Mean (SD)	Median (IQR)	T20 (17y1m) n	Mean (SD)	Median (IQR)
No. of bouts of MVPA	759	150 (10.7)	153 (145, 158)	744	152 (9.5)	154 (147, 159)	733	152 (8.7)	154 (147, 159)	731	151 (9.8)	154 (147, 158)
Males												
No. of bouts of MVPA	367	147 (12.2)	150 (142, 156)	359	149 (11.0)	152 (145, 157)	353	150 (10.0)	153 (146, 158)	344	149 (11.6)	153 (143, 157)
Females												
No. of bouts of MVPA	392	152 (8.5)	154 (149, 159)	385	154 (7.3)	155 (150, 160)	380	154 (6.8)	157 (150, 160)	387	153 (7.3)	155 (150, 159)

Depressive symptoms were measured using the KDSS at each wave of follow up. Depressive symptom score was positively skewed at each time. There was a decrease in depression score over time until around follow up occasion 8 (14y 2m), after which there was an increase in depression score over time (this was observed in both males and females) (Table 4.23). When stratified by sex mean depression score was consistently higher in females than in males (Table 4.23)

Table 4.23 – Descriptive statistics for the depression data collected in the NDIT cohort

Variable	T1 (12y9m) n	Mean (SD)	Median (IQR)	T2 (13y0m) n	Mean (SD)	Median (IQR)	T3 (13y2m) n	Mean (SD)	Median (IQR)	T4 (13y2m) n	Mean (SD)	Median (IQR)
Depressive symptom score	1216	2.10 (0.61)	2.00 (1.67, 2.50)	1188	2.03 (0.65)	2.00 (1.50, 2.50)	1177	1.93 (0.69)	1.83 (1.33, 2.33)	540	1.90 (0.69)	1.83 (1.33, 2.33)
Males												
Depressive symptom score	585	1.99 (0.61)	2.00 (1.50, 2.33)	569	1.88 (0.61)	1.83 (1.33, 2.33)	574	1.77 (0.66)	1.67 (1.33, 2.17)	264	1.75 (0.63)	1.67 (1.25, 2.17)
Females												
Depressive symptom score	631	2.20 (0.60)	2.17 (1.67, 2.67)	619	2.17 (0.66)	2.00 (1.67, 2.67)	603	2.07 (0.69)	2.00 (1.50, 2.50)	276	2.04 (0.71)	2.00 (1.50, 2.50)

Table continued

Variable	T5 (13y8m) n	Mean (SD)	Median (IQR)	T6 (13y10m) n	Mean (SD)	Median (IQR)	T7 (14y1m) n	Mean (SD)	Median (IQR)	T8 (14y2m) n	Mean (SD)	Median (IQR)
Depressive symptom score	1098	1.94 (0.70)	1.83 (1.33, 2.33)	1086	1.94 (0.73)	1.83 (1.33, 2.33)	942	1.90 (0.73)	1.83 (1.33, 2.33)	968	1.90 (0.77)	1.67 (1.33, 2.33)
Males												
Depressive symptom score	525	1.73 (0.64)	1.50 (1.17, 2.17)	522	1.71 (0.69)	1.50 (1.17, 2.00)	458	1.68 (0.70)	1.50 (1.17, 2.00)	479	1.69 (0.72)	1.50 (1.00, 2.00)
Females												
Depressive symptom score	573	2.14 (0.69)	2.00 (1.67, 2.67)	564	2.15 (0.71)	2.00 (1.67, 2.67)	484	2.11 (0.71)	2.00 (1.67, 2.67)	489	2.12 (0.76)	2.00 (1.50, 2.67)

Table continued

Variable	T9 (14y7m) n	Mean (SD)	Median (IQR)	T10 (14y10m) n	Mean (SD)	Median (IQR)	T11 (15y0m) n	Mean (SD)	Median (IQR)	T12 (15y2m) n	Mean (SD)	Median (IQR)
Depressive symptom score	1007	2.02 (0.78)	2.00 (1.33, 2.50)	984	2.01 (0.79)	1.83 (1.33, 2.50)	959	1.98 (0.78)	1.83 (1.33, 2.50)	968	1.96 (0.77)	1.83 (1.33, 2.50)
Males												
Depressive symptom score	480	1.74 (0.71)	1.50 (1.17, 2.17)	467	1.73 (0.71)	1.50 (1.00, 2.17)	451	1.69 (0.69)	1.50 (1.00, 2.00)	462	1.67 (0.70)	1.50 (1.00, 2.00)
Females												
Depressive symptom score	527	2.28 (0.75)	2.17 (1.67, 2.83)	517	2.26 (0.78)	2.17 (1.67, 2.83)	508	2.23 (0.76)	2.17 (1.67, 2.67)	506	2.22 (0.74)	2.17 (1.67, 2.80)

Table continued

Variable	T13 (15y7m) n	Mean (SD)	Median (IQR)	T14 (15y10m) n	Mean (SD)	Median (IQR)	T15 (16y0m) n	Mean (SD)	Median (IQR)	T16 (16y2m) n	Mean (SD)	Median (IQR)
Depressive symptom score	909	2.02 (0.76)	2.00 (1.50, 2.50)	898	2.16 (0.81)	2.17 (1.50, 2.83)	896	2.03 (0.79)	2.00 (1.33, 2.50)	882	2.07 (0.81)	2.00 (1.33, 2.67)
Males												
Depressive symptom score	434	1.73 (0.69)	1.50 (1.17, 2.00)	430	1.84 (0.75)	1.67 (1.17, 2.33)	430	1.73 (0.75)	1.50 (1.00, 2.17)	427	1.74 (0.74)	1.50 (1.00, 2.17)
Females												
Depressive symptom score	475	2.29 (0.71)	2.33 (1.67, 2.83)	468	2.46 (0.74)	2.50 (2.00, 3.00)	466	2.31 (0.72)	2.33 (1.83, 2.83)	455	2.39 (0.75)	2.33 (1.83, 3.00)

Table continued

Variable	T17 (16y6m) n	Mean (SD)	Median (IQR)	T18 (16y9m) n	Mean (SD)	Median (IQR)	T19 (17y0m) n	Mean (SD)	Median (IQR)	T20 (17y1m) n	Mean (SD)	Median (IQR)
Depressive symptom score	866	2.09 (0.80)	2.00 (1.50, 2.67)	845	2.25 (0.82)	2.17 (1.50, 2.83)	836	2.11 (0.79)	2.00 (1.50, 2.67)	837	2.13 (0.81)	2.17 (1.50, 2.67)
Males												
Depressive symptom score	413	1.77 (0.75)	1.50 (1.17, 2.17)	397	1.87 (0.74)	1.83 (1.17, 2.33)	395	1.80 (0.75)	1.67 (1.00, 2.33)	388	1.78 (0.75)	1.67 (1.00, 2.33)
Females												
Depressive symptom score	453	2.38 (0.73)	2.33 (1.83, 2.83)	448	2.58 (0.75)	2.67 (2.00, 3.17)	441	2.40 (0.71)	2.33 (2.00, 2.83)	449	2.44 (0.73)	2.50 (2.00, 3.00)

There was strong evidence at every follow up occasion ($p<0.001$ at every time point) that mean depression score differed by gender, with females having a higher depression score than males (Table 4.24). There was inconsistent evidence of an association between maternal history of depression and level of depression in the participants. Of the 20 follow up occasions there was only evidence of an association between level of depression and mother's level of education at two time points (follow up occasion 17 and 18) (Table 4.24). There was inconsistent evidence that maternal job status was associated with level of depressive symptoms (Table 4.24). Similarly, there was inconsistent evidence of an association between age and level of depression. There was also strong, consistent evidence that level of alcohol consumption was associated with participants' level of depression, with higher levels of alcohol consumption were associated with higher levels of depression (Table 4.25).

Table 4.24 – Univariable association between depressive symptom score and time invariant confounders/covariates in the NDIT cohort

Variable	Time 1			Time 2			Time 3			Time 4		
	n	Mean (SD)	p-value	n	Mean (SD)	p-value	n	Mean (SD)	p-value	n	Mean (SD)	p-value
Sex:	1216			1188			1177			540		
Male	585	1.99 (0.61)		569	1.88 (0.61)		574	1.77 (0.66)		264	1.75 (0.63)	
Female	631	2.20 (0.60)	<0.001	619	2.17 (0.66)	<0.001	603	2.07 (0.69)	<0.001	276	2.04 (0.71)	<0.001
Maternal history of depression:	554			541			537			240		
No	441	2.07 (0.57)		437	2.01 (0.62)		432	1.88 (0.63)		202	1.83 (0.65)	
Yes	113	2.18 (0.63)	0.077	104	2.05 (0.65)	0.545	105	2.03 (0.73)	0.032	38	1.99 (0.78)	0.173
Maternal highest level of education:	567			555			550			247		
High Schoool - attended	45	2.20 (0.55)		42	2.12 (0.56)		42	1.95 (0.64)		10	1.43 (0.30)	
High Schoool - graduated	102	2.13 (0.58)		99	2.01 (0.62)		101	1.99 (0.69)		29	1.94 (0.71)	
CEGEP - attended	48	2.11 (0.56)		47	1.89 (0.57)		46	1.76 (0.69)		19	1.65 (0.77)	
CEGEP - graduated	89	2.16 (0.70)		85	2.12 (0.66)		85	2.05 (0.74)		40	1.90 (0.72)	
University - attended	54	2.08 (0.46)		54	2.05 (0.64)		52	1.91 (0.61)		20	1.98 (0.68)	
University - graduated BSc	137	2.04 (0.60)		137	1.98 (0.64)		135	1.85 (0.58)		84	1.83 (0.63)	
University - graduated MSc	43	1.95 (0.52)		42	1.94 (0.56)		41	1.85 (0.48)		24	1.94 (0.66)	
University - graduated PhD	4	2.29 (0.25)		4	2.04 (0.48)		4	2.29 (0.58)		3	2.17 (0.44)	
Other	45	2.03 (0.47)	0.405	45	1.97 (0.66)	0.574	44	1.77 (0.66)	0.123	18	1.82 (0.67)	0.393
Maternal job status:	565			555			548			246		
Full-time job	333	2.10 (0.58)		325	2.03 (0.60)		326	1.91 (0.63)		155	1.89 (0.65)	
Part-time job	110	2.12 (0.57)		106	2.02 (0.61)		103	1.89 (0.60)		43	1.75 (0.65)	
Full-time student	2	1.83 (1.18)		2	1.75 (0.82)		2	1.75 (0.82)		0	NA	
Part-time student	1	1.50 (0)		1	2.33 (0)		1	1.83 (0)		1	1.67 (0)	
Homemaker	49	2.03 (0.57)		48	1.86 (0.68)		47	1.80 (0.69)		17	1.76 (0.75)	
Not working for health reasons	10	2.47 (0.80)		10	2.58 (1.02)		8	2.75 (0.94)		2	2.42 (1.77)	
Unemployed	14	2.18 (0.66)		16	2.10 (0.72)		14	2.08 (0.84)		8	1.77 (0.82)	
On welfare	3	2.30 (0.96)		2	1.75 (0.35)		3	2.39 (1.51)		0	NA	
Other	43	1.96 (0.50)	0.315	45	1.89 (0.56)	0.058	44	1.81 (0.58)	0.016	20	1.83 (0.57)	0.748

p-value from ANOVA

Table continued

Variable	Time 5 n	Time 5 Mean (SD)	Time 5 p-value	Time 6 n	Time 6 Mean (SD)	Time 6 p-value	Time 7 n	Time 7 Mean (SD)	Time 7 p-value	Time 8 n	Time 8 Mean (SD)	Time 8 p-value
Sex:	1098			1086			942			968		
Male	525	1.73 (0.64)		522	1.71 (0.69)		458	1.68 (0.70)		479	1.69 (0.72)	
Female	573	2.14 (0.69)	<0.001	564	2.15 (0.71)	<0.001	484	2.12 (0.71)	<0.001	489	2.12 (0.76)	<0.001
Maternal history of depression:	524			517			444			463		
No	427	1.93 (0.69)		421	1.90 (0.70)		364	1.87 (0.70)		380	1.85 (0.74)	
Yes	97	1.92 (0.71)	0.924	96	2.02 (0.74)	0.151	80	1.97 (0.82)	0.260	83	2.11 (0.85)	0.005
Maternal highest level of education:	536			528			453			475		
High Schoool - attended	44	2.13 (0.63)		41	2.05 (0.71)		40	2.19 (0.79)		30	2.04 (0.84)	
High Schoool - graduated	93	1.88 (0.67)		95	1.96 (0.69)		90	1.88 (0.66)		82	1.87 (0.66)	
CEGEP - attended	45	1.87 (0.81)		40	1.85 (0.81)		35	1.85 (0.86)		39	1.75 (0.83)	
CEGEP - graduated	83	2.05 (0.76)		84	1.98 (0.79)		75	1.94 (0.79)		79	1.93 (0.85)	
University - attended	53	2.05 (0.75)		52	1.95 (0.68)		41	1.83 (0.67)		48	1.97 (0.82)	
University - graduated BSc	129	1.81 (0.59)		130	1.80 (0.61)		99	1.81 (0.68)		120	1.84 (0.72)	
University - graduated MSc	41	1.95 (0.73)		41	2.02 (0.73)		36	1.98 (0.77)		39	2.06 (0.77)	
University - graduated PhD	4	1.92 (0.50)		4	2.17 (0.41)		2	2.00 (0.24)		3	1.61 (0.19)	
Other	44	1.84 (0.65)	0.091	41	1.92 (0.75)	0.466	35	1.70 (0.54)	0.156	35	1.84 (0.67)	0.605
Maternal job status:	536			528			453			475		
Full-time job	315	1.92 (0.67)		315	1.95 (0.69)		262	1.90 (0.70)		280	1.88 (0.71)	
Part-time job	103	1.89 (0.66)		101	1.84 (0.75)		89	1.83 (0.76)		91	1.84 (0.70)	
Full-time student	2	1.92 (1.06)		2	1.92 (1.01)		2	1.75 (0.82)		1	3.00 (0)	
Part-time student	1	1.67 (0)		1	1.17 (0)		0	NA		1	1.50 (0)	
Homemaker	45	2.01 (0.76)		44	1.87 (0.75)		42	1.88 (0.78)		40	1.77 (0.81)	
Not working for health reasons	10	2.72 (0.93)		9	2.24 (0.86)		8	2.38 (0.97)		8	3.04 (1.13)	
Unemployed	15	1.92 (0.84)		14	2.18 (0.76)		11	2.03 (0.82)		11	2.05 (1.04)	
On welfare	2	2.08 (0.82)		1	2.67 (0)		2	2.33 (0.47)		1	3.00 (0)	
Other	43	1.86 (0.61)	0.061	44	1.79 (0.60)	0.309	37	1.79 (0.59)	0.500	42	1.86 (0.75)	0.001

Table continued

Variable	Time 9 n	Time 9 Mean (SD)	Time 9 p-value	Time 10 n	Time 10 Mean (SD)	Time 10 p-value	Time 11 n	Time 11 Mean (SD)	Time 11 p-value	Time 12 n	Time 12 Mean (SD)	Time 12 p-value
Sex:	1007			984			959			968		
Male	480	1.74 (0.71)		467	1.73 (0.71)		451	1.69 (0.69)		462	1.67 (0.70)	
Female	527	2.28 (0.75)	<0.001	517	2.26 (0.78)	<0.001	508	2.23 (0.76)	<0.001	506	2.22 (0.74)	<0.001
Maternal history of depression:	501			503			489			483		
No	412	2.00 (0.76)		412	1.97 (0.76)		405	1.95 (0.77)		399	1.91 (0.75)	
Yes	89	2.13 (0.86)	0.152	91	2.31 (0.88)	<0.001	84	2.08 (0.84)	0.164	84	2.14 (0.83)	0.014
Maternal highest level of education:	513			515			501			493		
High Schoool - attended	36	2.14 (0.76)		37	2.19 (0.89)		37	1.99 (0.83)		31	2.20 (0.96)	
High Schoool - graduated	86	1.96 (0.71)		85	1.96 (0.72)		87	1.97 (0.71)		79	1.90 (0.69)	
CEGEP - attended	42	1.82 (0.84)		40	1.86 (0.89)		41	1.83 (0.83)		40	1.88 (0.90)	
CEGEP - graduated	78	2.07 (0.79)		83	2.02 (0.76)		78	1.93 (0.78)		79	1.98 (0.82)	
University - attended	53	2.09 (0.84)		53	2.21 (0.88)		52	2.05 (0.78)		52	1.96 (0.73)	
University - graduated BSc	127	2.03 (0.79)		127	1.96 (0.77)		125	1.93 (0.78)		122	1.86 (0.73)	
University - graduated MSc	43	2.07 (0.77)		44	2.20 (0.74)		42	2.23 (0.80)		44	2.13 (0.71)	
University - graduated PhD	4	2.13 (0.44)		4	2.25 (0.69)		4	1.96 (0.42)		4	2.00 (0.38)	
Other	44	1.99 (0.78)	0.749	42	1.98 (0.79)	0.241	41	1.86 (0.85)	0.426)	42	2.00 (0.77)	0.406
Maternal job status:	513			515			500			493		
Full-time job	304	2.03 (0.76)		310	2.04 (0.78)		297	2.00 (0.79)		299	1.99 (0.78)	
Part-time job	99	2.02 (0.77)		99	1.99 (0.81)		98	1.93 (0.77)		92	1.92 (0.77)	
Full-time student	2	1.92 (1.30)		2	2.17 (1.41)		1	2.83 (0)		2	2.00 (1.41)	
Part-time student	1	2.00 (0)		1	1.83 (0)		1	1.67 (0)		1	1.83 (0)	
Homemaker	44	2.03 (0.89)		43	2.06 (0.90)		41	1.86 (0.76)		40	1.90 (0.73)	
Not working for health reasons	7	2.45 (1.18)		6	2.58 (1.08)		7	2.00 (0.96)		6	2.00 (1.05)	
Unemployed	10	2.03 (0.88)		9	2.06 (0.91)		11	1.79 (0.89)		11	1.85 (0.81)	
On welfare	1	1.17 (0)		1	1.00 (0)		1	1.00 (0)		0	NA	
Other	45	1.97 (0.71)	0.889	44	1.93 (0.63)	0.656	43	2.00 (0.69)	0.735	42	1.88 (0.68)	0.973

Table continued

Variable	Time 13 n	Time 13 Mean (SD)	Time 13 p-value	Time 14 n	Time 14 Mean (SD)	Time 14 p-value	Time 15 n	Time 15 Mean (SD)	Time 15 p-value	Time 16 n	Time 16 Mean (SD)	Time 16 p-value
Sex:	909			898			896			882		
Male	434	1.73 (0.69)		430	1.84 (0.75)		430	1.73 (0.75)		427	1.74 (0.74)	
Female	475	2.29 (0.71)	<0.001	468	2.46 (0.74)	<0.001	466	2.31 (0.72)	<0.001	455	2.39 (0.75)	<0.001
Maternal history of depression:	477			482			474			474		
No	393	2.02 (0.75)		396	2.13 (0.79)		389	1.99 (0.76)		391	2.05 (0.81)	
Yes	84	2.13 (0.78)	0.224	86	2.30 (0.83)	0.064	85	2.13 (0.83)	0.131	83	2.16 (0.84)	0.258
Maternal highest level of education:	487			491			485			484		
High Schoool - attended	30	2.14 (0.92)		30	2.29 (0.93)		28	2.18 (0.88)		28	2.20 (0.87)	
High Schoool - graduated	79	2.07 (0.71)		79	2.15 (0.77)		78	2.06 (0.71)		78	2.04 (0.78)	
CEGEP - attended	42	1.82 (0.73)		41	2.03 (0.92)		39	1.87 (0.87)		38	1.87 (0.88)	
CEGEP - graduated	81	2.05 (0.81)		81	2.22 (0.85)		80	2.04 (0.78)		82	2.06 (0.85)	
University - attended	53	2.09 (0.66)		53	2.19 (0.75)		54	2.13 (0.78)		54	2.11 (0.83)	
University - graduated BSc	115	1.98 (0.76)		121	2.09 (0.76)		120	1.93 (0.73)		121	1.95 (0.76)	
University - graduated MSc	42	2.16 (0.74)		43	2.28 (0.73)		43	2.11 (0.74)		41	2.30 (0.75)	
University - graduated PhD	4	2.21 (0.76)		4	2.17 (0.82)		4	1.87 (0.75)		3	2.00 (0.67)	
Other	41	2.05 (0.77)	0.600	39	2.09 (0.79)	0.776	39	1.88 (0.78)	0.465	39	2.24 (0.83)	0.219
Maternal job status:	487			491			484			484		
Full-time job	288	2.05 (0.77)		294	2.12 (0.76)		290	2.00 (0.76)		288	2.05 (0.81)	
Part-time job	93	1.96 (0.68)		97	2.24 (0.86)		92	2.03 (0.80)		92	2.06 (0.81)	
Full-time student	2	1.75 (1.06)		2	1.83 (0.71)		1	1.00 (0)		2	1.83 (1.18)	
Part-time student	1	1.83 (0)		1	1.33 (0)		0	NA		1	1.33 (0)	
Homemaker	40	2.03 (0.75)		39	2.09 (0.72)		39	2.06 (0.74)		41	2.09 (0.80)	
Not working for health reasons	8	2.37 (0.73)		7	2.83 (0.89)		8	2.27 (0.90)		8	2.19 (0.48)	
Unemployed	10	1.90 (0.84)		9	2.30 (0.96)		9	2.02 (1.14)		9	2.20 (1.09)	
On welfare	1	1.00 (0)		0	NA		1	2.33 (0)		1	1.83 (0)	
Other	44	2.17 (0.84)	0.584	42	2.21 (0.77)	0.248	44	1.99 (0.65)	0.868	42	2.16 (0.84)	0.970

Table continued

Variable	Time 17 n	Time 17 Mean (SD)	Time 17 p-value	Time 18 n	Time 18 Mean (SD)	Time 18 p-value	Time 19 n	Time 19 Mean (SD)	Time 19 p-value	Time 20 n	Time 20 Mean (SD)	Time 20 p-value
Sex:	866			845			836			837		
Male	413	1.77 (0.75)		397	1.87 (0.74)		395	1.80 (0.75)		388	1.78 (0.75)	
Female	453	2.38 (0.73)	<0.001	448	2.58 (0.75)	<0.001	441	2.40 (0.70)	<0.001	449	2.44 (0.73)	<0.001
Maternal history of depression:	469			459			457			461		
No	386	2.08 (0.77)		376	2.23 (0.79)		376	2.09 (0.75)		379	2.13 (0.79)	
Yes	83	2.17 (0.88)	0.331	83	2.22 (0.83)	0.908	81	2.16 (0.77)	0.414	82	2.21 (0.84)	0.418
Maternal highest level of education:	479			469			467			472		
High Schoool - attended	28	2.37 (0.88)		28	2.57 (0.90)		29	2.26 (0.82)		30	2.17 (0.90)	
High Schoool - graduated	76	2.18 (0.77)		78	2.26 (0.78)		76	2.12 (0.68)		80	2.18 (0.78)	
CEGEP - attended	40	1.79 (0.84)		39	1.97 90.86)		36	1.91 (0.84)		36	1.94 (0.93)	
CEGEP - graduated	80	2.06 (0.78)		75	2.21 (0.78)		73	2.12 (0.79)		78	2.12 (0.79)	
University - attended	52	2.22 (0.79)		51	2.34 (0.78)		51	2.17 (0.83)		48	2.30 (0.83)	
University - graduated BSc	116	1.92 (0.69)		113	2.07 (0.69)		117	1.99 (0.71)		115	2.08 (0.74)	
University - graduated MSc	43	2.30 (0.80)		43	2.48 (0.78)		43	2.25 (0.75)		42	2.29 (0.79)	
University - graduated PhD	4	2.04 (0.98)		4	2.67 (0.68)		4	2.63 (0.44)		3	2.39 (1.00)	
Other	40	2.15 (0.81)	0.007	38	2.24 (0.83)	0.007	38	2.13 (0.65)	0.225	40	2.09 (0.69)	0.510
Maternal job status:	479			469			467			471		
Full-time job	284	2.07 (0.77)		278	2.22 (0.78)		275	2.05 (0.73)		278	2.12 (0.76)	
Part-time job	94	2.11 (0.79)		91	2.25 (0.81)		92	2.22 (0.75)		93	2.24 (0.85)	
Full-time student	2	1.50 (0.47)		1	2.33 (0)		1	1.67 (0)		1	1.50 (0)	
Part-time student	1	1.83 (0)		1	2.17 (0)		1	1.50 (0)		1	1.67 (0)	
Homemaker	39	2.22 (0.85)		40	2.22 (0.81)		39	2.14 (0.82)		39	2.12 (0.81)	
Not working for health reasons	6	1.94 (0.91)		7	2.55 (0.81)		6	2.47 (0.64)		6	2.53 (0.79)	
Unemployed	10	1.97 (0.98)		9	2.11 (1.28)		9	2.31 (1.09)		9	2.07 (1.07)	
On welfare	1	1.67 (0)		1	1.83 (0)		1	2.17 (0)		1	1.67 (0)	
Other	42	2.17 (0.80)	0.878	41	2.21 (0.73)	0.974	43	2.08 (0.70)	0.555	43	2.15 (0.75)	0.789

Table 4.25 - Univariable association between depressive symptom score and time-varying confounders/covariates in the NDIT cohort

Variable	T1 (12y9m) n	Mean (SD)	Correlation	p-value*	T2 (13y0m) n	Mean (SD)	Correlation	p-value*	T3 (13y2m) n	Mean (SD)	Correlation	p-value*	T4 (13y2m) n	Mean (SD)	Correlation	p-value*
age	1216	2.10 (0.61)	0.103	<0.001	1188	2.03 (0.65)	0.066	0.022	1177	1.93 (0.69)	0.058	0.047	540	1.90 (0.69)	0.072	0.093
Alcohol	1204				1181				1163				534			
Never	685	2.00 (0.57)			623	1.92 (0.61)			699	1.78 (0.63)			318	1.77 (0.65)		
A bit to try	372	2.15 (0.60)			340	2.08 (0.64)			254	2.09 (0.70)			121	2.07 (0.68)		
One or a couple of times a month	113	2.43 (0.67)			174	2.26 (0.71)			160	2.13 (0.70)			70	2.15 (0.67)		
One or a couple of times a weeek	27	2.33 (0.72)			33	2.39 90.72)			39	2.49 (0.84)			20	1.91 (0.77)		
Everyday	7	2.24 (1.01)		<0.001	11	2.23 (0.62)		<0.001	11	2.05 (1.06)		<0.001	5	2.47 (1.23)		<0.001

Table continued

Variable	T5 (13y8m) n	Mean (SD)	Correlation	p-value*	T6 (13y10m) n	Mean (SD)	Correlation	p-value*	T7 (14y1m) n	Mean (SD)	Correlation	p-value*	T8 (14y2m) n	Mean (SD)	Correlation	p-value*
age	1098	1.94 (0.70)	0.026	0.385	1086	1.94 (0.73)	0.084	0.007	942	1.90 (0.73)	0.090	0.006	968	1.90 (0.77)	0.086	0.008
Alcohol	1088				1075				929				953			
Never	548	1.81 (0.65)			493	1.74 (0.65)			477	1.71 (0.65)			477	1.76 (0.73)		
A bit to try	268	2.01 (0.67)			227	2.00 (0.71)			186	1.94 (0.68)			180	1.99 (0.72)		
One or a couple of times a month	198	2.06 (0.71)			244	2.12 (0.74)			170	2.16 (0.77)			206	2.06 (0.74)		
One or a couple of times a weeek	62	2.39 (0.78)			91	2.29 (0.79)			83	2.29 (0.84)			66	2.13 (0.92)		
Everyday	12	2.45 (1.10)		<0.001	20	2.32 (1.07)		<0.001	13	2.59 (0.86)		<0.001	24	1.99 (1.02)		<0.001

Table continued

Variable	T9 (14y7m) n	Mean (SD)	Correlation	p-value*	T10 (14y10m) n	Mean (SD)	Correlation	p-value*	T11 (15y0m) n	Mean (SD)	Correlation	p-value*	T12 (15y2m) n	Mean (SD)	Correlation	p-value*
age	1007	2.02 (0.78)	0.064	0.044	984	2.01 (0.79)	0.041	0.203	959	1.98 (0.78)	0.055	0.087	968	1.96 (0.77)	0.070	0.030
Alcohol	993				971				937				955			
Never	452	1.86 (0.74)			425	1.84 (0.75)			431	1.80 (0.75)			452	1.83 (0.77)		
A bit to try	177	2.10 (0.75)			173	2.06 (0.72)			146	2.07 (0.77)			151	1.99 (0.73)		
One or a couple of times a month	238	2.13 (0.78)			263	2.19 (0.84)			252	2.22 (0.78)			255	2.14 (0.75)		
One or a couple of times a weeek	107	2.28 (0.81)			95	2.15 (0.82)			86	1.98 (0.78)			87	2.11 (0.84)		
Everyday	19	2.23 (1.14)		<0.001	15	2.14 (0.84)		<0.001	22	2.03 (0.82)		<0.001	10	1.83 (1.09)		<0.001

Table continued

Variable	T13 (15y7m) n	Mean (SD)	Correlation	p-value*	T14 (15y10m) n	Mean (SD)	Correlation	p-value*	T15 (16y0m) n	Mean (SD)	Correlation	p-value*	T16 (16y2m) n	Mean (SD)	Correlation	p-value*
age	909	2.02 (0.76)	0.009	0.785	898	2.16 (0.81)	-0.013	0.688	896	2.03 (0.79)	0.001	0.983	882	2.07 (0.81)	0.056	0.099
Alcohol	899				887				886				869			
Never	337	1.87 (0.78)			299	1.98 (0.82)			327	1.88 (0.80)			315	1.95 (0.87)		
A bit to try	139	2.03 (0.73)			132	2.15 (0.74)			115	2.12 (0.82)			124	2.19 (0.76)		
One or a couple of times a month	292	2.11 (0.70)			316	2.32 (0.78)			322	2.17 (0.76)			291	2.17 (0.79)		
One or a couple of times a weeek	117	2.19 (0.76)			126	2.26 (0.80)			107	2.04 (0.74)			130	2.12 (0.72)		
Everyday	14	2.27 (0.79)		<0.001	14	1.92 (0.94)		<0.001	15	1.90 (0.87)		<0.001	9	1.76 (0.65)		0.003

Table continued

Variable	T17 (16y6m) n	Mean (SD)	Correlation	p-value*	T18 (16y9m) n	Mean (SD)	Correlation	p-value*	T19 (17y0m) n	Mean (SD)	Correlation	p-value*	T20 (17y1m) n	Mean (SD)	Correlation	p-value*
age	866	2.09 (0.80)	0.013	0.706	845	2.25 (0.82)	0.007	0.851	836	2.11 (0.79)	0.056	0.104	837	2.13 (0.81)	0.028	0.417
Alcohol	847				840				826				824			
Never	251	1.98 (0.82)			261	2.05 (0.89)			222	1.91 (0.90)			258	1.96 (0.89)		
A bit to try	85	2.01 (0.74)			77	2.28 (0.81)			84	2.20 (0.70)			74	2.10 (0.75)		
One or a couple of times a month	326	2.14 (0.79)			325	2.33 (0.76)			353	2.17 (0.72)			312	2.28 (0.73)		
One or a couple of times a weeek	170	2.17 (0.76)			167	2.36 (0.79)			154	2.21 (0.74)			166	2.13 (0.76)		
Everyday	15	2.27 (0.95)		0.044	10	2.90 (0.95)		<0.001	13	2.12 (1.08)		<0.001	14	2.27 (1.20)		<0.001

CHAPTER 5. RESULTS AND DISCUSSION – OBJECTIVE 1; OBESITY AND DEPRESSION

5.1. Linear Regression

5.1.1. BMI

ALSPAC

In the ALSPAC cohort the linear regression analyses investigating the relationship between BMI and later depressive symptoms showed strong evidence of an association between BMI at (mean) age 10 years 8 months and depressive symptom score two years later (mean age 12 years 10 months) (Table 5.1 column 4 row 2). The analysis suggested that a one kg/m^2 unit increase in BMI was associated with a 0.018 standard deviation (SD) (95% CI 0.008, 0.027, p-value <0.001) increase in depressive symptom score. A similar result was obtained when investigating the association between BMI at age 13 years 10 months and depressive symptoms at age 16 years 8 months: a one unit increase in BMI was associated with a 0.016 standard deviation increase in later depressive symptom score (95% CI: 0.005, 0.027). However the analyses looking at the association between BMI (exposure) at age 12 years 10 months and depressive symptoms (outcome) at 13 years 10 months, and BMI (exposure) at 17 years 10 months and depressive symptoms (outcome) at 18 years 8 months showed no evidence of an association (0.007 SD increase in depressive symptom score per unit increase in BMI, 95% CI: -0.002, 0.015, p-value: 0.119, and 0.001 SD increase in depressive symptom score per unit increase in BMI, 95%CI -0.009, 0.012, p-value: 0.805 respectively) (Table 5.1).

The regression models at the different time points were each repeated including an interaction between BMI and sex to test for differences in the association between obesity and depression in males and females. There was some evidence of an interaction by sex when BMI (as exposure) was measured at 12y 10m and depressive symptoms (as outcome) was measured at 13y 10m (a one unit increase in BMI was associated with a greater increase in depressive symptoms, of 0.018 SDs, in females compared to males), and again when BMI (as exposure) was measured at 13y 10m and depressive symptoms (as outcome) was measured at 16y 8m (Table 5.1 column 5). The coefficients in column 5 of Table 5.1 represent the BMI*sex interaction coefficient. When the regression analyses were carried out stratified by sex the results suggested that there was a positive association between BMI and depressive symptoms at the next follow up occasion in females (except in the oldest age group) (Table 5.1 continued column 4). In males however, there was evidence of a positive association in the first regression model (the earliest time points), and no evidence to support an association between BMI and depressive symptoms in older boys (Table 5.1 continued column 5).

In order to test for a potential non-linear "U" shaped relationship between BMI and depression, the ALSPAC linear regression analysis was repeated including a BMI squared term in the model. There was no evidence of an association between the quadratic BMI term and depressive symptom score (range of p-values; 0.510 to 0.828) (see Appendix 4). Another sensitivity analysis was carried out to investigate the impact of puberty. The linear regression analysis in ALSPAC was repeated including a measure of puberty. There was no evidence of an association between puberty and depressive symptom score, and the inclusion of puberty did not alter the conclusions being drawn on the relationship between BMI and depression (see Appendix 5). As such these effects were not considered in subsequent analyses. A further sensitivity analysis was carried out including physical activity as a confounder in the regression model. The inclusion of PA had little impact on findings (see Appendix 6).

Table 5.1 - Results from the linear regression analyses investigating the association between BMI (exposure) on depression (Z score) (outcome) at next follow up in the ALSPAC cohort

Timepoint	Exposure	Outcome	Model 1 - main model				Model 2 - Including interaction (coefficient for interaction effect)			
			n	coeff.	95%CI	p-value	n	coeff[#].	95%CI	p-value
F10 to TF1 10y8m to 12y10m	BMI	Depression	4264	0.018	(0.008, 0.027)	<0.001	4264	-0.002	(-0.020, 0.017)	0.853
TF1 to TF2 12y10m to 13y10m	BMI	Depression	3964	0.007	(-0.002, 0.015)	0.119	3964	0.018	(0.001, 0.035)	0.036
TF2 to CCS 13y10m to 16y8m	BMI	Depression	2864	0.016	(0.005, 0.027)	0.004	2864	0.024	(0.002, 0.045)	0.030
TF4 to CCT 17y10m to 18y8m	BMI	Depression	1723	0.001	(-0.009, 0.012)	0.805	1723	0.011	(-0.009, 0.032)	0.276

Model 1 is adjusted for age (at outcome), sex, previous depression, maternal depression, maternal education and maternal profession

*Model 2 is Model 1 plus the inclusion of a BMI*Sex interaction term*

[#] *Results presented are for the BMI* Sex interaction term*

Table continued

Timepoint	Exposure	Outcome	Model 3 - stratified (females)				Model 4 - stratified (males)			
			n	coeff.	95%CI	p-value	n	coeff.	95%CI	p-value
F10 to TF1 10y8m to 12y10m	BMI	Depression	2172	0.018	(0.004, 0.031)	0.009	2092	0.018	(0.005, 0.030)	0.008
TF1 to TF2 12y10m to 13y10m	BMI	Depression	2013	0.015	(0.003, 0.028)	0.016	1951	-0.002	(-0.013, 0.010)	0.749
TF2 to CCS 13y10m to 16y8m	BMI	Depression	1621	0.024	(0.009, 0.039)	0.002	1243	0.004	(-0.011, 0.019)	0.594
TF4 to CCT 17y10m to 18y8m	BMI	Depression	1091	0.005	(-0.009, 0.018)	0.495	632	-0.007	(-0.023, 0.008)	0.352

Model 3 is Model 1 stratified by gender - analysing females only and without including sex as a confounder

Model 4 is Model 1 stratified by gender - analysing males only and without including sex as a confounder

TRAILS

In the TRAILS cohort there was evidence of a positive association between BMI and depressive symptoms at the next follow up occasion in the first two regression models (age 10 years 7 months to 13 years 1 month, and 13 years 1 month to 15 years 10 months) (Table 5.2 column 4 rows 2 and 3). In the third model however (15 year 10 months to 18 years 7 months), the direction of the association had reversed (regression coefficient now negative), suggesting that a one unit increase in BMI was associated with a 0.012 standard deviation decrease in depressive symptom score at the next follow up occasion. However, the 95% CI included the null in this regression model (Table 5.2 column 4 row 4).

The linear regression models in TRAILS were repeated including an interaction between BMI and sex to test for differences in the association between BMI and depression by sex. There was no evidence of an interaction by sex with the lowest p-value from the interaction models being 0.077; when BMI (as exposure) was measured at 10y 7m and depressive symptoms (as outcome) was measured at 13y 1m (a one unit increase in BMI was associated with a greater increase in depressive symptoms, of 0.025 SDs, in females compared to males) (Table 5.2 column 5 row 2). When this regression model with BMI (as exposure) at age 10y 7m and depressive symptom score (as outcome) at age 13y 1m was stratified by sex, there was evidence of an association between BMI and depressive symptom score in females but not in males (Table 5.2 columns 6 and 7, row 2). There was no evidence of an interaction between BMI and sex at the later time points (Table 5.2 column 5 rows 3 and 4). Stratifying the regression models by gender showed that there was no consistent evidence for a difference between males and females. It should be noted that, as was seen in the regression models on all subjects (males and females together), the direction of the association between BMI and depression reversed in the last regression model (age 15y 10m to 18y 7m), the regression coefficients were positive in the first two models (age 10y 7m to 13y 1m, and 13y 1m to 15y 10m) but negative in the last model (15y 10m to 18y 7m) (Table 5.2 continued).

Table 5.2 - Results from the linear regression analyses investigating the association between BMI (exposure) on depression (Z score) (outcome) at next follow up in the TRAILS cohort

Timepoint	Exposure	Outcome	Model 1 - main model				Model 2 - including interaction (coefficient for interaction effect)			
			n	coeff.	95%CI	p-value	n	coeff[#].	95%CI	p-value
T1 to T2 10y7m to 13y1m	BMI	Depression	1836	0.016	(0.002, 0.030)	0.030	1836	0.025	(-0.003, 0.053)	0.077
T2 to T3 13y1m to 15y10m	BMI	Depression	1475	0.016	(0.001, 0.030)	0.033	1475	-0.01	(-0.032, 0.022)	0.714
T3 to T4 15y10m to 18y7m	BMI	Depression	1276	-0.012	(-0.025, 0.002)	0.086	1276	0.001	(-0.025, 0.027)	0.948

Model 1 is adjusted for age, sex, previous depression, maternal depression, SEP, alcohol and smoking

*Model 2 is Model 1 plus the inclusion of a BMI*Sex interaction term*

[#] *Results presented are for the BMI* Sex interaction term*

Table continued

Timepoint	Exposure	Outcome	Model 3 - stratified (females)				Model 4 - stratified (males)			
			n	coeff.	95%CI	p-value	n	coeff.	95%CI	p-value
T1 to T2 10y7m to 13y1m	BMI	Depression	947	0.025	(0.003, 0.046)	0.023	889	0.002	(-0.017, 0.022)	0.806
T2 to T3 13y1m to 15y10m	BMI	Depression	773	0.013	(-0.009, 0.036)	0.237	702	0.020	(0.002, 0.037)	0.029
T3 to T4 15y10m to 18y7m	BMI	Depression	697	-0.010	(-0.030, 0.010)	0.325	579	-0.02	(-0.034, 0.003)	0.108

Model 3 is Model 1 stratified by gender - analysing females only and without including sex as a confounder

Model 4 is Model 1 stratified by gender - analysing males only and without including sex as a confounder

NDIT

It should be noted that in the NDIT cohort the interval between waves of follow-up was much shorter than in the other cohorts (on average ~3 months). In the NDIT cohort there was evidence for a positive association between BMI and depressive symptom score at one of the time points (age 15 years 2 months to 15 years 7 months) but not the others (12 years 10 months to 13 years, and 17 years to 17 years 1 month), in fact at these other time points the regression coefficients were negative (although the confidence intervals were wide) (Table 5.3 column 4).

When the linear regression model was repeated including an interaction between BMI and sex to test for differences in the association between obesity and depression by sex there was no evidence for a difference in the association between BMI and depressive symptom score in males and females (Table 5.3 column 5). When stratified by sex there was evidence of an inverse relationship between BMI and depressive symptoms in females at the earliest time point (12 years 10 months to 13 years): a one unit increase in BMI was associated with a 0.027 standard deviation decrease in depressive symptom score (95% CI -0.051, -0.003, p-value 0.025) (Table 5.3 column 6).

Table 5.3 - Results from the linear regression analyses investigating the association between BMI (exposure) on depression (Z score) (outcome) at next follow up in the NDIT cohort

Timepoint	Exposure	Outcome	Model 1 - main model				Model 2 - including interaction (coefficient for interaction effect)			
			n	coeff.	95%CI	p-value	n	coeff[#].	95%CI	p-value
T1 to T2 12y10m to 13y0m	BMI	Depression	496	-0.008	(-0.025, 0.009)	0.355	496	-0.026	(-0.060, 0.006)	0.108
T12 to T13 15y2m to 15y7m	BMI	Depression	433	0.024	(0.006, 0.043)	0.011	433	2E-04	(-0.037, 0.038)	0.993
T19 to T20 17y0m to 17y1m	BMI	Depression	416	-0.003	(-0.019, 0.012)	0.659	416	0.002	(-0.029, 0.033)	0.913

Model 1 is adjusted for age, sex, previous depression, maternal education, maternal profession and alcohol

*Model 2 is Model 1 plus the inclusion of a BMI*Sex interaction term*

[#] *Results presented are for the BMI* Sex interaction term*

Table continued

Timepoint	Exposure	Outcome	Model 3 - stratified (females)				Model 4 - stratified (males)			
			n	coeff.	95%CI	p-value	n	coeff.	95%CI	p-value
T1 to T2 12y10m to 13y0m	BMI	Depression	263	-0.027	(-0.051, -0.003)	0.025	233	0.004	(-0.019, 0.027)	0.723
T12 to T13 15y2m to 15y7m	BMI	Depression	222	0.025	(-0.002, 0.052)	0.069	211	0.024	(-0.006, 0.053)	0.117
T19 to T20 17y0m to 17y1m	BMI	Depression	218	-0.004	(-0.028, 0.020)	0.759	198	-0.003	(-0.025, 0.020)	0.823

Model 3 is Model 1 stratified by gender - analysing females only and without including sex as a confounder

Model 4 is Model 1 stratified by gender - analysing males only and without including sex as a confounder

5.1.2. Other Measures of Obesity

ALSPAC

In the ALSPAC cohort as well as BMI, waist circumference and DXA body fat percentage were also collected (at a smaller number of time points than BMI) as alternative measures of obesity. Analyses were repeated using these additional measures of obesity and are presented below.

Waist Circumference

The analyses utilising waist circumference as the measure of obesity provided evidence of an association between waist circumference at (mean) age 10 years 8 months and depressive symptom score at (mean) age 12 years 10 months (as was also found when using BMI as the exposure): a one centimetre increase in waist circumference was associated with a 0.005 standard deviation increase in depressive symptoms (95% CI: 0.002, 0.009, p-value: 0.002) (Table 5.4 row 2). A similar result was obtained when investigating the association between waist circumference at age 13 years 10 months and depression at age 16 years 8 months: a one centimetre increase in waist circumference was associated with a 0.006 standard deviation increase in later depressive symptom score (95% CI: 0.002, 0.010, p-value: 0.003) (Table 5.4 row 4). However the analyses looking at the association between waist circumference at age 12 years 10 months and depression at 13 years 10 months showed no evidence of an association (0.001 SD increase in depressive symptom score per one centimetre increase in waist circumference, 95%CI: -0.002, 0.004, p-value: 0.363) (Table 5.4 row 3). This is the same pattern of results as was found when using BMI as the exposure variable.

The regression models at the different time points were each repeated including an interaction between waist circumference and sex to test for differences in the association between obesity and depression by sex. The results suggested that, in the

regression model at the latest time point (waist circumference at 13 years 10 months and depressive symptoms at age 16 years 8 months), there was a more positive association in females than males (Table 5.4 column 5). When the regression analysis was stratified by sex at this time point the results provided strong evidence that there was a positive association between waist circumference and depressive symptom score at the next follow up occasion in females but not males (Females: 0.010 SD increase in depressive symptom score per one centimetre increase in waist circumference, 95% CI 0.004, 0.016, p-value 0.001. Males: 0.001 SD increase in depressive symptoms per one centimetre increase in waist circumference, 95% CI -0.005, 0.007, p-value 0.759) (Table 5.4 column 6). This is the same pattern of results as was found when using BMI as the exposure variable.

Table 5.4 - Results from the linear regression analyses investigating the association between Waist Circumference (exposure) on depression (Z score) (outcome) at next follow up in the ALSPAC cohort

Timepoint	Exposure	Outcome	Model 1 - main model				Model 2 - Including interaction (coefficient for interaction effect)			
			n	coeff.	95%CI	p-value	n	coeff[#].	95%CI	p-value
F10 to TF1 10y8m to 12y10m	Waist Circumference	Depression	4286	0.005	(0.002, 0.009)	0.002	4286	-0.001	(-0.008, 0.005)	0.677
TF1 to TF2 12y10m to 13y10m	Waist Circumference	Depression	3973	0.001	(-0.002, 0.004)	0.363	3973	0.004	(-0.002, 0.010)	0.236
TF2 to CCS 13y10m to 16y8m	Waist Circumference	Depression	2861	0.006	(0.002, 0.010)	0.003	2861	0.010	(0.002, 0.019)	0.012

Model 1 is adjusted for age, sex, previous depression, maternal depression, maternal education and maternal profession

*Model 2 is Model 1 plus the inclusion of a Waist Circumference*Sex interaction term*

[#] *Results presented are for the Waist Circumference* Sex interaction term*

Table continued

Timepoint	Exposure	Outcome	Model 3 - stratified (females)				Model 4 - stratified (males)			
			n	coeff.	95%CI	p-value	n	coeff.	95%CI	p-value
F10 to TF1 10y8m to 12y10m	Waist Circumference	Depression	2189	0.005	(-0.0003, 0.010)	0.064	2097	0.006	(0.001, 0.010)	0.010
TF1 to TF2 12y10m to 13y10m	Waist Circumference	Depression	2019	0.003	(-0.001, 0.008)	0.161	1954	-0.0002	(-0.004, 0.003)	0.914
TF2 to CCS 13y10m to 16y8m	Waist Circumference	Depression	1619	0.010	(0.004, 0.016)	0.001	1242	0.001	(-0.005, 0.007)	0.759

Model 3 is Model 1 stratified by gender - analysing females only and without including sex as a confounder

Model 4 is Model 1 stratified by gender - analysing males only and without including sex as a confounder

DXA Fat Percentage

The analyses utilising DXA fat percentage as the measure of obesity (exposure) provided evidence of an association between percentage body fat at (mean) age 13 years 10 months and depressive symptom score at (mean) age 16 years 8 months: a one percent increase in body fat was associated with a 0.005 standard deviation increase in depressive symptoms (95% CI: 0.001, 0.009, p-value: 0.008). However the analysis looking at the association of body fat percentage at age 17 years 10 months and depression at 18 years 8 months showed no evidence of an association (Table 5.5 column 4). This is the same pattern of result as was observed when BMI was used as the exposure variable.

The regression models at the different time points were each repeated including an interaction between DXA fat percentage and gender. In the regression model investigating body fat percentage at 13 years 10 months and depression at age 16 years 8 months, there was evidence that the effect of DXA fat percentage on depressive symptoms differed between males and females (Table 5.5 column 5). When the regression analysis was carried out stratified by sex at this time point there was evidence of a positive association between fat percentage and depressive symptoms at the next follow up occasion in females but not males (Females: 0.011 SD increase in depressive symptom score per one percent increase in body fat, 95% CI 0.005, 0.017, p-value <0.001. Males: -0.0003 reduction in depressive symptom score per one percent increase in body fat, 95% CI -0.005, 0.004, p-value 0.872) (Table 5.5 column 6). This is the same pattern of results as was found when using BMI as the exposure variable.

Table 5.5 - Results from the linear regression analyses investigating the association between DXA Fat Percentage (exposure) on depression (Z score) (outcome) at next follow up in the ALSPAC cohort

Timepoint	Exposure	Outcome	Model 1 - main model				Model 2 - Including interaction (coefficient for interaction effect)			
			n	coeff.	95%CI	p-value	n	coeff[#].	95%CI	p-value
TF2 to CCS 13y10m to 16y8m	DXA Fat Percentage	Depression	2827	0.005	(0.001, 0.009)	0.008	2827	0.01	(0.005, 0.020)	0.001
TF4 to CCT 17y10m to 18y8m	DXA Fat Percentage	Depression	1692	1E-04	(-0.004, 0.005)	0.775	1692	0.01	(-0.005, 0.014)	0.330

Model 1 is adjusted for age, sex, previous depression, maternal depression, maternal education and maternal profession

*Model 2 is Model 1 plus the inclusion of a DXA*Sex interaction term*

[#] *Results presented are for the DXA* Sex interaction term*

Table 5.5 continued

Timepoint	Exposure	Outcome	Model 3 - stratified (females)				Model 4 - stratified (males)			
			n	coeff.	95%CI	p-value	n	coeff.	95%CI	p-value
TF2 to CCS 13y10m to 16y8m	DXA Fat Percentage	Depression	1605	0.011	(0.005, 0.017)	<0.001	1222	-0	(-0.005, 0.004)	0.872
TF4 to CCT 17y10m to 18y8m	DXA Fat Percentage	Depression	1071	0.003	(-0.004, 0.009)	0.433	621	-0	(-0.010, 0.004)	0.439

Model 3 is Model 1 stratified by gender - analysing females only and without including sex as a confounder

Model 4 is Model 1 stratified by gender - analysing males only and without including sex as a confounder

TRAILS

In the TRAILS cohort as well as BMI, waist circumference and subscapular skinfold thickness were also collected (at a smaller number of time points than BMI) as alternative measures of obesity. Analyses were repeated using these additional measures of obesity and are presented below.

Waist Circumference

There was evidence of an inverse association between waist circumference at (mean) age 15 years 10 months and depressive symptoms at (mean) age 18 years 7 months: a one centimetre increase in waist circumference was associated with a 0.006 standard deviation decrease in depressive symptom score (95% CI: -0.011, 0.001, p-value: 0.017). This is similar to what was observed when BMI was used as the exposure variable, a negative regression coefficient was obtained at this time point when BMI was used as the measure of obesity. When an interaction between waist circumference and sex was included in the regression model there was no evidence that the results differed by sex (p = 0.604) (Table 5.6).

Table 5.6 - Results from the linear regression analyses investigating the association between Waist Circumference (exposure) on depression (Z score) (outcome) at next follow up in the TRAILS cohort

			T3 to T4 15y10m to 18y7m			
Model	**Exposure**	**Outcome**	**n**	**coeff.**	**95% CI**	**p-value**
Model 1 - main model	Waist Circumference	Depression	1274	-0.006	(-0.011, -0.001)	0.017
Model 2 - interaction model[#]	Waist Circumference	Depression	1274	-0.003	(-0.012, 0.007)	0.604
Model 3 - stratified (females)	Waist Circumference	Depression	577	-0.007	(-0.014, 0.001)	0.073
Model 4 - stratified (males)	Waist Circumference	Depression	697	-0.006	(-0.013, 0.001)	0.078

Model 1 is adjusted for age, sex, previous depression, maternal depression, SEP, alcohol and smoking

*Model 2 is Model 1 plus the inclusion of a Waist Circumference*Sex interaction term*

Model 3 is Model 1 stratified by gender - analysing females only and without including sex as a confounder

Model 4 is Model 1 stratified by gender - analysing males only and without including sex as a confounder

[#]*Results presented are for the Waist Circumference* Sex interaction term*

Subscapular Skinfold Thickness

The analyses utilising subscapular skinfold thickness as the measure of obesity (exposure) provided no evidence of an association with depressive symptoms (Table 5.7 column 3). When an interaction between subscapular skinfold thickness and sex was introduced into the regression model there was no evidence that the results differed by sex. When stratified by sex there was no evidence of an inverse relationship between subscapular skinfold thickness and depressive symptoms in females or males (Table 5.7 column 6 and 7).

Table 5.7 - Results from the linear regression analyses investigating the association between Subscapular Skinfold Thickness (exposure) on depression (Z score) (outcome) at next follow up in the TRAILS cohort

Timepoint	Exposure	Outcome	Model 1 - main model				Model 2 - including interaction (coefficient for interaction effect)			
			n	coeff.	95%CI	p-value	n	coeff[#].	95%CI	p-value
T1 to T2 10y7m to 13y1m	Subscapular Skinfold Thickness	Depression	1367	0.001	(-0.001, 0.003)	0.267	1367	0.001	(-0.003, 0.005)	0.497
T3 to T4 15y10m to 18y7m	Subscapular Skinfold Thickness	Depression	1268	-0.001	(-0.003, 0.001)	0.194	1268	-0.004	(-0.008, 0.0004)	0.057

Model 1 is adjusted for age, sex, previous depression, maternal depression, SEP, alcohol and smoking

*Model 2 is Model 1 plus the inclusion of a Subscapular Skinfold Thickness*Sex interaction term*

[#] *Results presented are for the Subscapular Skinfold Thickness* Sex interaction term*

Table 5.7 continued

Timepoint	Exposure	Outcome	Model 3 - stratified (females)				Model 4 - stratified (males)			
			n	coeff.	95%CI	p-value	n	coeff.	95%CI	p-value
T1 to T2 10y7m to 13y1m	Subscapular Skinfold Thickness	Depression	681	0.001	(-0.002, 0.005)	0.420	686	0.001	(-0.002, 0.003)	0.637
T3 to T4 15y10m to 18y7m	Subscapular Skinfold Thickness	Depression	692	-0.003	-0.006, 0.0001	0.058	576	1E-04	(-0.002, 0.003)	0.892

Model 3 is Model 1 stratified by gender - analysing females only and without including sex as a confounder

Model 4 is Model 1 stratified by gender - analysing males only and without including sex as a confounder

NDIT

In the NDIT cohort as well as BMI, waist circumference and subscapular skinfold thickness were also collected as alternative measures of obesity. Analyses were repeated using these additional measures of obesity and are presented below.

Waist Circumference

When the linear regression analysis was carried out using waist circumference as the measure of obesity (exposure) there was no evidence of an association between waist circumference and depressive symptoms. The regression models at the different time points were each repeated including an interaction between waist circumference and sex. The results provided no evidence that the relationship between waist circumference and depressive symptoms differed by sex (p-value for interaction term between waist circumference and sex at the three time points were 0.180, 0.871 and 0.779 respectively) (Table 5.8 column 5).

Table 5.8 - Results from the linear regression analyses investigating the association between Waist Circumference (exposure) on depression (Z score) (outcome) at next follow up in the NDIT cohort

Timepoint	Exposure	Outcome	Model 1 - main model				Model 2 - including interaction (coefficient for interaction effect)			
			n	coeff.	95%CI	p-value	n	coeff[#].	95%CI	p-value
T1 to T2 12y10m to 13y0m	Waist Circumference	Depression	496	-0.004	(-0.010, 0.002)	0.158	496	-0.008	(-0.019, 0.003)	0.180
T12 to T13 15y2m to 15y7m	Waist Circumference	Depression	433	0.007	(-0.0004, 0.015)	0.064	433	0.001	(-0.015, 0.017)	0.871
T19 to T20 17y0m to 17y1m	Waist Circumference	Depression	416	-0.002	(-0.008, 0.004)	0.595	416	-0.002	(-0.014, 0.010)	0.779

Model 1 is adjusted for age, sex, previous depression, maternal education, maternal profession and alcohol

*Model 2 is Model 1 plus the inclusion of a Waist Circumference*Sex interaction term*

[#] *Results presented are for the Waist Circumference* Sex interaction term*

Table continued

Timepoint	Exposure	Outcome	Model 3 - stratified (females)				Model 4 - stratified (males)			
			n	coeff.	95%CI	p-value	n	coeff.	95%CI	p-value
T1 to T2 12y10m to 13y0m	Waist Circumference	Depression	263	-0.010	(-0.019, -0.001)	0.024	233	-0.001	(-0.008, 0.007)	0.838
T12 to T13 15y2m to 15y7m	Waist Circumference	Depression	222	0.010	(-0.001, 0.020)	0.083	211	0.008	(-0.004, 0.020)	0.199
T19 to T20 17y0m to 17y1m	Waist Circumference	Depression	218	-0.003	(-0.012, 0.007)	0.562	198	4E-04	(-0.008, 0.008)	0.926

Model 3 is Model 1 stratified by gender - analysing females only and without including sex as a confounder

Model 4 is Model 1 stratified by gender - analysing males only and without including sex as a confounder

Subscapular Skinfold Thickness

There was no evidence of an association between subscapular skinfold thickness and depressive symptoms. When an interaction between subscapular skinfold thickness and sex was introduced into the regression model there was no evidence that the results differed by sex (Table 5.9 column 5). When the analysis was stratified by sex there was no evidence of an association between subscapular skinfold thickness and depressive symptoms in females or in males.

Table 5.9 - Results from the linear regression analyses investigating the association between Subscapular Skinfold Thickness (exposure) on depression (Z score) (outcome) at next follow up in the NDIT cohort

Timepoint	Exposure	Outcome	Model 1 - main model				Model 2 - including interaction (coefficient for interaction effect)			
			n	coeff.	95%CI	p-value	n	coeff[#].	95%CI	p-value
T1 to T2 12y10m to 13y0m	Subscapular Skinfold Thickness	Depression	495	-0.005	(-0.016, 0.006)	0.385	495	-0.01	(-0.036, 0.007)	0.189
T12 to T13 15y2m to 15y7m	Subscapular Skinfold Thickness	Depression	432	0.014	(-0.0003, 0.028)	0.056	432	0.008	(-0.022, 0.039)	0.588
T19 to T20 17y0m to 17y1m	Subscapular Skinfold Thickness	Depression	415	0.001	(-0.008, 0.010)	0.855	415	0.002	(-0.016, 0.019)	0.833

Model 1 is adjusted for age, sex, previous depression, maternal education, maternal profession and alcohol

*Model 2 is Model 1 plus the inclusion of a Subscapular Skinfold Thickness*Sex interaction term*

[#] *Results presented are for the Subscapular Skinfold Thickness*Sex interaction term*

Table continued

Timepoint	Exposure	Outcome	Model 3 - stratified (females)				Model 4 - stratified (males)			
			n	coeff.	95%CI	p-value	n	coeff.	95%CI	p-value
T1 to T2 12y10m to 13y0m	Subscapular Skinfold Thickness	Depression	262	-0.019	(-0.037, -0.0002)	0.047	233	0.001	(-0.012, 0.014)	0.889
T12 to T13 15y2m to 15y7m	Subscapular Skinfold Thickness	Depression	222	0.019	(-0.001, 0.039)	0.065	210	0.012	(-0.011, 0.036)	0.302
T19 to T20 17y0m to 17y1m	Subscapular Skinfold Thickness	Depression	217	0.002	(-0.013, 0.016)	0.832	198	0.001	(-0.010, 0.013)	0.799

Model 3 is Model 1 stratified by gender - analysing females only and without including sex as a confounder

Model 4 is Model 1 stratified by gender - analysing males only and without including sex as a confounder

5.1.3. Generalized Estimating Equations

BMI

Following on from the linear regression analyses, GEEs were used to model the repeated exposure-outcome association – in other words, the average effect of BMI on future depressive symptoms (Table 5.10). The results of the GEE analysis from two of the three cohorts (ALSPAC and NDIT) suggested that there was a positive association between lagged BMI and future depressive symptoms (Table 5.10). The greater the increase in BMI over time, the greater the increase in later depressive symptom score. In the ALSPAC cohort, averaged across the population, a one unit increase in lagged BMI was associated with a 0.014 standard deviation increase in depressive symptom score at the next time point (i.e. the effect of BMI on later level of depression) (95% CI 0.003, 0.025, p-value 0.010). In the NDIT cohort a one unit increase in lagged BMI was associated with a 0.032 standard deviation increase in depressive symptoms (95% CI 0.007, 0.058, p-value 0.013). In the TRAILS cohort however, there was no evidence of an association between lagged BMI and depressive symptoms (lagged BMI was associated with a -0.002 SD decrease in depressive symptom score, 95% CI -0.007, 0.002, p-value 0.312).

The GEE model was repeated in each cohort including an interaction term between lagged BMI and sex in order to test for a difference in the relationship between BMI and depression in males and females. There was evidence of an interaction between BMI and sex in the ALSPAC cohort (females had a stronger positive association between BMI and depressive symptoms compared with males: interaction coefficient 0.010, 95% CI 0.0002, 0.020, p-value 0.046) but not in the TRAILS (p-value for interaction term 0.745) or NDIT (p-value for interaction term 0.783) cohorts.

The GEE results from the different cohorts were meta-analysed to provide a pooled estimate across the three cohorts. The pooled estimates provided no evidence of an

association between lagged BMI and later depressive symptoms when the analysis was carried out on all participants or when stratified by gender (Table 5.11). There was however considerable heterogeneity, 85% of the variation across the studies was due to heterogeneity (p-value 0.001). When the meta-analysis was stratified by sex it could be seen that the heterogeneity was observed in females, but for males there is very little heterogeneity between studies and both when considered separately by cohort and pooled there was no evidence of association. As such the results of the GEE analyses in the different cohorts should be considered separately rather than drawing conclusions from the meta-analysed pooled estimates for females.

Table 5.10 - Results from the GEE analyses investigating the association between lagged BMI (exposure) on depression (Z score) (outcome) at next follow up

	ALSPAC				TRAILS				NDIT			
Variable	**n**	**Coeff.**	**95% CI**	**p-value**	**n**	**Coeff.**	**95% CI**	**p-value**	**n**	**Coeff.**	**95% CI**	**p-value**
Lagged BMI	4397	0.014	(0.003, 0.025)	0.010	1847	-0.002	(-0.007, 0.002)	0.312	513	0.032	(0.007, 0.058)	0.013
Sex Interaction												
Lagged BMI*Female	4397	0.010	(0.0002, 0.020)	0.046	1847	0.001	(-0.004, 0.005)	0.745	513	-0.003	(-0.023, 0.017)	0.783
Males												
Lagged BMI	2148	0.0004	(-0.014, 0.015)	0.959	892	-0.003	(-0.009, 0.003)	0.314	242	0.016	(-0.016, 0.047)	0.335
Females												
Lagged BMI	2249	0.023	(0.008, 0.037)	0.002	955	-0.001	(-0.007, 0.005)	0.766	271	0.052	(0.013, 0.090)	0.009

Table 5.11 - Meta-analysis of the GEE results of the investigation of the association between lagged BMI (kg/m²) (exposure) on depression (Z score) (outcome) at next follow up

	Estimates				Heterogeneity	
Variable	**Coeff.**	**95% CI**	**Weight**	**p-value**	**I-squared**	**p-value**
Model1 - main model						
Lagged BMI						
ALSPAC	0.014	(0.003, 0.025)	36.66			
TRAILS	-0.002	(-0.007, 0.002)	41.62			
NDIT	0.032	(0.007, 0.058)	21.72			
Pooled	0.011	(-0.006, 0.028)	100	0.195	85.2%	0.001
Model 2 - interaction model						
Lagged BMI*Female						
ALSPAC	0.010	(0.0002, 0.020)	28.87			
TRAILS	0.001	(-0.004, 0.005)	61.44			
NDIT	-0.003	(-0.023, 0.017)	9.69			
Pooled	0.003	(-0.004, 0.010)	100	0.365	36.9%	0.205
Model 3 - stratified analysis						
Females						
Lagged BMI						
ALSPAC	0.023	(0.008, 0.037)	37.33			
TRAILS	-0.001	(-0.007, 0.005)	41.87			
NDIT	0.052	(0.013, 0.090)	20.80			
Pooled	0.019	(-0.006, 0.043)	100	0.132	86.1%	0.001
Model 4 - stratified analysis						
Males						
Lagged BMI						
ALSPAC	0.0004	(-0.014, 0.015)	14.62			
TRAILS	-0.003	(-0.009, 0.003)	82.17			
NDIT	0.016	(-0.016, 0.047)	3.21			
Pooled	-0.002	(-0.008, 0.004)	100	0.471	0.0%	0.489

Other measures of obesity

Waist Circumference

The ALSPAC and NDIT cohorts collected repeated measures data on waist circumference, hence the GEE analysis was repeated using waist circumference as the exposure variable. As was the case when investigating BMI, in both the ALSPAC and NDIT cohorts there was evidence of a positive association between lagged waist circumference and later depressive symptoms. In the ALSPAC cohort, averaged across the population, a one centimetre increase in lagged waist circumference was associated with a 0.005 standard deviation increase in later depressive symptoms (95% CI 0.001, 0.009, p-value 0.013) (Table 5.12). In the NDIT cohort a one centimetre increase in lagged waist circumference was associated with a 0.010 standard deviation increase in depressive symptoms (95% CI 0.002, 0.018, p-value 0.013) (Table 5.12).

When the GEE analyses were repeated including an interaction term between lagged waist circumference and sex, there was evidence in the ALSPAC cohort that there was a greater positive association in females compared to males (interaction coefficient 0.006, 95% CI 0.002, 0.010, p-value 0.005) (Table 5.12). When stratified by sex, the ALSPAC cohort showed no evidence of an association between lagged waist circumference and later depressive symptoms in males, in females however there was strong evidence of an association between lagged waist circumference and later depressive symptom score (a one centimetre increase in lagged waist circumference was associated with a 0.009 SD increase in later depressive symptoms in females, 95% CI 0.003, 0.015, p-value 0.003). In the NDIT cohort there was very weak evidence of an association between lagged waist circumference and later depressive symptoms in males: a one centimetre increase in lagged waist circumference was associated with a 0.011 standard deviation increase in depressive symptoms (95% CI 0.00004, 0.022, p-value 0.049). In females there was no evidence of an association

between lagged waist circumference and depressive symptoms as the 95% CI was wider than in males and encompassed the null (95% CI -0.002 to 0.021).

Table 5.12 - Results from the GEE analyses investigating the association between lagged Waist Circumference (cm) (exposure) on depression (Z score) (outcome) at next follow up

	ALSPAC				NDIT			
Variable	**n**	**Coeff.**	**95% CI**	**p-value**	**n**	**Coeff.**	**95% CI**	**p-value**
Lagged Waist Circumference	4376	0.005	(0.001, 0.009)	0.013	513	0.010	(0.002, 0.018)	0.013
Sex Interaction								
Waist Circumference*Female	4376	0.006	(0.002, 0.010)	0.005	513	-0.002	(-0.010, 0.005)	0.519
Males								
Lagged Waist Circumference	2140	0.001	(-0.004, 0.007)	0.577	242	0.011	(0.00004, 0.022)	0.049
Females								
Lagged Waist Circumference	2236	0.009	(0.003, 0.015)	0.003	271	0.010	(-0.002, 0.021)	0.094

The GEE results from the two cohorts were meta-analysed to provide a pooled estimate of the association between lagged waist circumference and later depressive symptoms. When the estimates were pooled across the cohorts the results provided strong evidence that lagged waist circumference was associated with later depressive symptoms (Table 5.13). A one centimetre increase in lagged waist circumference was associated with a 0.006 standard deviation increase in later depression (95% CI 0.002, 0.011, p-value 0.004). The meta-analysis was largely weighted towards the ALSPAC analysis due to the large sample size and narrow 95% CI compared to the NDIT analysis. There was no evidence of heterogeneity in the meta-analysis (p-value 0.271) and the I-squared percentage was low (17.5%).

When the pooled analysis was repeated including an interaction term between lagged waist circumference and sex there was no evidence that the relationship between waist circumference and depressive symptoms differed by sex (Table 5.13). When the analysis was stratified by sex, there was no evidence of an association between lagged waist circumference and later depressive symptoms in males. In females however, the pooled results suggested a positive association between lagged waist circumference and later depressive symptoms: a one centimetre increase in lagged waist circumference was associated with a 0.009 standard deviation increase in depressive symptom score (95% CI 0.004, 0.014, p-value 0.001).

Table 5.13 - Meta-analysis of the GEE results of the investigation of the association between lagged Waist Circumference (exposure) on depression (Z score) (outcome) at next follow up

	Estimates				Heterogeneity	
Variable	**Coeff.**	**95% CI**	**Weight**	**p-value**	**I-squared**	**p-value**
Model 1 - main model						
Lagged Waist Circumference						
ALSPAC	0.005	(0.001, 0.009)	74.50	0.013		
NDIT	0.010	(0.002, 0.018)	25.50	0.013		
Pooled	0.006	(0.002, 0.011)	100	0.004	17.6%	0.271
Model 2 - interaction model						
Waist Circumference*Female						
ALSPAC	0.006	(0.002, 0.010)	55.94	0.005		
NDIT	-0.002	(-0.010, 0.005)	44.06	0.519		
Pooled	0.002	(-0.006, 0.011)	100	0.570	75.0%	0.046
Model 3 - stratified analysis						
Females						
Lagged Waist Circumference						
ALSPAC	0.009	(0.003, 0.015)	78.87	0.003		
NDIT	0.010	(-0.002, 0.021)	21.13	0.094		
Pooled	0.009	(0.004, 0.014)	100	0.001	0.0%	0.906
Model 4 - stratified analysis						
Males						
Lagged Waist Circumference						
ALSPAC	0.001	(-0.004, 0.007)	63.42	0.577		
NDIT	0.011	(0.00004, 0.022)	36.58	0.049		
Pooled	0.005	(-0.004, 0.014)	100	0.281	58.0%	0.123

5.1.4. Cross-lagged Structural Equation Modelling

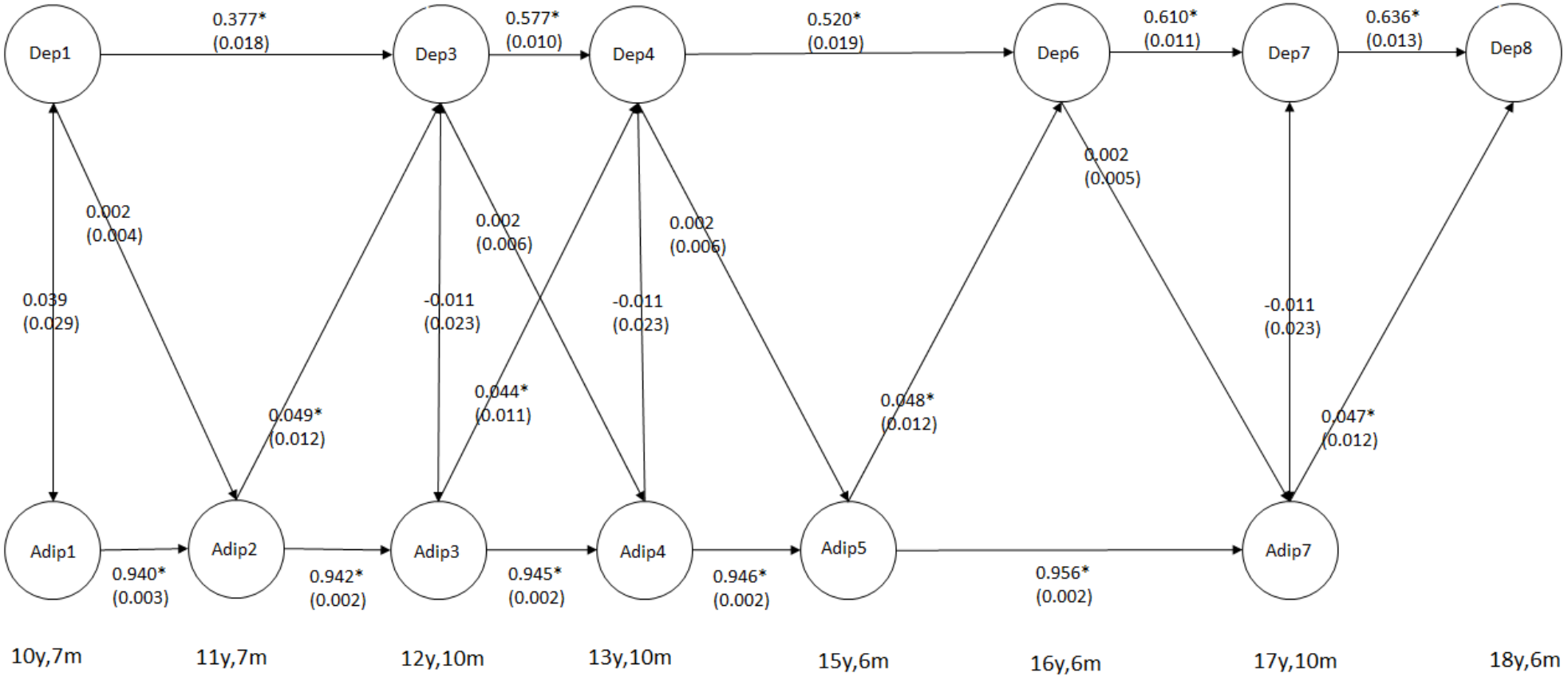
A cross-lagged structural equation model (SEM) was fitted to the data from each of the cohorts in order to test the potential bi-directionality of the relationship between obesity and depression in adolescents (see Section 3.7.1).

ALSPAC

The results from the SEM model using ALSPAC data suggested that there was strong evidence that later depression latent trait score was associated with previous depression latent trait score (an autoregressive relationship in depressive symptoms). For example, a one standard deviation increase in depression latent trait score at time 3 was associated with a 0.577 standard deviation increase in depression latent trait score at time 4, and the magnitude of this association increased as participants got older (Figure 5.1). There was also evidence that later obesity latent trait score was associated with previous obesity latent trait score (i.e. an autoregressive obesity relationship) and that this relationship was fairly stable over time (Figure 5.1). There was also evidence of a cross-lagged association between obesity latent trait score and depression latent trait score at the next follow up occasion, for example a one standard deviation increase in obesity latent trait score at follow up occasion two was associated with a 0.049 standard deviation (SE 0.012) increase in depression latent trait score at time point three. When looking at the relationship in the other direction however there was no evidence of a cross-lagged association between depression and obesity at the next measurement occasion (e.g. a one standard deviation increase in depression latent trait score at time point 1 was associated with a 0.002 standard deviation increase in obesity latent trait score at time point 2 but with a standard error of 0.004).

When the cross-lagged SEM was carried out separately by sex the same associations were found in females (Figure 5.2) as were found when analysing all participants together (i.e. evidence of autoregressive relationships and a cross-lagged relationship between obesity latent trait and depression latent trait at next follow up but no evidence of a cross-lagged relationship in the other direction). In males however, although there was evidence of autoregressive pathways for both obesity and depression, there was no evidence of a cross-lagged relationship between obesity and depression (Figure 5.3).

Figure 5.1 - Cross lagged SEM in the ALSPAC cohort describing the relationship between obesity (Adip) and depressive (Dep) latent traits

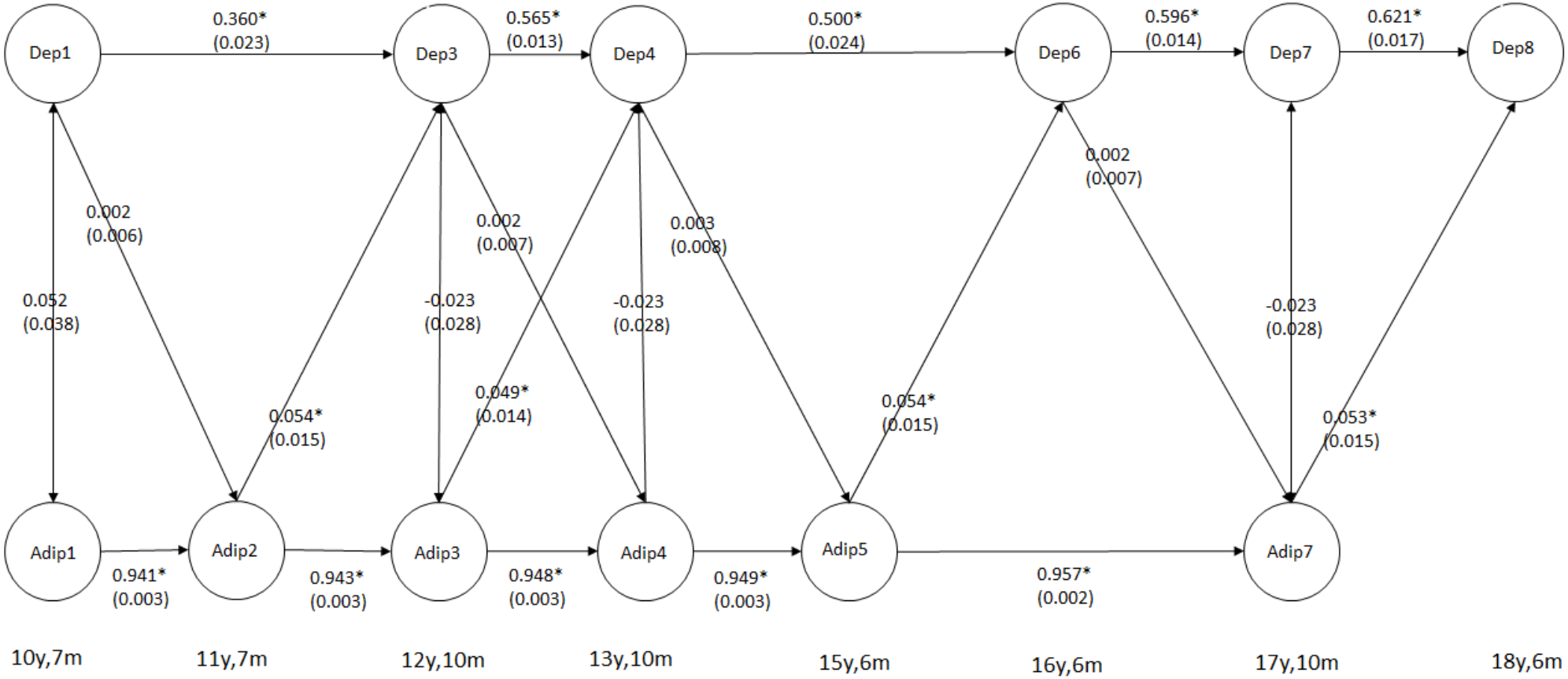


Model adjusted for age, sex, socio-economic position (maternal education and maternal profession) and maternal depression.

Coefficients represent standard deviation change, figures in brackets represent standard error

**represents p-value <0.05*

Figure 5.2 - Cross lagged SEM in the ALSPAC cohort describing the relationship between obesity (Adip) and depressive (Dep) latent traits in females

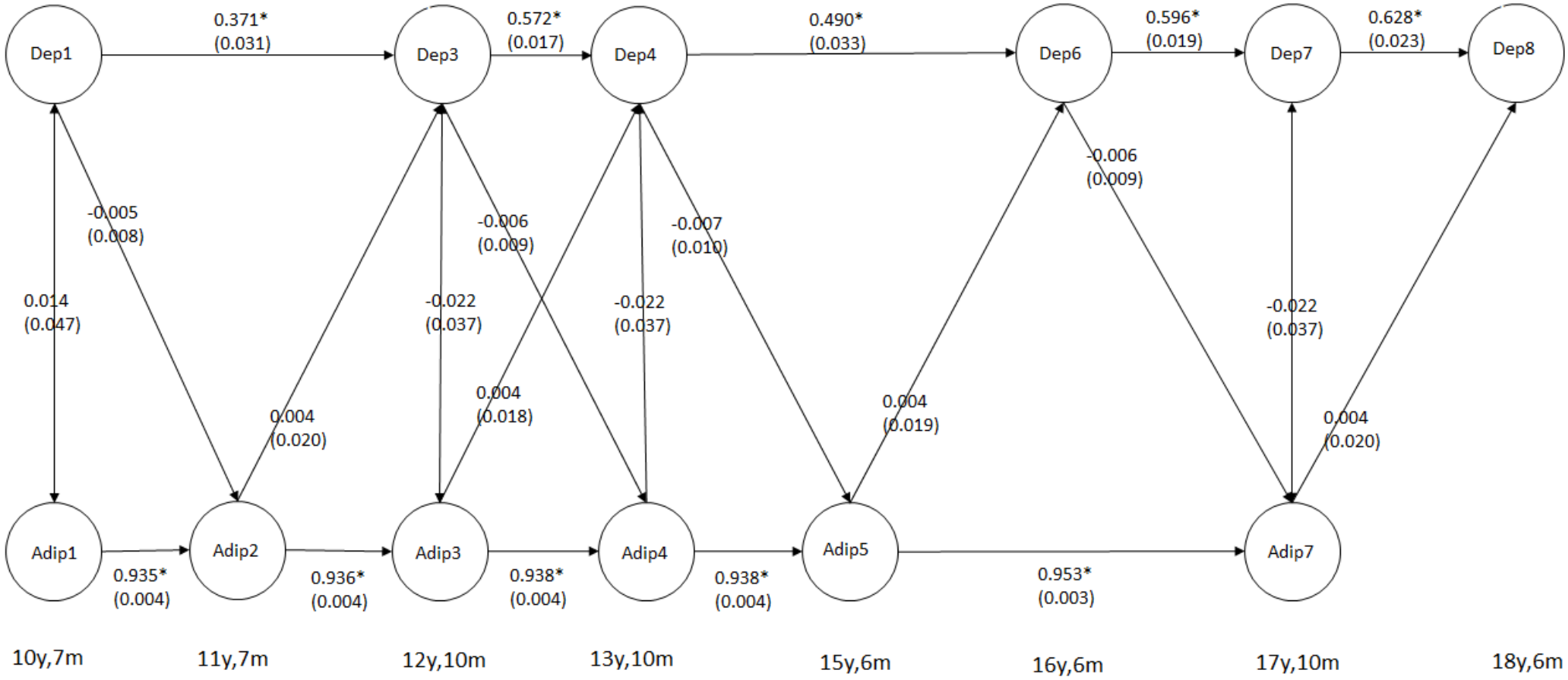


Model adjusted for age, socio-economic position (maternal education and maternal profession) and maternal depression.

Coefficients represent standard deviation change, figures in brackets represent standard error

**represents p-value <0.05*

Figure 5.3 - Cross lagged SEM in the ALSPAC cohort describing the relationship between obesity (Adip) and depressive (Dep) latent traits in males



Model adjusted for age, socio-economic position (maternal education and maternal profession) and maternal depression.

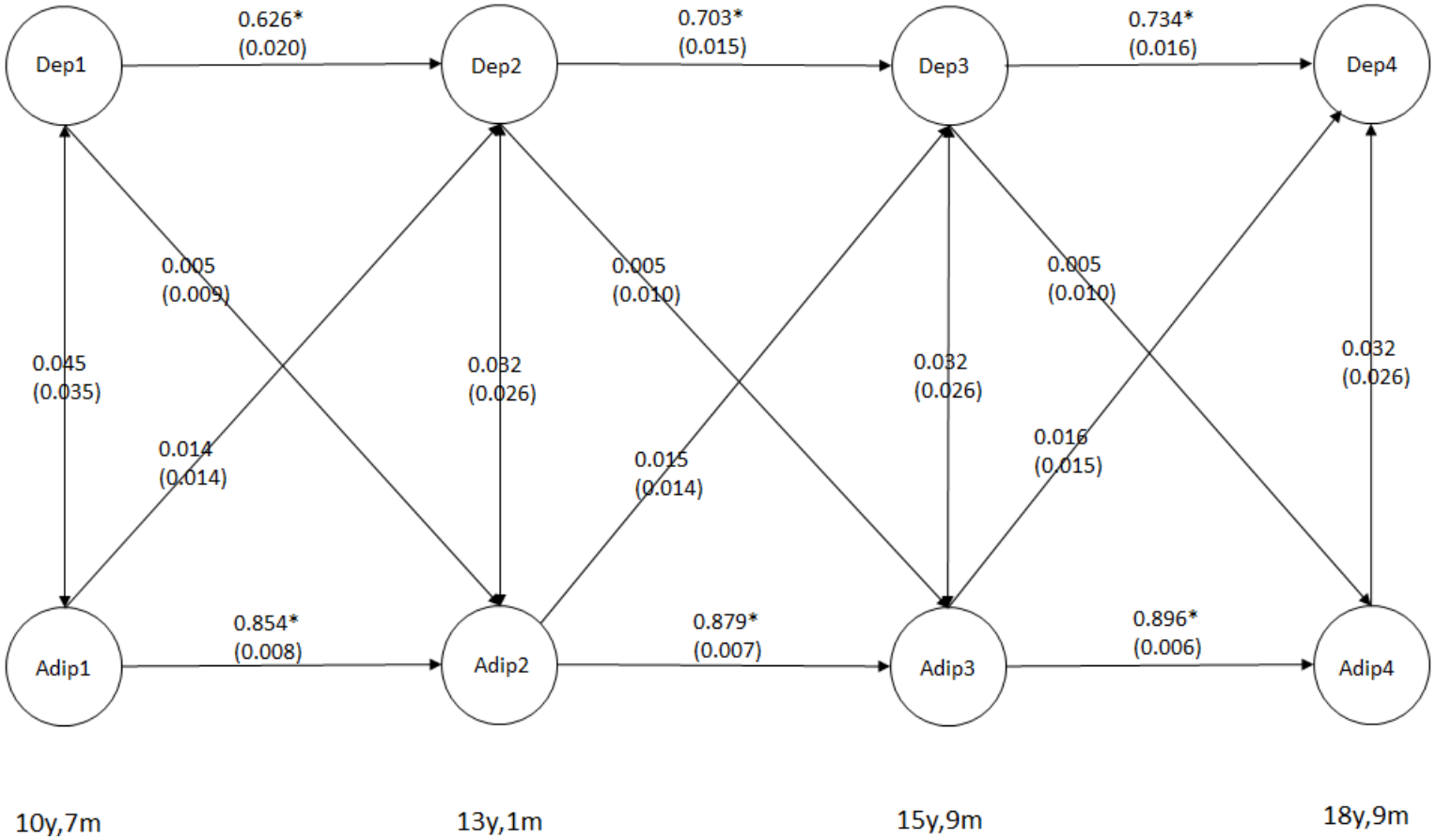
Coefficients represent standard deviation change, figures in brackets represent standard error

**represents p-value <0.05*

TRAILS

The results from the SEM model for the TRAILS cohort suggested that there was evidence of an autoregressive relationship in both obesity and depression latent trait scores. There was however no evidence of a cross-lagged relationship, in other words no evidence that obesity was associated with depression at the next follow up occasion, or vice-versa (Figure 5.4). When the analysis was carried out stratified by gender, in both females (Figure 5.5) and males (Figure 5.6), there was evidence of an autoregressive relationship in both obesity and depression but no evidence of a cross-lagged relationship.

Figure 5.4 - Cross lagged SEM in the TRAILS cohort describing the relationship between obesity (Adip) and depressive (Dep) latent traits

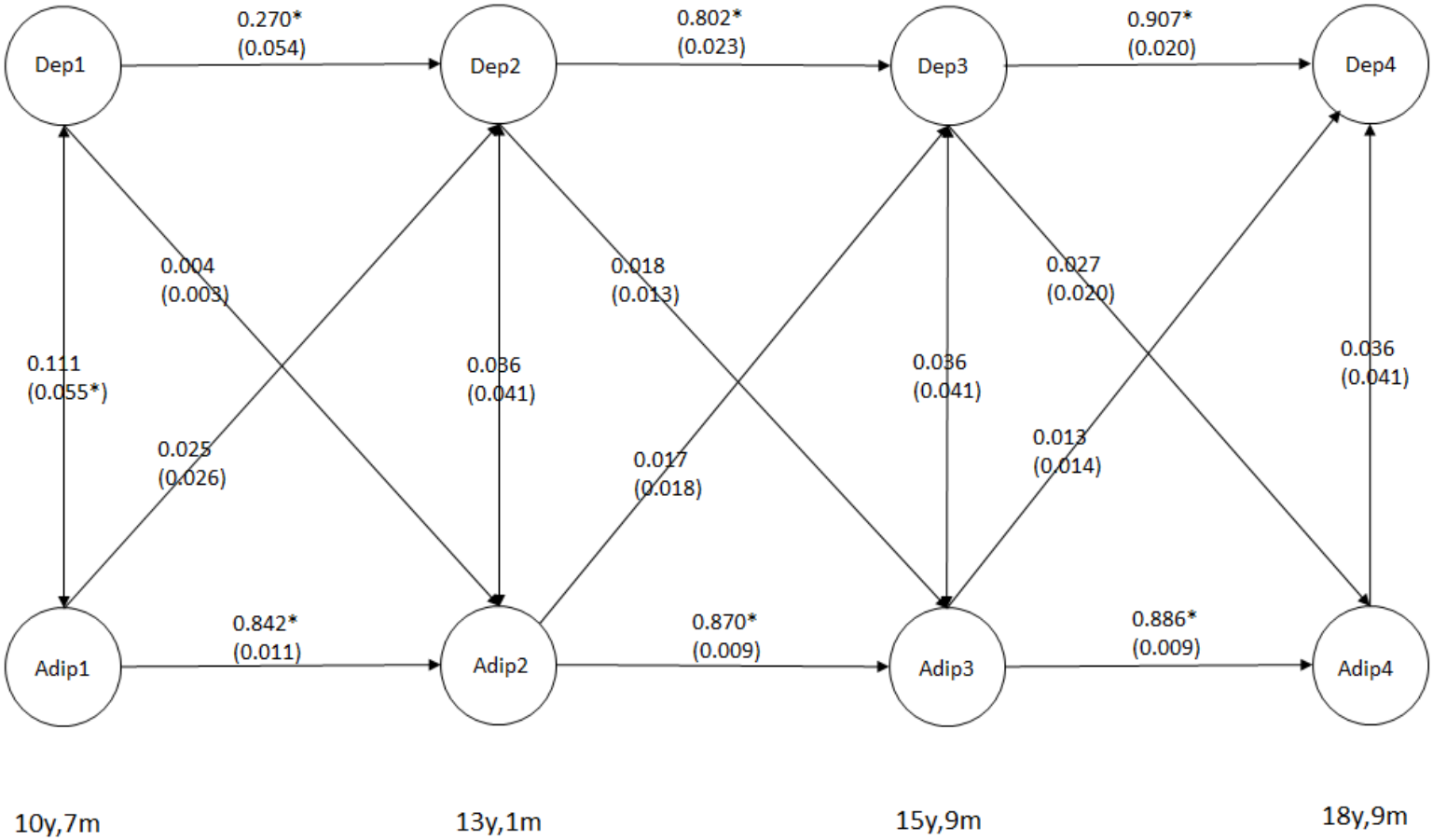


Model adjusted for age, sex, socio-economic position, maternal depression, smoking and alcohol use.

Coefficients represent standard deviation change, figures in brackets represent standard error

**represents p-value <0.05*

Figure 5.5 - Cross lagged SEM in the TRAILS cohort describing the relationship between obesity (Adip) and depressive (Dep) latent traits in females

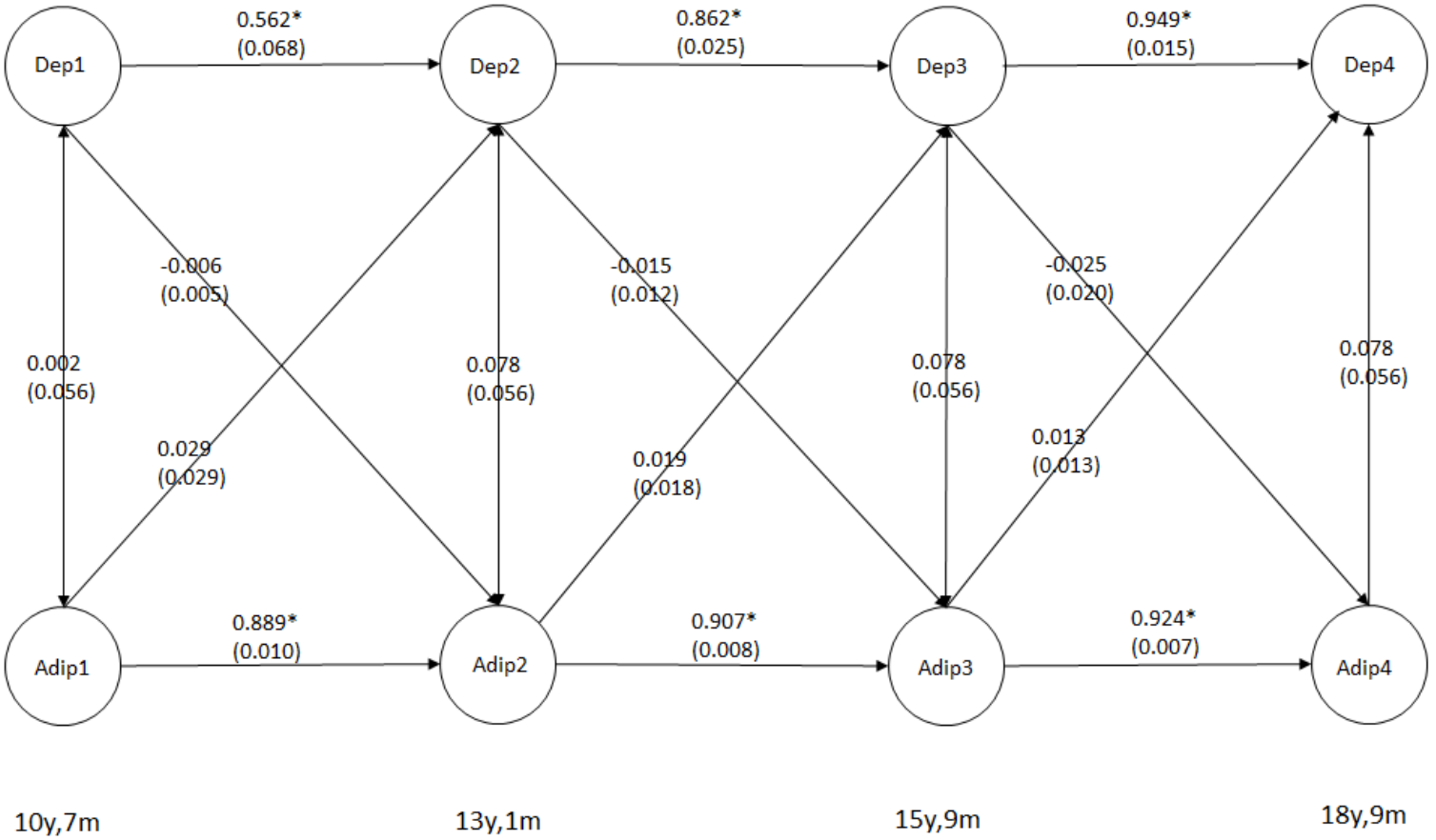


Model adjusted for age, socio-economic position, maternal depression, smoking and alcohol use.

Coefficients represent standard deviation change, figures in brackets represent standard error

**represents p-value <0.05*

Figure 5.6 - Cross lagged SEM in the TRAILS cohort describing the relationship between obesity (Adip) and depressive (Dep) latent traits in males



Model adjusted for age, socio-economic position, maternal depression, smoking and alcohol use.

Coefficients represent standard deviation change, figures in brackets represent standard error

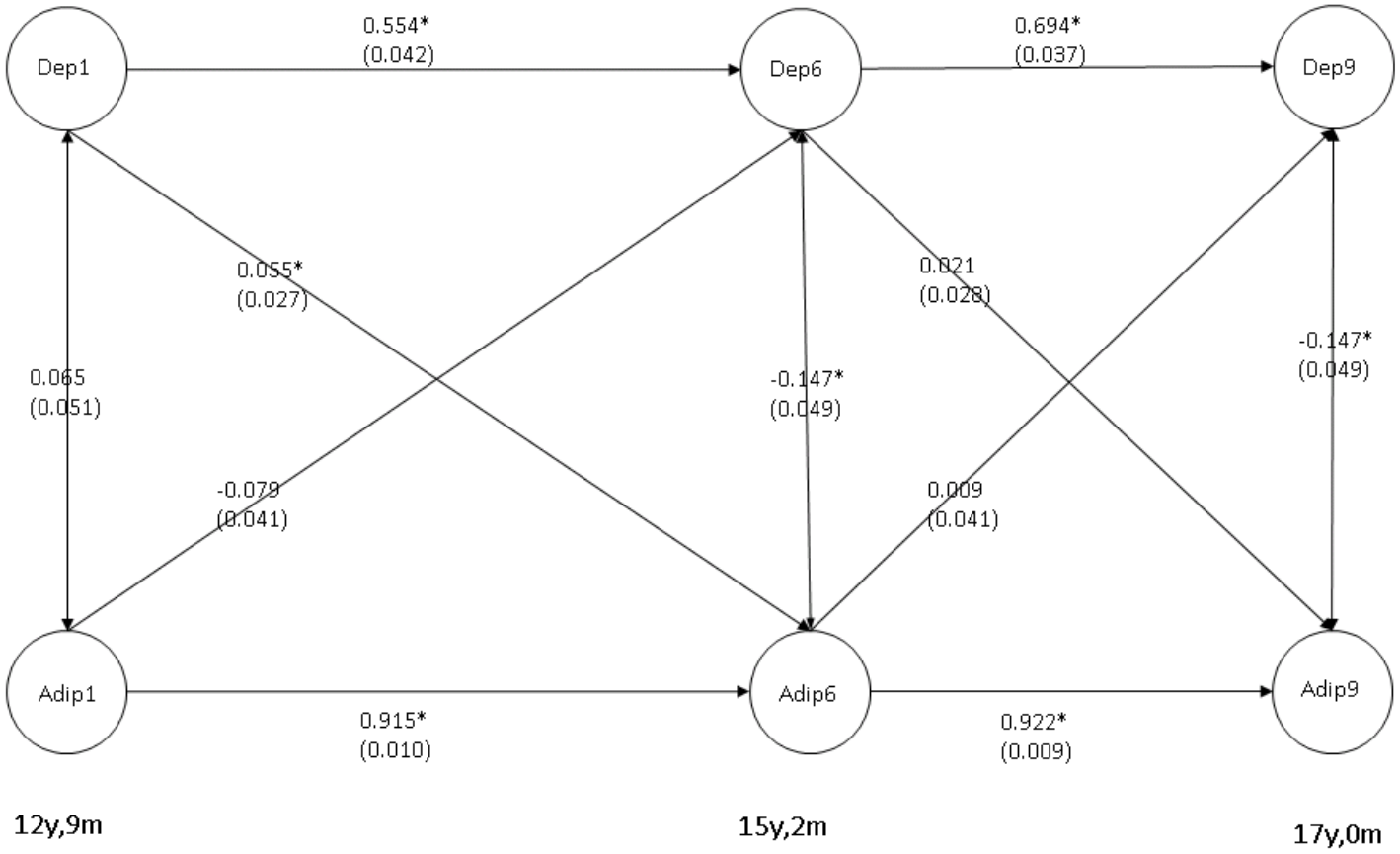
**represents p-value <0.05*

NDIT

In order to achieve model convergence a reduced number of depression measurement time points were used for the NDIT cohort. Instead of depression measured (approximately) every 3 months the SEM only used measures of depression recorded at concurrent time points to measures of obesity. When the cross-lagged SEM was fitted to the NDIT cohort there was evidence of an autoregressive relationship in both obesity and depression. There was also evidence of a cross-lagged relationship between depression at age 12 years 10 months and obesity at 15 years 2 months (Figure 5.7). A one standard deviation increase in depression latent trait score at age 12 years 10 months was associated with a 0.055 (standard error 0.027) increase in obesity latent trait score at age 15 years 2 months. No depression to obesity cross-lagged relationship was observed between the later time points however. There was no evidence of an obesity to future depression cross-lagged relationship.

When the cross-lagged SEM was fitted to females and males separately the autoregressive relationships for obesity and depression were observed in both females (Figure 5.8) and males (Figure 5.9) as for the entire cohort. However, the cross-lagged association between depression at age 12 years 10 months and obesity at 15 years 2 months was observed only in females and not in males (and again only between these two first earliest time points and not between the later time points) (Figure 5.8 and 5.9).

Figure 5.7 - Cross lagged SEM in the NDIT cohort describing the relationship between obesity (Adip) and depressive (Dep) latent traits

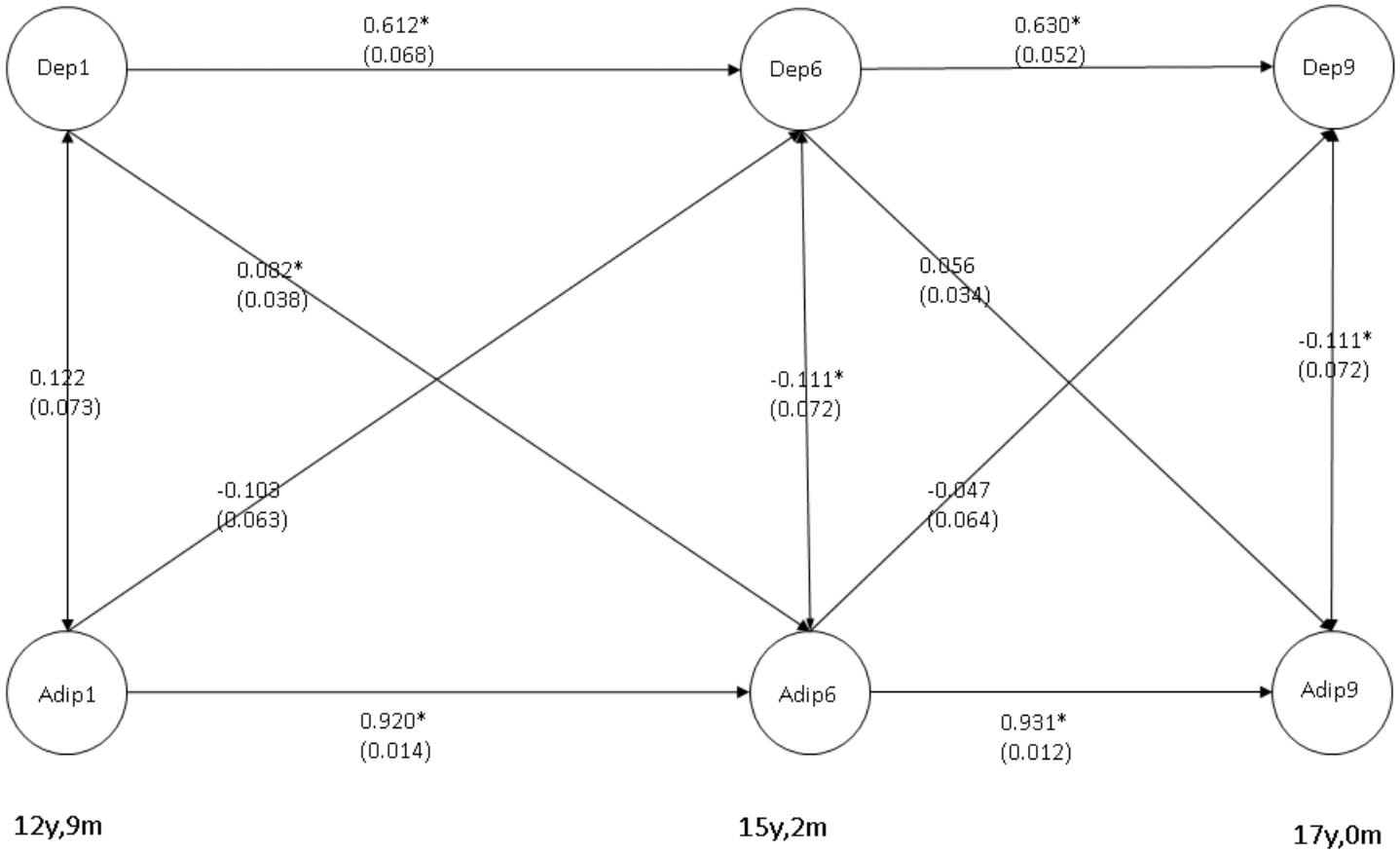


Model adjusted for age, sex, socio-economic position and alcohol use.

Coefficients represent standard deviation change, figures in brackets represent standard error

**represents p-value <0.05*

Figure 5.8 - Cross lagged SEM in the NDIT cohort describing the relationship between obesity (Adip) and depressive (Dep) latent traits in females

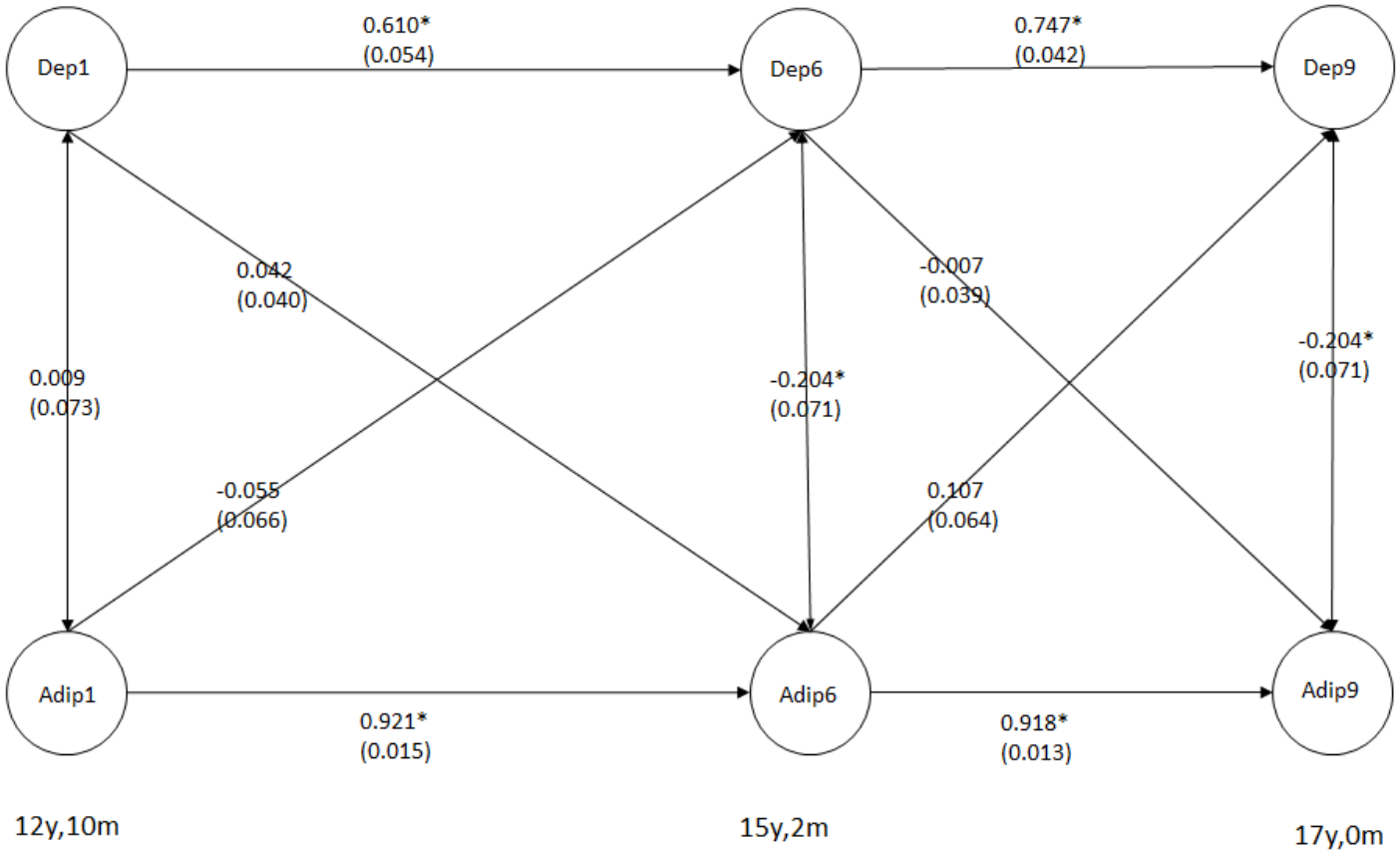


Model adjusted for age, socio-economic position and alcohol use.

Coefficients represent standard deviation change, figures in brackets represent standard error

**represents p-value <0.05*

Figure 5.9 - Cross lagged SEM in the NDIT cohort describing the relationship between obesity (Adip) and depressive (Dep) latent traits in males



Model adjusted for age, socio-economic position and alcohol use.

Coefficients represent standard deviation change, figures in brackets represent standard error

**represents p-value <0.05*

Meta-analysis

In order to summarise the results of the SEM models examining the association between obesity and depression across the cohorts, results from time points with measurements of obesity and depression most closely aligned (and hence of comparable ages) in the ALSPAC and TRAILS cohort, were meta-analysed. The meta-analysed results showed evidence for autoregressive relationships in both depression and obesity. There was also evidence for a cross-lagged association between depression at age 10 years 7 months and obesity at age 13 years: a one standard deviation increase in depression latent trait was associated with a 0.013 (95% CI 0.000, 0.025) standard deviation increase in obesity latent trait. There was also evidence of a cross-lagged association in the other direction; obesity to depression: a one standard deviation increase in obesity latent trait was associated with a 0.035 (95%CI 0.007, 0.077) standard deviation increase in depression latent trait. When males and females were analysed separately there was evidence of an association between level of obesity at age 10 years 7 months and depression at age 13 years for both sexes.

Figure 5.10 - Meta-analysed cross lagged SEM investigating the relationship between obesity and depression

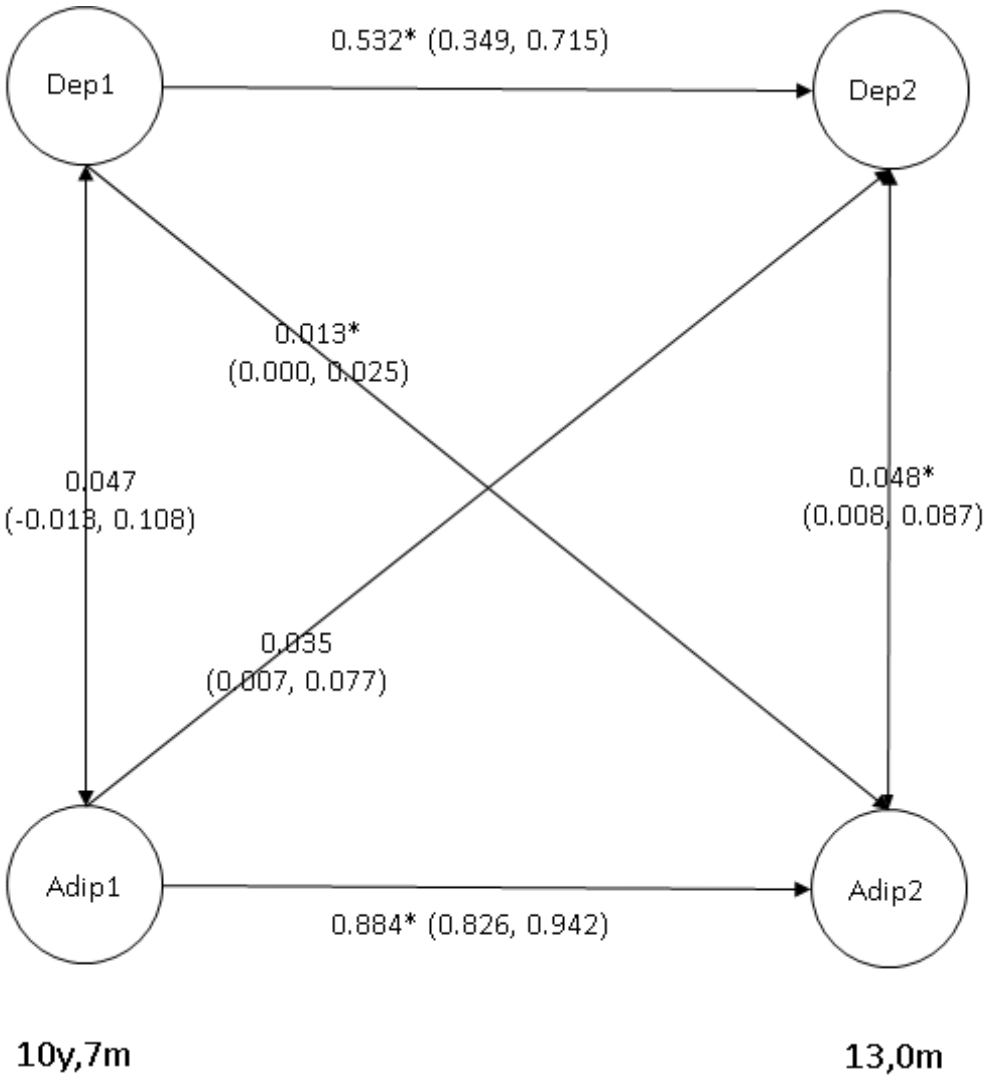


Figure 5.11 - Meta-analysed cross lagged SEM investigating the relationship between obesity and depression, females only

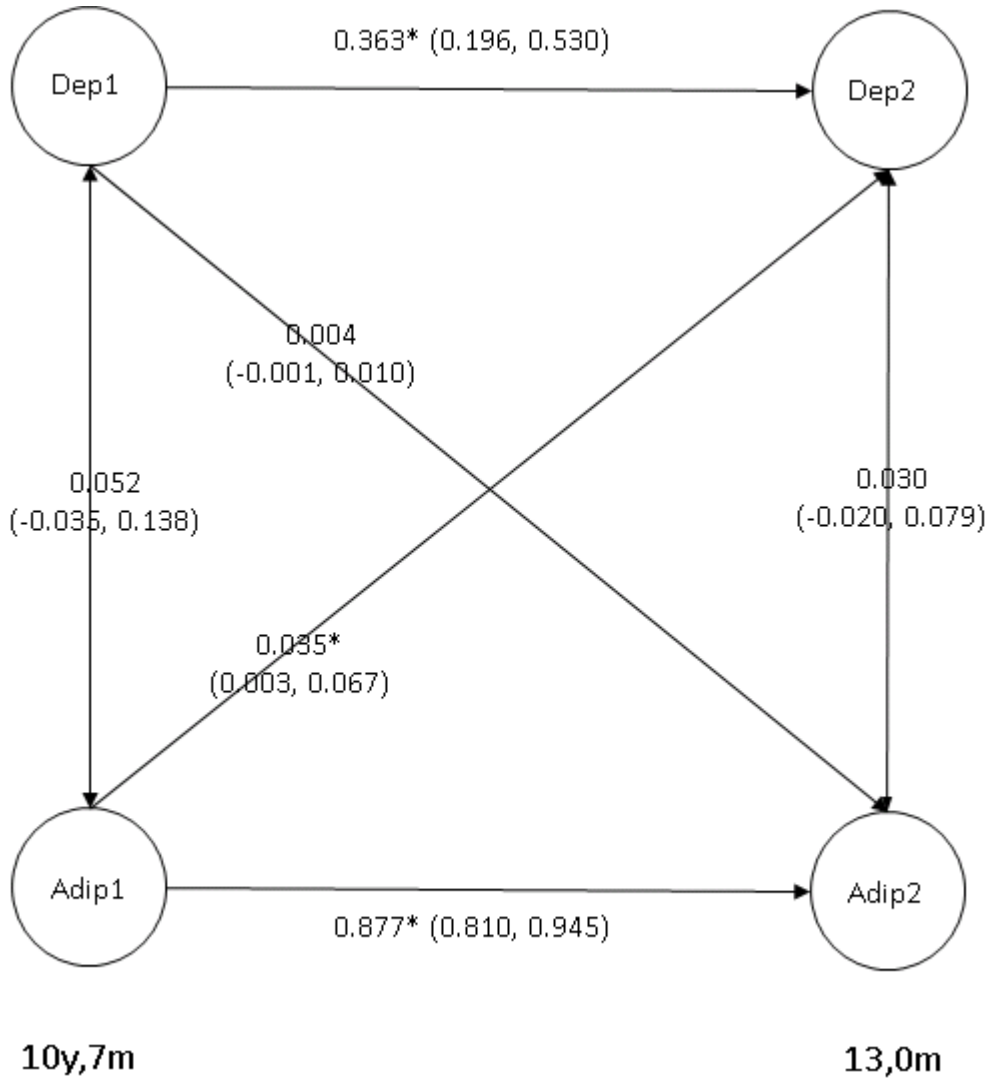
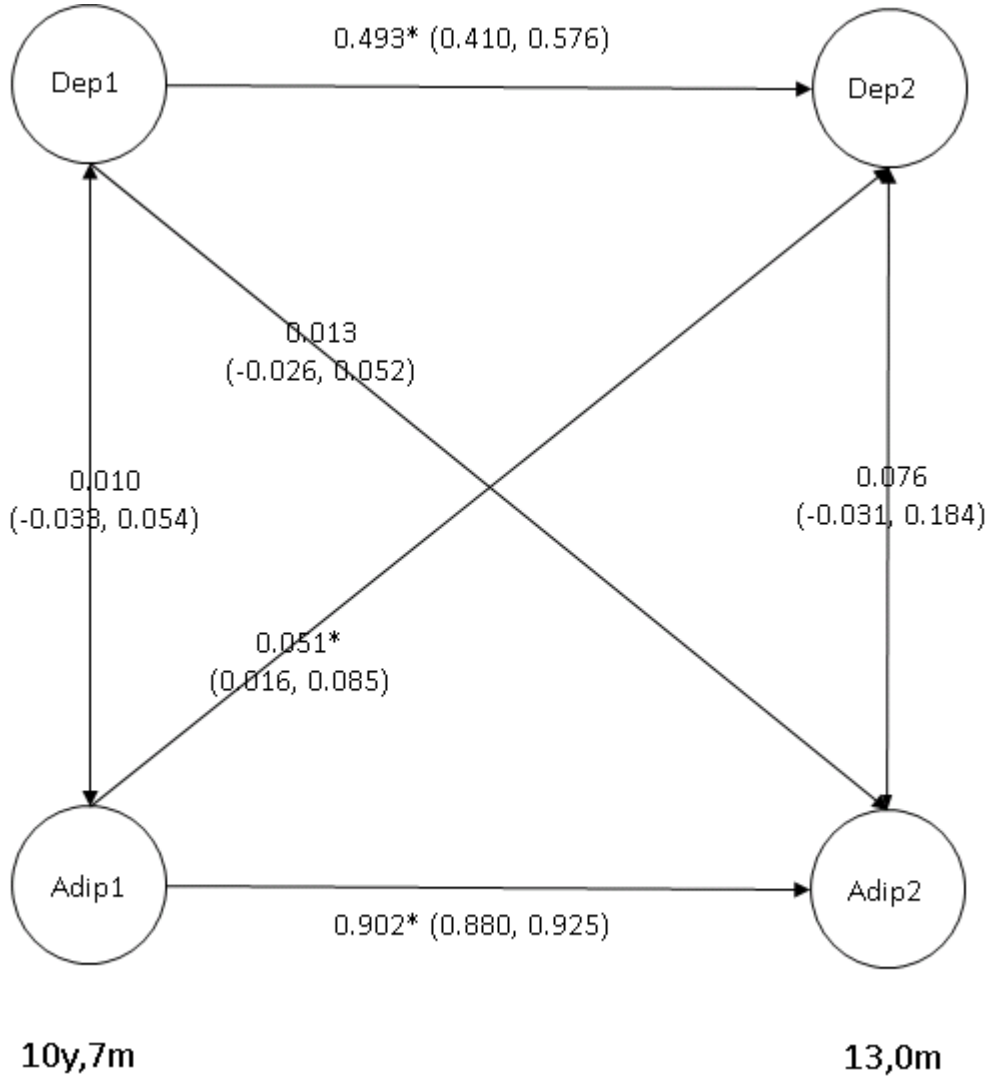


Figure 5.12 - Meta-analysed cross lagged SEM investigating the relationship between obesity and depression, males only



5.1.5. Mendelian Randomization

Mendelian randomization analysis was carried out to analyse the relationship between obesity and depression in adolescence free from the problems of residual confounding. Genetic data for use in the MR analysis was only available in the ALSPAC and TRAILS cohorts. In the ALSPAC cohort, there was evidence of an association between the weighted allele score and BMI (see Appendix 7) and there was no evidence of an association between the weighted allele score and any confounding variables (see Appendix 7). The MR analyses suggested that the

weighted allele score generated was a good genetic instrument for BMI (F-statistics range from 125 to 258, an F-statistic greater than 10 is generally considered to indicate a good instrument) (Table 5.14 column 3). However, there was no evidence of an association between BMI and depressive symptom score at any of the time points (Table 5.14).

Table 5.14 - Mendelian Randomization analysis investigating the relationship between obesity and depression in the ALSPAC cohort

Time point	n	F statistic	coefficient	95% CI	p-value
F10: 10y 8m	5461	258	-0.026	(-0.067, 0.014)	0.205
TF1: 12y 10m	5011	229	0.006	(-0.033, 0.044)	0.771
TF2: 13y 10m	4626	201	0.011	(-0.030, 0.052)	0.612
TF4: 17y 10m	3186	125	-0.001	(-0.046, 0.043)	0.950

In the TRAILS cohort, the Mendelian randomization analyses suggested again that the weighted allelic score generated was a good genetic instrument for BMI (F-statistics range from 89 to 188) (Table 5.15 column 3) but that there was no evidence of an association between BMI and level of depression at any of the time points (Table 5.15).

Table 5.15 - Mendelian Randomization analysis investigating the relationship between obesity and depression in the TRAILS cohort

Time point	n	F statistic	coefficient	95% CI	p-value
T1: 10y 7m	1847	188	0.009	(-0.035, 0.020)	0.763
T2: 13y 1m	1522	173	0.014	(-0.018, 0.032)	0.545
T3: 15y 10m	1364	103	0.004	(-0.044, 0.022)	0.638
T4: 18y 7m	1243	89	-0.003	(-0.044, 0.037)	0.847

In the ALSPAC and TRAILS cohort there was one time point where the ages of the participants were comparable between the two cohorts: ALSPAC TF1 mean age 12 years 10 months and TRAILS T2 mean age 13 years and 1 month. The results from the Mendelian Randomization analyses at these time points were meta-analysed to produce a pooled result across cohorts. The results of the meta-analysis showed no evidence of an association between BMI and level of depression (pooled coefficient: 0.013, 95% CI -0.003, 0.028, p-value: 0.704).

5.2. Summary of findings

There was some evidence of a positive association between obesity and future depression. When stratified by sex, there was no evidence of an association between obesity and later depression in males, this was consistent across the three cohorts with no evidence of heterogeneity between studies when results were meta-analysed (see section 5.1.3). There was however evidence of a positive association between obesity and depression in females, but this was not consistent across all three cohorts. Focussing on the findings from the most robust analyses, using a SEM approach in the ALSPAC cohort there was evidence that an increase in obesity latent trait score was associated with an increase in depression latent trait score at the next time point (a 1 SD increase in obesity latent trait score was associated with a 0.049 SD increase in depression latent trait score at the next time point). When the ALSPAC SEM analysis was stratified by sex the same associations were found in females (i.e. a positive cross-lagged association between obesity and later depression), however in males there was no evidence of any cross-lagged associations, in either direction. In the TRAILS and NDIT cohorts there was no evidence of any cross-lagged associations either overall or when stratified by sex, although in the TRAILS cohort the coefficients were in the same direction but the confidence interval crossed the null.

Where possible the results of the analyses in the different cohorts were meta-analysed to produce a pooled estimate. When meta-analysed there was a large amount of heterogeneity between the cohorts when analysing females only, suggesting that the findings in females of the different cohorts should be considered separately not pooled (see Section 5.1.3). The heterogeneity observed may be due to differences between the three cohorts (e.g. in terms of differences in the populations studied) and/or due to differences in the length of follow up that was analysed in the different cohorts.

As I have discussed previously there are many potential problems when trying to investigate a causal relationship using observational data (see section 3.7). The availability of genetic data in two of the three cohorts (ALSPAC and TRAILS) provided the opportunity for causal modelling using MR. Using MR analysis there was no evidence of an association between BMI and future depressive symptoms in either the ALSPAC or TRAILS cohorts. The size and direction of the coefficients were consistent with the observational findings. However, it should be highlighted that the statistical power of these MR analyses was very low.

5.3. Strengths and limitations

In this study, data from three population based cohort studies which have collected longitudinal, repeated measures of depression and several different (objective) measures of obesity in adolescence were analysed, using appropriate statistical methods, in order to investigate the potential causal relationship between obesity and depression in adolescence. There are however some limitations of the current study. One limitation is confounding, although certain known confounders were adjusted for, the study is limited by what confounders were measured in the data sets. There is also the possibility of measurement error within the confounders that were adjusted for. If a confounding variable is measured with error, then adjusting for it in analyses will not remove all the confounding effect – so residual confounding will remain. If the error is systematic, for example in the self-report of amount of alcohol consumed it may be socially desirable for participants to under report the number of drinks they have consumed, this has the potential to introduce bias into the study.

The current study is also limited in the measurement of depression. Depression is a construct that we cannot directly observe, as such we must make the assumption that the instruments used to measure depression are truly measuring the construct

in the way we believe. The different cohorts used different instruments to measure depression. The assumption here is that all three depression instruments are measuring the same underlying concept. Although the different measures were Z-transformed to compare findings across the cohorts this is still making the assumption that a standard deviation change in one instrument means the same to an individual as a standard deviation change in another depression instrument. This may be a potential source of heterogeneity observed between the cohorts.

Another potential issue that may be important and may (at least partly) explain the heterogeneity observed in females between the cohorts could be the social context/environment/group of the participants. Previous studies have shown that there are cultural differences in the tolerance of, or even preference for, higher body weights and differing body shapes [160, 161]. This may be of relevance in a potential obesity – depression association in females (particularly if the causal pathway is psycho-social in nature) due to differences in how weight status is perceived by both the participants themselves and those around them, leading to differences in levels of body satisfaction, bullying/peer victimisation. As such, not accounting for social context/environment/group (i.e. through adjusting for participant race/ethnicity as a potential proxy for cultural relationship with body weight/shape [160, 161]) may at least in part explain the inconsistent findings of this investigation.

A potential limitation of this study is that in all the analyses examining whether there was an association between obesity and later depression, obesity was considered as a continuous variable, with a linear association with depression. It is possible that any relationship may not in fact be linear in nature, for example it has been suggested that a quadratic "U shaped" relationship may be present whereby depression is related to being both over- and underweight [162]. To investigate this a sensitivity analysis was carried out. In the ALSPAC cohort the linear regression models were repeated including a BMI squared term in the models (see Sections 3.7.1 and 5.1.1). There was no evidence of an association between quadratic BMI and

depressive symptom score at the next follow up occasion and hence this was not explored further in later models.

Although there was no evidence of a quadratic relationship this does not necessarily mean that conceptualizing obesity as a continuous variable, with a linear association with depression, is the best way to think about obesity as an exposure variable. For example it is possible that there could be a threshold effect, i.e. there may be no effect of obesity on depression until a certain level of obesity is reached. Or perhaps it is the size of change in obesity that is important; maybe a small increase in obesity has no effect on depression but a large change may have an effect. Alternatively, it may be more appropriate to conceptualise obesity as a chronic stressor, i.e. use longitudinal repeated measures to identify different "classes" of obesity (for example never obese, early obesity, later obesity, persistently obese) in a latent class analysis and investigate the impact of class membership and/or the impact of change in class membership on the development of depression [42].

If we consider obesity as more of a chronic life stressor this also opens the door to questions relating to gene-environment interactions and other factors, such as resilience. It has been suggested that a number of genetic and psychosocial factors are associated with increased/decreased sensitivities to adversity, particularly in females, as such the impact of obesity could be moderated by certain inherited genetic factors [163]. A similar argument could be made relating to resilience, some individuals who are at high-risk for depression do not develop it. Factors such as high intelligence, social support (through friends and family) and cognitive styles, such as the explanatory style, whereby individuals are able to view problems as temporary issues that only affect limited areas of their life and do not automatically blame themselves for them, are believed to be protective and confer resilience against risk factors for depression [163, 164].

The first strength of this study was that longitudinal data was available allowing investigation into direction of causality between obesity and depression in adolescence rather than examining a cross-sectional association between these variables. Another strength of this study was that data was available on a large number of participants across multiple time points in the three cohorts. Having data available from three different cohorts allowed comparison of findings across three different cohorts, although the confounding structures were similar in the cohorts (see Chapter 4) limiting our ability to triangulate results. There was also a large amount of information on important confounders, including confounders that have not been adjusted for in previous studies (e.g. maternal depression) (see Section 2.1.1).

The information available on potential confounders also allowed a sensitivity analysis into the effect of puberty to be carried out. There is prior evidence that in females pubertal stage may be associated with depression [142, 143], and there is also evidence that puberty is associated with changes in body composition and BMI [165, 166]. Therefore one potential explanation for the inconsistency in findings in prior analyses and the heterogeneity observed in females between cohorts in the current study may be lack of adjustment for puberty as a confounder. Sensitivity analyses were carried out to investigate this. In the ALSPAC cohort data on age of menarche was available. Therefore whether or not a participant had experienced menarche was entered into the linear regression models to test the impact on the BMI – depression score association. As reported earlier (see Section 5.1.1), the inclusion of a marker of pubertal stage did not alter the conclusions drawn from the analyses and as such puberty is unlikely to be the cause of the inconsistent findings for females reported in this studies.

A further strength of the study was that an objective measure of obesity (BMI from measured height and weight) was available in all three cohorts and could be used in the analyses. The advantages of an objective rather than a subjective measure of

obesity is that self-reported estimates are likely to be less precise than an objective measurement and may also introduce bias into the analysis by participants over estimating height and under reporting weight due to social desirability (see Section 3.3.1). There were also other objective measures of obesity available in the cohorts (DXA fat mass, waist circumference and subscapular skinfold thickness). This allowed the analyses to be repeated with these alternative measures to investigate whether the findings were robust to the different measures of obesity. The findings were robust to the measure of obesity used and the different measures were incorporated (as factor indicators) into the obesity latent trait used in the SEM analyses.

The analytical approach taken in this project was to start fairly simply using a series of longitudinal linear regression models using BMI as the exposure and depression as the outcome, then to begin to take advantage of the repeated measures data available using GEE to estimate the average effect of BMI on future depression. Subsequently, the approach used cross-lagged SEM in order to investigate a potential bi-directional relationship between obesity and depression. MR analysis was also used as well in an attempt to strengthen the evidence for a causal relationship [61, 167]. A strength of using a variety of analytical approaches is that the different analyses make different assumptions about the data, thus if results are consistent across the various analyses it allows us to see how robust the conclusions are to the different assumptions. The results were consistent across the different methods employed and therefore viewed as robust.

MR analyses were used to try to overcome the problems of confounding and reverse causation associated with traditional epidemiological methods (see Section 3.7.1). The genetic instrumental variable for obesity that was used in the MR analysis was a good instrument, however the analysis was still limited in terms of power. A MR analysis requires a large sample size for the analysis to be carried out with adequate statistical power. In the current study the sample size of the analysis was not large

enough to provide sufficient power, as such the results of the MR analyses should be interpreted with caution.

5.4. Comparison with previous literature

The previous literature investigating the association between obesity and depression in adolescence is sparse and inconsistent (see Table 2.2 in Section 2.1.1). Moreover, the previous literature suffers from serious methodological weaknesses that this study has improved upon; for example earlier studies relied on a subjective measure of obesity, namely BMI from self-reported height and weight. The two previous studies that used self-report BMI both used a 3-level categorical obesity exposure variable ("obese", "overweight" "non-overweight") in their analysis [37, 43]. The first of these studies found evidence of a positive association between the highest category of BMI and a (binary) classification of depression in both males and females [37] whilst the second study found no evidence of an association between BMI category and (binary) depression classification in males but found evidence of a positive association in females [43].

In the current study objectively measured BMI has been used as the measure of obesity, the advantages of which have already been discussed (see Section 3.3.1). In the previous literature three studies used objectively measured BMI in their analyses [38, 42, 44]. In common with the findings from this project, Herva et al [37], Anderson et al 2007 [43], Boutelle et al [168] and Anderson et al 2011 [44] all found evidence of an association between obesity and later depression in females. It should be noted that these previous studies all used a classification of "obese" as the exposure variable, not a continuous variable, and all but one of these was investigating an association with a binary classification of depression rather than a continuous symptoms score. In contrast to the current study Mustillo et al [42] found evidence of increased risk of depression in chronically obese males (and did not

present findings in females). This difference to the current study may possibly be due to Mustillo et al using different trajectories of obesity as the exposure variable rather than a continuous score, or it may be influenced by what confounders were/were not adjusted for. Mustillo et al adjusted for age, sex, family income and other psychiatric disorder, whilst in the current study analyses were adjusted for age, sex, previous depressive symptoms, maternal depression, SES (measure of SES varied between cohorts), alcohol and smoking use (where available).

It should be noted that in two [44, 168] of the five previous studies data was only available on females not on males and for the three studies [37, 42] [43] where data was available on both males and females, one study only presented results for males whilst the other two presented results stratified by sex. None of the previous work investigated males and females together or formally tested for an interaction by sex. In this study analyses were carried out both on males and females combined, stratified by sex and also formally tested for an interaction. There was no evidence for an effect in males but some evidence for an association between obesity and later depression in females.

In the previous literature there has also often been inadequate adjustment for important confounders, for example none of the previous studies adjusted for maternal depression. The large amount of information collected by the three cohorts used in this investigation allowed for greater adjustment for potential confounding factors. Similarly, only one [38] of the previous studies adjusted for an earlier measure of the individual's depressive symptoms, therefore any observed associations may be persistence of symptoms rather than a causal association. However, this factor was adjusted for in all of the analyses in this study.

5.5. Implications and future work

Understanding the causes of depression in adolescents is challenging. A number of different risk factors are implicated in the aetiology of adolescent depression and as such assessing the impact of a single risk factor is difficult. From this investigation there was some evidence of a positive association between obesity and later depression in adolescent females, however this was not consistent across all the cohorts studied and hence further work is clearly needed. Future work needs to explore how best to conceptualize obesity as a risk factor for depression, investigating the potential of a threshold effect or considering obesity as a chronic stressor. In addition, future studies should investigate the potential interplay between obesity, puberty, social context and resilience in the relationship with adolescent depression.

CHAPTER 6. RESULTS AND DISCUSSION - OBJECTIVE 2; PHYSICAL ACTIVITY AND DEPRESSION

6.1. Linear Regression

ALSPAC

In the ALSPAC cohort linear regression analyses were carried out investigating the relationship between physical activity (exposure) and depressive symptoms (outcome) at the next follow up occasion. There was no evidence of an association between total daily minutes of physical activity (as measured by accelerometer total number of minutes spent in light, moderate or vigorous activity per day averaged across a week) (at mean age 13y 10m) and depressive symptoms (at mean age 16y 6m): a one minute increase in total daily amount of PA resulted in a 0.0002 SD increase in depressive symptoms, however the 95% CI was wide and encompassed the null (-0.0004 to 0.0009) and the p-value was large (0.466) (Table 6.1). The analysis was repeated including an interaction between physical activity and sex to test for differences in the association between PA and depression between males and females. There was no evidence that the association varied by sex (interaction coefficient: 0.001, 95% CI: -0.0003, 0.002, p-value: 0.137). When the analysis was stratified by sex, again, there was no evidence of an association between total daily minutes of physical activity and depressive symptoms in either males (wide 95% CI crossing the null: -0.0008, 0.0006, and a large p-value: 0.779) or females (wide 95% CI crossing the null: -0.0005, 0.002, and a large p-value: 0.276) (Table 6.1).

In line with the pre-specified analysis plan the linear regression model was repeated with alternative accelerometer measures of PA as the exposure: accelerometer counts per minute (CPM), daily minutes of MVPA, percentage of time spent in MVPA and whether or not the participant achieved the recommendation of at least one hour of MVPA a day.

When the analysis was repeated using CPM as the exposure variable there was no evidence of an association between CPM (at mean age 13y 10m) and depression (at mean age 16y 6m) (an increase in one count per minute was associated with a 0.00005 SD increase in depressive symptoms score, 95% CI: -0.0003, 0.0006, p-value: 0.685) (Table 6.1). The analysis was repeated including an interaction between CPM and sex to test for differences in the association between PA and depression between males and females. However, there was no evidence that the association between PA and depression varied by gender (interaction coefficient: 0.0001, 95% CI: -0.0003, 0.0006, p-value: 0.520). When the analysis was stratified by gender, there was no evidence of an association between CPM and depression in either males (95% CI: -0.0003, 0.0002, p-value: 0.816) or females (95% CI: -0.0003, 0.0005, p-value: 0.541) (Table 6.1).

In order to examine the effect of intensity (rather than amount) of PA, additional analyses focused on MVPA were carried out. The linear regression analysis was repeated using daily minutes of MVPA as the exposure variable. There was no evidence of an association between daily minutes of MVPA (at mean age 13y 10m) and depression (at mean age 16y 6m) (95% CI: -0.002, 0.003, p-value: 0.600) (Table 6.1). The analysis was repeated including an interaction between daily minutes of MVPA and sex to test for differences in the association between PA and depression between males and females. Again using daily minutes of MVPA as the measure of PA, there was no evidence that the association between PA and depression varied by sex (interaction coefficient: -0.0005, 95% CI: -0.005, 0.004, p-value: 0.832). When the model was stratified by sex, there was no evidence of an association between daily

minutes of MVPA and depression in either males (95% CI: -0.002, 0.003, p-value: 0.646) or females (95% CI: -0.003, 0.004, p-value: 0.857) (Table 6.1).

The analysis was repeated using percentage of time spent in MVPA as the exposure variable. There was no evidence of an association between percentage of time spent in MVPA (at mean age 13y 10m) and depression (at mean age 16y 6m) (95% CI: -0.015, 0.021, p-value: 0.731) (Table 6.1). The analysis was repeated including an interaction between percentage of time spent in MVPA and sex to test for differences in the association between PA and depression between males and females. However, there was no evidence that the association varied by gender (interaction coefficient: -0.008, 95% CI: -0.045, 0.029, p-value: 0.662). When stratified by sex, there was no evidence of an association between percentage of time spent in MVPA and depression in either males (95% CI: -0.016, 0.026, p-value: 0.653) or females (95% CI: -0.031, 0.029, p-value: 0.955) (Table 6.1).

The analysis was repeated using a binary variable indicating whether the participants achieved the recommended level of at least one hour of MVPA a day as the exposure variable. There was no evidence of an association between the recommended level of PA (≥1 hour of MVPA per day) (at mean age 13y 10m) and depression (at mean age 16y 6m) (95% CI: -0.123, 0.230, p-value: 0.552) (Table 6.1). The analysis was repeated including an interaction between physical activity and sex to test for differences in the association between PA and depression between males and females. Again, there was no evidence that the association between PA and depression varied by sex (interaction coefficient: -0.227, 95% CI: -0.660, 0.205, p-value: 0.303). When the analysis was stratified by sex, there was no evidence of an association between at least one hour of MVPA a day and depression in either males (95% CI: -0.085, 0.292, p-value: 0.282) or females (95% CI: -0.505, 0.279, p-value: 0.573) (Table 6.1).

In the ALSPAC cohort self-report physical activity questionnaire data were collected as well as accelerometry data. In order to compare the findings from objective and self-report PA, analyses were repeated using this self-report measure of frequency of PA. There was no evidence of an association between self-reported frequency of physical activity in the past year (at mean age 13y 10m) and later depressive symptoms (mean age 16y 6m). A greater frequency of self-reported PA was associated with lower levels of depressive symptoms (p-value from test for trend in regression coefficients: 0.058) (Table 6.1). The analysis was repeated including an interaction between self-reported PA and sex to test for differences in the association between self-reported PA and depression between males and females. There was no evidence that the association between self-reported PA and depression varied by gender (interaction p-value: 0.937). When the analysis was stratified by sex, there was no evidence of an association between self-reported frequency of PA and depression in either males (p-value from test for trend in regression coefficients 0.222) or females (p-value from test for trend in regression coefficients 0.404) (Table 6.1).

The linear regression analysis was carried out again between self-report frequency of PA (exposure) at (mean) age 16 years 6 months and later depressive symptoms (outcome) (mean age 17 years 10 months). There was no evidence of an association between self-reported frequency of physical activity and later depression (p-value from test for trend in regression coefficients 0.132). The analysis was repeated including an interaction between self-reported PA and sex to test for differences in the association between self-report frequency of PA and depressive symptoms between males and females. There was no evidence that the association between self-reported PA and depression varied by sex (interaction p-value: 0.194). When stratified by sex, there was no evidence of an association between self-reported frequency of PA and depression in either males (p-value from test for trend in regression coefficients 0.386) or females (p-value from test for trend in regression coefficients 0.107) (Table 6.1).

Table 6.1 - Results from the linear regression analyses investigating the association between PA (exposure) on depression (Z score) (outcome) at next follow up in the ALSPAC cohort

Timepoint	Exposure	Outcome	Model 1 - main model				Model 2 - Including interaction (coefficient for interaction effect)			
			n	coeff.	95% CI	p-value	n	coeff[#].	95% CI	p-value
TF2 to CCS 13y10m to 16y6m	Total daily minutes of PA	Depression	2025	0.0002	(-0.0004, 0.0009)	0.466	2025	0.001	(-0.0003, 0.002)	0.137
TF2 to CCS 13y10m to 16y6m	Accelerometer counts per minute	Depression	2025	0.00005	(-0.0002, 0.0003)	0.685	2025	0.0001	(-0.0003, 0.0006)	0.520
TF2 to CCS 13y10m to 16y6m	Daily minutes of MVPA	Depression	2025	0.0006	(-0.002, 0.003)	0.600	2025	-0.0005	(-0.005, 0.004)	0.832
TF2 to CCS 13y10m to 16y6m	Percentage of time spent in MVPA	Depression	2025	0.003	(-0.015, 0.021)	0.731	2025	-0.008	(-0.045, 0.029)	0.662
TF2 to CCS 13y10m to 16y6m	At least 1 hour of MVPA a day	Depression	2025	0.053	(-0.123, 0.230)	0.552	2025	-0.227	(-0.660, 0.205)	0.303
TF2 to CCS 13y10m to 16y6m	Self reported frequency of PA in past year	Depression	2559				2559			
	Never		47	ref			47	ref		
	Less than one a month		29	0.038	(-0.441, 0.516)		29	-0.362	(-1.345, 0.622)	
	1-3 times a month		155	-0.261	(-0.602, 0.080)		155	-0.079	(-0.788, 0.630)	
	1-4 times a week		1360	-0.203	(-0.523, 0.118)		1360	-0.131	(-0.821, 0.526)	
	5 or more times a week		968	-0.277	(-0.597, 0.043)	0.058	968	-0.148	(-0.821, 0.525)	0.937
CCS to TF4 16y6m to 17y10m	Self reported frequency of PA in past year	Depression	2357				2357			
	Never		81	ref			81	ref		
	Less than one a month		118	0.071	(-0.188, 0.330)		118	0.213	(-0.324, 0.751)	
	1-3 times a month		331	-0.119	(-0.333, 0.095)		331	-0.056	(-0.502, 0.390)	
	1-4 times a week		1220	-0.143	(-0.344, 0.058)		1220	0.031	(-0.380, 0.443)	
	5 or more times a week		607	-0.131	(-0.339, 0.077)	0.132	607	0.191	(-0.233, 0.615)	0.194

Model 1 is adjusted for age, sex, previous depression, maternal depression, maternal education, maternal profession and accelerometer weartime (where appropriate)

*Model 2 is Model 1 plus the inclusion of a PA*Sex interaction term*

[#] *Results presented are for the PA* Sex interaction term*

Table continued

Timepoint	Exposure	Outcome	Model 3 - stratified (females)				Model 4 - stratified (males)			
			n	coeff.	95% CI	p-value	n	coeff.	95% CI	p-value
TF2 to CCS 13y10m to 16y6m	Total daily minutes of PA	Depression	1157	0.001	(-0.0005, 0.002)	0.276	868	-0.0001	(-0.0008, 0.0006)	0.779
TF2 to CCS 13y10m to 16y6m	Accelerometer counts per minute	Depression	1157	0.0001	(-0.0003, 0.0005)	0.541	868	-3E-05	(-0.0003, 0.0002)	0.816
TF2 to CCS 13y10m to 16y6m	Daily minutes of MVPA	Depression	1157	0.0004	(-0.003, 0.004)	0.857	868	0.0006	(-0.002, 0.003)	0.646
TF2 to CCS 13y10m to 16y6m	Percentage of time spent in MVPA	Depression	1157	-0.0009	(-0.031, 0.029)	0.955	868	0.005	(-0.016, 0.026)	0.653
TF2 to CCS 13y10m to 16y6m	At least 1 hour of MVPA a day	Depression	1157	-0.113	(-0.505, 0.279)	0.573	868	0.103	(-0.085, 0.292)	0.282
TF2 to CCS 13y10m to 16y6m	Self reported frequency of PA in past year	Depression	1462				1097			
	Never		18	ref			29	ref		
	Less than one a month		17	-0.146	(-0.863, 0.570)		12	0.217	(-0.499, 0.883)	
	1-3 times a month		87	-0.318	(-0.881, 0.245)		68	-0.233	(-0.650, 0.185)	
	1-4 times a week		867	-0.272	(-0.809, 0.265)		493	-0.171	(-0.563, 0.220)	
	5 or more times a week		473	-0.350	(-0.888, 0.188)	0.404	493	-0.240	(-0.629, 0.149)	0.222
CCS to TF4 16y6m to 17y10m	Self reported frequency of PA in past year	Depression	1376				981			
	Never		54	ref			27	ref		
	Less than one a month		83	0.135	(-0.193, 0.462)		35	-0.066	(-0.493, 0.360)	
	1-3 times a month		247	-0.122	(-0.392, 0.148)		84	-0.047	(-0.405, 0.311)	
	1-4 times a week		743	-0.127	(-0.385, 0.131)		477	-0.141	(-0.465, 0.182)	
	5 or more times a week		249	-0.033	(-0.307, 0.240)	0.107	358	-0.203	(-0.531, 0.125)	0.386

Model 3 is Model 1 stratified by gender - analysing females only and without including sex as a confounder

Model 4 is Model 1 stratified by gender - analysing males only and without including sex as a confounder

TRAILS

In the TRAILS cohort, physical activity data was collected via self-report of total frequency of PA, as measured by number of days per week physical activity was carried out ("never", "once a week", "2-3 times a week", "4-5 times a week", "6-7 times a week"). There was no evidence of an association between self-reported frequency of PA (exposure) (at mean age 10y 7m) and depressive symptom score (outcome) (at mean age 13y 1m) (p-value from test for trend in regression coefficients: 0.676) (Table 6.2). The analysis was repeated including an interaction between self-report PA and sex to test for differences in the association between PA and depression between males and females, there was no evidence that the association between self-reported frequency of PA and later depression varied by gender (interaction coefficient p-value: 0.918) (Table 6.2). When the analysis was stratified by sex, there was no evidence of an association between self-reported frequency of PA and depression in either males (p-value from test for trend in regression coefficients: 0.897) or females (p-value from test for trend in regression coefficients: 0.637) (Table 6.2).

There was no evidence of an association between self-reported frequency of PA (exposure) (at mean age 13y 1m) and depression (outcome) (at mean age 15y 10m) (p-value from test for trend in regression coefficients: 0.688) (Table 6.2). The analysis was repeated including an interaction between self-reported frequency of PA and sex to test for differences in the association between self-reported frequency of PA and depression between males and females. There was no evidence that the association varied by gender (interaction coefficient p-value: 0.614) (Table 6.2). When stratified by sex, there was no evidence of an association between self-reported frequency of PA and depression in either males (p-value from test for trend in regression coefficients: 0.155) or females (p-value from test for trend in regression coefficients: 0.992) (Table 6.2).

There was no evidence of an association between self-reported frequency of PA (exposure) (at mean age 15y 10m) and depression (outcome) (at mean age 18y 7m) (p-value from test for trend in regression coefficients: 0.244) (Table 6.2). The analysis was repeated including an interaction between PA and sex to test for differences in the association between PA and depression between males and females. There was no evidence that the association between self-reported frequency of PA and later depression varied by sex (interaction coefficient p-value: 0.251) (Table 6.2). When the analysis was stratified by sex, there was no evidence of an association between self-reported frequency of PA and depression in either males (p-value from test for trend in regression coefficients: 0.448) or females (p-value from test for trend in regression coefficients: 0.212) (Table 6.2).

Table 6.2 - Results from the linear regression analyses investigating the association between PA (exposure) on depression (Z score) (outcome) at next follow up in the TRAILS cohort

Timepoint	Exposure	Outcome	Model 1 - main model				Model 2 - Including interaction (coefficient for interaction effect)			
			n	coeff.	95% CI	p-value	n	coeff[#].	95% CI	p-value
T1 to T2	No. of days of									
10y7m to 13y1m	PA a week:	Depression	1873				1873			
	Never		221	ref			221	ref		
	Once a week		459	0.039	(-0.115, 0.193)		459	-0.066	(-0.368, 0.236)	
	2 or 3 days a week		663	0.014	(-0.130, 0.159)		663	-0.092	(-0.374, 0.190)	
	4 or 5 days a week		289	-0.009	(-0.167, 0.148)		289	-0.121	(-0.432, 0.191)	
	6 or 7 days a week		241	-0.056	(-0.213, 0.101)	0.676	241	-0.148	(-0.485, 0.188)	0.918
T2 to T3	No. of days of									
13y1m to 15y10m	PA a week:	Depression	1504				1504			
	Never		118	ref			118	ref		
	Once a week		194	-0.078	(-0.279, 0.124)		194	-0.005	(-0.388, 0.378)	
	2 or 3 days a week		593	-0.051	(-0.231, 0.129)		593	0.033	(-0.305, 0.370)	
	4 or 5 days a week		404	-0.071	(-0.254, 0.113)		404	0.048	(-0.299, 0.395)	
	6 or 7 days a week		195	0.021	(-0.184, 0.226)	0.688	195	-0.175	(-0.569, 0.219)	0.614
T3 to T4	No. of days of									
15y10m to 18y7m	PA a week:	Depression	1352				1352			
	Never		126	ref			126	ref		
	Once a week		176	-0.050	(-0.248, 0.149)		176	0.150	(-0.246, 0.547)	
	2 or 3 days a week		500	-0.017	(-0.194, 0.159)		500	0.145	(-0.208, 0.499)	
	4 or 5 days a week		306	0.041	(-0.143, 0.226)		306	0.360	(-0.006, 0.726)	
	6 or 7 days a week		244	-0.101	(-0.280, 0.079)	0.244	244	0.193	(-0.168, 0.553)	0.251

Model 1 is adjusted for age, sex, previous depression, maternal depression, SEP, alcohol and smoking

*Model 2 is Model 1 plus the inclusion of a PA*Sex interaction term*

[#] *Results presented are for the PA* Sex interaction term*

Table continued

Timepoint	Exposure	Outcome	Model 3 - stratified (females)				Model 4 - stratified (males)			
			n	coeff.	95% CI	p-value	n	coeff.	95% CI	p-value
T1 to T2	No. of days of									
10y7m to 13y1m	PA a week:	Depression	964				909			
	Never		119	ref			102	ref		
	Once a week		301	0.006	(-0.219, 0.232)		158	0.046	(-0.153, 0.246)	
	2 or 3 days a week		349	-0.028	(-0.249, 0.193)		314	0.031	(-0.152, 0.213)	
	4 or 5 days a week		131	-0.080	(-0.328, 0.168)		158	0.012	(-0.182, 0.205)	
	6 or 7 days a week		64	-0.164	(-0.445, 0.117)	0.637	177	-0.027	(-0.213, 0.159)	0.897
T2 to T3	No. of days of									
13y1m to 15y10m	PA a week:	Depression	793				711			
	Never		72	ref			46	ref		
	Once a week		106	-0.053	(-0.348, 0.241)		88	-0.085	(-0.330, 0.160)	
	2 or 3 days a week		331	-0.011	(-0.274, 0.252)		262	-0.077	(-0.290, 0.136)	
	4 or 5 days a week		191	-0.017	(-0.291, 0.256)		213	-0.123	(-0.340, 0.094)	
	6 or 7 days a week		93	-0.046	(-0.354, 0.262)	0.992	102	0.083	(-0.161, 0.327)	0.155
T3 to T4	No. of days of									
15y10m to 18y7m	PA a week:	Depression	743				609			
	Never		73	ref			53	ref		
	Once a week		123	0.014	(-0.257, 0.285)		53	-0.119	(-0.404, 0.165)	
	2 or 3 days a week		271	0.029	(-0.218, 0.277)		229	-0.091	(-0.342, 0.159)	
	4 or 5 days a week		158	0.191	(-0.073, 0.454)		148	-0.138	(-0.392, 0.116)	
	6 or 7 days a week		118	-0.028	(-0.287, 0.231)	0.212	126	-0.201	(-0.449, 0.047)	0.448

Model 3 is Model 1 stratified by gender - analysing females only and without including sex as a confounder

Model 4 is Model 1 stratified by gender - analysing males only and without including sex as a confounder

NDIT

In the NDIT cohort the linear regression analyses investigating the relationship between physical activity (as measured by self-report number of bouts of MVPA per week) and later depression showed little consistent evidence of an association between MVPA and depressive symptoms but some weak evidence of a very small effect in later adolescence (bouts of MVPA at mean age 16y 2m and depression at 16y 6m (p-value 0.041)). This analysis suggested that an increase in one bout of MVPA (a bout was defined as at least 5 minutes in an activity previously defined as MVPA) per week was associated with a very small (0.007 standard deviation (95% CI 0.0003, 0.131)) increase in depressive symptoms. There was also evidence of a positive association between PA at age 16y 9m and depression at 17y 0m (coefficient; 0.008, 95% CI; 0.0004, 0.016, p-value; 0.038).

The regression models at the different time points were each repeated including an interaction between bouts of MVPA and sex to test for differences in the association between PA and depression between males and females. At one time point (PA at 17 years and depression at 17 years and 1 month) there was evidence of a difference in association in females compared to males (interaction coefficient -0.019, 95% CI -0.035, -0.003, p-value 0.019), however there was no evidence of a difference between males and females at any of the other time points (18 other time points, p-values range from 0.074 to 0.949). When the regression analyses were carried out stratified by sex there was evidence of a positive association between bouts of MVPA at 16y 0m and depressive symptoms at 16y 2m in males (an increase in one bout of MVPA per week was associated with a 0.009 SD increase in depressive symptom score, 95% CI; 0.002, 0.016, p-value; 0.017). In females there was evidence of a positive association between self-reported bouts of MVPA at 16y 9m and depressive symptoms at 17y 0m, and evidence of an inverse association between bouts of MVPA at 17y 0m and depressive symptoms at 17y 1m (an increase of one bout of MVPA per week was associated with a -0.015 SD decrease in depressive symptoms score, 95% CI; -0.028, -0.001, p-value; 0.036).

Table 6.3 - Results from the linear regression analyses investigating the association between PA (exposure) on depression (Z score) (outcome) at next follow up in the NDIT cohort

Timepoint	Exposure	Outcome	Model 1 - main model				Model 2 - interaction model[#]				Model 3 - stratified (females)				Model 4 - stratified (males)			
			n	coeff.	95% CI	p-value	n	coeff.	95% CI	p-value	n	coeff.	95% CI	p-value	n	coeff.	95% CI	p-value
T1 to T2 12y9m to 13y0m	No. Bouts of MVPA	Depression	516	0.0002	(-0.004, 0.004)	0.945	516	-0.007	(-0.016, 0.001)	0.096	274	-0.005	(-0.012, 0.001)	0.113	242	0.004	(-0.002, 0.009)	0.196
T2 to T3 13y0m to 13y2m	No. Bouts of MVPA	Depression	507	0.002	(-0.002, 0.006	0.274	507	-0.007	(-0.017, 0.003)	0.183	265	-0.002	(-0.010, 0.006)	0.631	242	0.003	(-0.001, 0.008)	0.151
T3 to T4 13y2m to 13y2m	No. Bouts of MVPA	Depression	221	-0.001	(-0.008, 0.005)	0.662	221	-0.013	(-0.027, 0.001)	0.074	121	-0.009	(-0.021, 0.003)	0.153	100	0.003	(-0.005, 0.012)	0.453
T4 to T5 13y2m to 13y8m	No. Bouts of MVPA	Depression	214	0.003	(-0.004, 0.010)	0.387	214	-0.007	(-0.022, 0.008)	0.342	119	0.001	(-0.013, 0.014)	0.940	95	0.008	(-0.002, 0.017)	0.123
T5 to T6 13y8m to 13y10m	No. Bouts of MVPA	Depression	492	-0.002	(-0.008, 0.003)	0.392	492	0.001	(-0.009, 0.012)	0.797	256	-0.0004	(-0.008, 0.007)	0.910	236	-0.004	(-0.011, 0.004)	0.336
T6 to T7 13y10m to 14y1m	No. Bouts of MVPA	Depression	415	0.005	(-0.001, 0.011)	0.131	415	-0.0004	(-0.012, 0.011)	0.949	218	0.004	(-0.004, 0.012)	0.295	197	0.006	(-0.003, 0.016)	0.171
T7 to T8 14y1m to 14y2m	No. Bouts of MVPA	Depression	370	-0.004	(-0.010, 0.002)	0.211	370	0.0006	(-0.012, 0.013)	0.919	185	-0.006	(-0.016, 0.005)	0.280	185	-0.002	(-0.009, 0.004)	0.513
T8 to T9 14y2m to 14y7m	No. Bouts of MVPA	Depression	418	0.001	(-0.004, 0.006)	0.594	418	-0.002	(-0.014, 0.009)	0.680	214	0.0003	(-0.010, 0.011)	0.951	204	0.002	(-0.004, 0.008)	0.500
T9 to T10 14y7m to 14y10m	No. Bouts of MVPA	Depression	473	0.002	(-0.003, 0.007)	0.432	473	0.001	(-0.013, 0.014)	0.918	245	0.003	(-0.010, 0.016)	0.639	228	0.001	(-0.004, 0.006)	0.700

Table continued on next page …

Timepoint	Exposure	Outcome	Model 1 - main model				Model 2 - interaction model[#]				Model 3 - stratified (females)				Model 4 - stratified (males)			
			n	coeff.	95% CI	p-value	n	coeff.	95% CI	p-value	n	coeff.	95% CI	p-value	n	coeff.	95% CI	p-value
T10 to T11 14y10m to 15y0m	No. Bouts of MVPA	Depression	458	0.001	(-0.004, 0.006)	0.711	458	-0.003	(-0.013, 0.008)	0.606	239	0.0003	(-0.010, 0.010)	0.959	219	0.003	(-0.003, 0.009)	0.376
T11 to T12 15y0m to 15y2m	No. Bouts of MVPA	Depression	446	0.003	(-0.003, 0.008)	0.338	446	0.001	(-0.010, 0.011)	0.925	232	0.004	(-0.005, 0.014)	0.390	214	0.001	(-0.005, 0.007)	0.727
T12 to T13 15y2m to 15y7m	No. Bouts of MVPA	Depression	444	-0.001	(-0.006, 0.004)	0.793	444	-0.005	(-0.014, 0.005)	0.338	231	-0.004	(-0.012, 0.003)	0.232	213	0.002	(-0.005, 0.009)	0.592
T13 to T14 15y7m to 15y10m	No. Bouts of MVPA	Depression	448	0.004	(-0.003, 0.010)	0.239	448	0.002	(-0.011, 0.016)	0.743	236	0.006	(-0.006, 0.018)	0.317	212	0.002	(-0.006, 0.009)	0.682
T14 to T15 15y10m to 16y0m	No. Bouts of MVPA	Depression	450	-0.001	(-0.007, 0.005)	0.736	450	-0.003	(-0.016, 0.010)	0.654	231	-0.004	(-0.015, 0.007)	0.462	219	-0.001	(-0.008, 0.007)	0.889
T15 to T16 16y0m to 16y2m	No. Bouts of MVPA	Depression	446	0.004	(-0.002, 0.010)	0.190	446	-0.012	(-0.026, 0.001)	0.080	231	-0.006	(-0.018, 0.006)	0.366	215	0.009	(0.002, 0.016)	0.017
T16 to T17 16y2m to 16y6m	No. Bouts of MVPA	Depression	438	0.007	(0.0003, 0.131)	0.041	438	0.002	(-0.011, 0.015)	0.733	230	0.009	(-0.002, 0.019)	0.106	208	0.006	(-0.003, 0.015)	0.192
T17 to T18 16y6m to 16y9m	No. Bouts of MVPA	Depression	370	-0.002	(-0.011, 0.008)	0.730	370	-0.007	(-0.027, 0.014)	0.528	196	-0.007	(-0.025, 0.010)	0.407	174	-0.001	(-0.011, 0.009)	0.848
T18 to T19 16y9m to 17y0m	No. Bouts of MVPA	Depression	364	0.008	(0.0004, 0.016)	0.038	364	0.014	(-0.005, 0.033)	0.139	192	0.017	(0.001, 0.034)	0.039	172	0.001	(-0.008, 0.010)	0.793
T19 to T20 17y0m to 17y1m	No. Bouts of MVPA	Depression	370	-0.003	(-0.009, 0.004)	0.452	370	-0.019	(-0.035, -0.003)	0.019	196	-0.015	(-0.028, -0.001)	0.036	174	0.001	(-0.010, 0.011)	0.911

Model 1 is adjusted for age, sex, previous depression, maternal education, maternal profession and alcohol

*Model 2 is Model 1 plus the inclusion of a PA*Sex interaction term*

Model 3 is Model 1 stratified by gender - analysing females only and without including sex as a confounder

Model 4 is Model 1 stratified by gender - analysing males only and without including sex as a confounder

[#] *Results presented are for the PA* Sex interaction term*

6.1.1. Generalized Estimating Equations

Following on from the linear regression analyses, GEEs were used to model the repeated exposure-outcome association – i.e. the average effect of level of physical activity on future depressive symptoms.

ALSPAC

In the ALSPAC cohort the results of the GEE analysis suggested that there was strong evidence of an association between lagged self-report frequency of PA and later depression (i.e. the effect on depression of PA at the previous time point) (p-value for test of trend of regression coefficients <0.001). The results suggest that very low frequency of physical activity may be associated with increased depressive symptoms, whereas a high frequency of physical activity was associated with a reduced later depressive symptoms (as the regression coefficients are positive – representing increased depression for the categories representing low frequency of PA, but are then negative – representing reduced depression for the categories representing high frequencies of PA) (Table 6.4). The analysis was repeated including an interaction between PA and sex to test for differences in the association between PA and depression between males and females. There was no evidence that the association varied by sex (interaction coefficient p-value: 0.973). When the GEE analysis was carried out stratified by gender there was evidence of the same association observed in the main model between lagged frequency of PA and later levels of depressive symptoms (i.e. low frequency of PA associated with increased depression and high frequency of PA associated with reduced levels of depression) in both males and females (p-value for test of trend of regression coefficients in males and females respectively: <0.001, 0.035) (Table 6.4). The GEE analysis could not be carried out using the objective accelerometer data due to the timings at which these measurements were made.

TRAILS

When the GEE analysis was carried out in the TRAILS cohort there was no evidence of an association between lagged PA and future levels of depression (p-value for test of trend of regression coefficients 0.737). The analysis was repeated including an interaction between PA and sex to test for differences in the association between PA and depression by sex, there was no evidence that the association varied by gender (interaction coefficient p-value: 0.217). When the GEE analysis was carried out separately in males and females there was no evidence of an association between lagged PA and future levels of depression in either males or females (p-value for test of trend of regression coefficients in males and females respectively: 0.374, 0.111) (Table 6.4).

NDIT

When the GEE analysis was carried out in the NDIT cohort, there was no evidence of an association between self-reported lagged bouts of MVPA and future depressive symptoms (95% CI: -0.001, 0.001, p-value: 0.991). The analysis was repeated including an interaction between lagged bouts of MVPA and sex to test for differences in the association between PA and depression between males and females, there was no evidence that the association varied by sex (interaction coefficient p-value: 0.097). When the GEE analysis was carried out separately in males and females there was no evidence of an association between self-reported lagged bouts of MVPA and future of depressive symptoms in males (95% CI: -0.001, 0.003, p-value: 0.301), or females (95% CI: -0.002, 0.001, p-value: 0.772). However, the result in females when the model was stratified by sex should be interpreted with caution as the model failed to converge (even after increasing the maximum number of iterations from the default 100 to 50000) (Table 6.4).

Given these difficulties with model convergence, the NDIT GEE analyses were repeated using a reduced number of follow up occasions (data was used from measurements taken at every other follow up occasion - i.e. every 6 months rather

than every 3 months) to try and produce a model that achieved convergence when the analysis was stratified by sex. When the analysis was repeated using the reduced number of follow up occasions there was no evidence of an association between self-reported lagged bouts of MVPA and later depressive symptoms in the main model (95% CI -0.008, 0.0004, p-value 0.074). The analysis was repeated including an interaction between PA and sex to test for differences in the association between PA and depression between males and females. There was evidence that the association between self-reported lagged bouts of MVPA and depressive symptom score differed between males and females (interaction coefficient: -0.005, 95% CI: -0.009, -0.0002, p-value: 0.040). When the analysis was stratified by sex there was no evidence of an association between lagged bouts of MVPA and depression in males (95% CI: -0.002, 0.004, p-value: 0.704) or females (an increase in one bout of lagged MVPA was associated with a -0.004 SD decrease in depressive symptom score, 95% CI: -0.008, 0.0004, p-value: 0.074) (Table 6.4).

Table 6.4 - Results from the GEE analyses investigating the association between lagged PA (exposure) on depression (Z score) (outcome) at next follow up in all three cohorts

			Model 1 - main model				Model 2 - interaction model[#]				Model 3 - stratified (females)				Model 4 - stratified (males)			
Cohort	**Exposure**	**Outcome**	**n**	**coeff.**	**95% CI**	**p-value**	**n**	**coeff.**	**95% CI**	**p-value**	**n**	**coeff.**	**95% CI**	**p-value**	**n**	**coeff.**	**95% CI**	**p-value**
ALSPAC	Lagged self report frequency of PA:	Depression	6915				6915				3832				3083			
	Never			ref				ref				ref				ref		
	Less than once a month			0.225	(0.029, 0.422)			-0.018	(-0.413, 0.376)			0.215	(-0.048, 0.478)			0.245	(-0.049, 0.540)	
	1-3 times a month			0.029	(-0.123, 0.182)			-0.076	(-0.381, 0.228)			0.008	(-0.201, 0.216)			0.072	(-0.149, 0.293)	
	1-4 times a week			-0.038	(-0.179, 0.102)			-0.046	(-0.325, 0.233)			-0.034	(-0.230, 0.161)			-0.026	(-0.226, 0.173)	
	5 or more times a week			-0.116	(-0.259, 0.027)	<0.001		-0.028	(-0.311, 0.256)	0.973		-0.093	(-0.298, 0.112)	0.035		-0.120	(-0.319, 0.078)	<0.001
TRAILS	Lagged self report frequency of PA:	Depression	1887				1887				976				911			
	Never			ref				ref				ref				ref		
	Once a week			0.008	(-0.024, 0.041)			0.009	(-0.047, 0.065)			0.023	(-0.024, 0.070)			-0.009	(-0.050, 0.032)	
	2 or 3 days a week			0.012	(-0.020, 0.044)			0.0002	(-0.054, 0.054)			0.018	(-0.028, 0.064)			0.004	(-0.038, 0.046)	
	4 or 5 days a week			0.020	(-0.012, 0.052)			0.035	(-0.020, 0.090)			0.050	(0.002, 0.098)			-0.016	(-0.057, 0.026)	
	6 or 7 days a week			0.001	(-0.025, 0.044)	0.737		-0.010	(-0.069, 0.049)	0.217		0.009	(-0.046, 0.064)	0.111		0.005	(-0.038, 0.047)	0.374
NDIT[1]	Lagged no. of bouts of MVPA	Depression	541	-8.45E-06	(-0.001, 0.001)	0.991	541	-0.002	(-0.005, 0.0004)	0.097	286	-0.0002	(-0.002, 0.001)	0.772	255	0.001	(-0.001, 0.003)	0.301
NDIT[2]	Lagged no. of bouts of MVPA	Depression	538	-0.001	(-0.004, 0.002)	0.439	538	-0.005	(-0.009, -0.0002)	0.040	283	-0.004	(-0.008, 0.0004)	0.074	255	0.001	(-0.002, 0.004)	0.704

Model 1 is adjusted for age, sex and previous depression in all cohorts (Plus maternal depression, maternal education and maternal profession in ALSPAC)

(Plus maternal depression, SEP, alcohol and smoking in TRAILS) (Plus maternal education, maternal profession and alcohol in NDIT).

*Model 2 is Model 1 plus the inclusion of a PA*Sex interaction term*

Model 3 is Model 1 stratified by gender - analysing females only and without including sex as a confounder

Model 4 is Model 1 stratified by gender - analysing males only and without including sex as a confounder

[#] *Results presented are for the PA* Sex interaction term*

6.1.2. Partial least squares regression

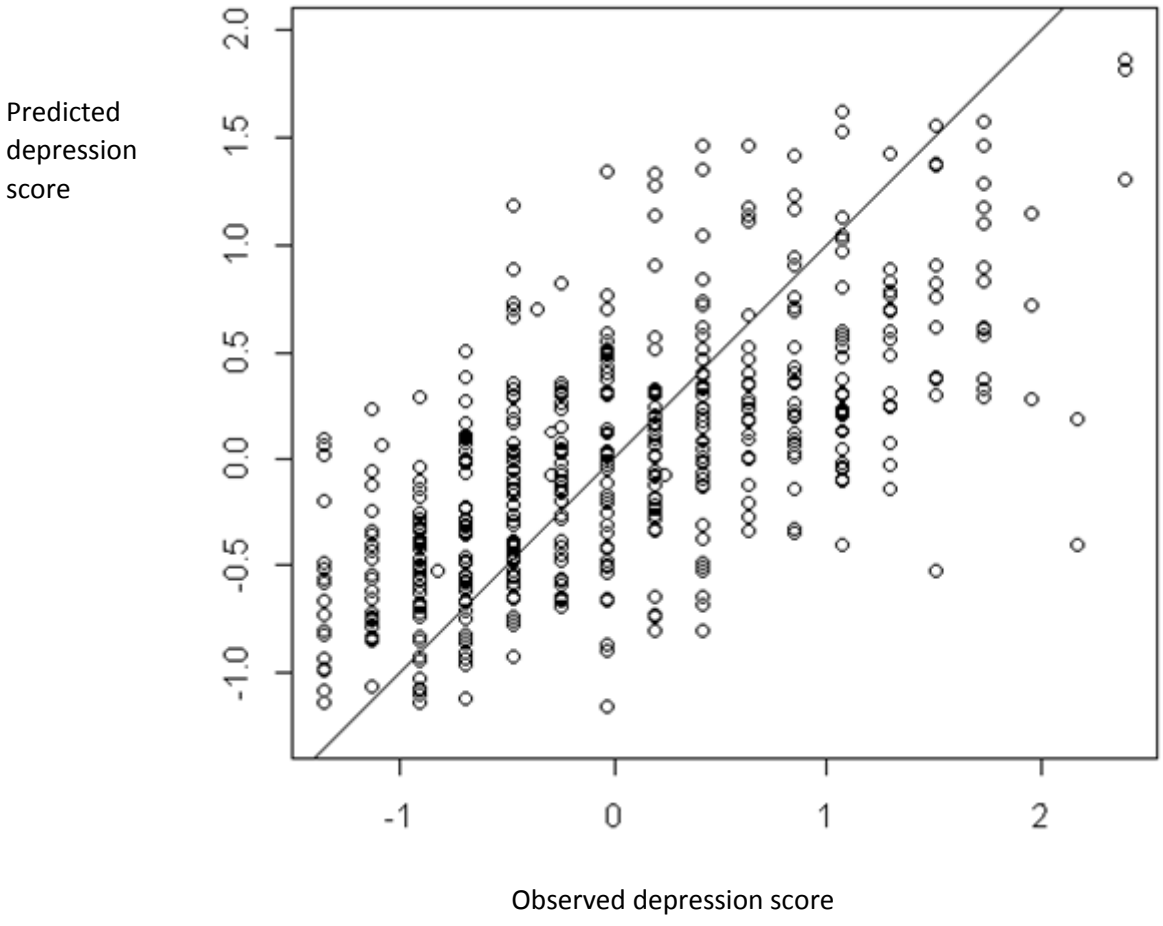
Partial least squares regression was used to try and identify which aspects of physical activity may be important in adolescent depression. The only cohort that collected sufficient physical activity questionnaire data to make PLS-R possible was the NDIT cohort.

As outlined in Section 3.7.2 the first aspect of PLS-R is to choose the appropriate number of components to retain in the model. The number of components to retain can be judged based on what number maximises the percentage of variance in the exposure and outcome explained and minimizes the size of the root mean squared error of prediction (RMSEP). Based on the percentage of variance explained, retaining between 5 and 7 components was viewed as sufficient, as retaining any further components only increased the amount of variance explained by a very small amount (Table 6.5). It should be noted that although retaining six components explained a large amount of variance in the exposure (92.35%), there was still a large amount of unexplained variance in the outcome (retaining 6 components explained 51.29% of the variance in the outcome). Retaining 6 components produced the smallest RMSEP value (Table 6.5). Based on both these findings it was decided to retain the first 6 components in the model. To validate the decision to retain 6 components the measured values of depression Z scores were plotted against the predicted scores using the 6 component model (Figure 6.1). The points of the plot followed the target line quite well and there was no evidence of the points fanning out, curvature or any other anomalies, this supported the decision to retain 6 components.

Table 6.5 Table outlining information required when deciding how many components to retain

No. of Components	RMSEP	Cumulative % of variance in exposure explained	Cumulative % of variance in outcome explained
1	0.8167	82.90	1.53
2	0.7845	88.71	10.94
3	0.7214	89.78	31.26
4	0.6611	90.82	44.80
5	0.6287	91.68	49.73
6	0.6193	92.35	51.29
7	0.6215	92.80	51.93
8	0.6214	93.25	52.43
9	0.6281	93.75	52.78
10	0.6336	94.38	53.07

Figure 6.1 Plot of the observed versus the predicted depression Z score based on retaining six components in the PLS-R model



The orthogonal rotation loadings (see Section 3.7.2) of the questionnaire items onto the six components did not reveal a clear, discernible pattern where each component could been seen to clearly represent a certain aspect of physical activity (Table 6.6). The first component, to which the "Bouts of MVPA" and "Total PA" items loaded very strongly may represent overall amount of PA. This first component explained the largest amount of variance in the exposures of all the six components, however it only explained a very small amount (1.53%) of variance in the outcome (depression). The second component may be viewed as representing a low intensity PA component as, other than total PA, the two items that loaded strongest onto this component were "walking" and "indoor chores", which represented the least intense forms of activity asked about on the questionnaire. The other components could not be interpreted as reflecting any particular aspect(s) of physical activity or contrasts between different types of PA.

Table 6.6 Loadings of different types of PA onto PLS regression components

	Loadings					
	Component 1	**Component 2**	**Component 3**	**Component 4**	**Component 5**	**Component 6**
Bouts of MVPA	0.796	0.152	-0.112	-0.145		
Total PA	0.918	-0.890		0.174		
Running/Jogging			0.416	-0.458	0.317	-0.279
Walking		-0.559	1.067	-0.434	-0.246	0.195
Mixed Walking						-0.531
Indoor Chores		-0.276	-0.451	0.609		-0.507
Outdoor Chores				0.135		
Boxing/Wrestling			-0.175			0.223
Outdoor Play			-0.253			0.336
Dancing			-0.257			0.301
Physical Exercise				0.123	-0.325	0.211
Gymnastics			-0.109			
Rollerblading				0.113	-0.151	0.134
Ice Skating					-0.122	
Hockey		0.104	0.159	0.204	-0.450	-0.236
Cycling			-0.168	0.174		-0.220

When the six components were fitted as exposure variables with depressive symptom score (Z score) as the outcome, there was strong evidence that all six components were associated with later depression (all p-values <0.001) (Table *6.7*). The regression coefficient for component 1 (which could be viewed as representing overall activity) was however very small; a one unit increase in component 1 score was associated with a 0.007 standard deviation increase in depression score (95%CI: 0.004, 0.010). The regression results suggested that components 3, 4 and 5 explained the largest amount of variation in the depression outcome.

Table 6.7 - Results of the PLS-R investigation into the relationship between the identified components of self-reported PA and depression in the NDIT cohort

n=516	Regression coefficient	95% CI	p-value
Component 1	0.007	(0.004, 0.010)	<0.001
Component 2	0.059	(0.047, 0.070)	<0.001
Component 3	0.252	(0.218, 0.287)	<0.001
Component 4	0.170	(0.142, 0.198)	<0.001
Component 5	0.111	(0.079, 0.142)	<0.001
Component 6	0.068	(0.034, 0.101)	<0.001

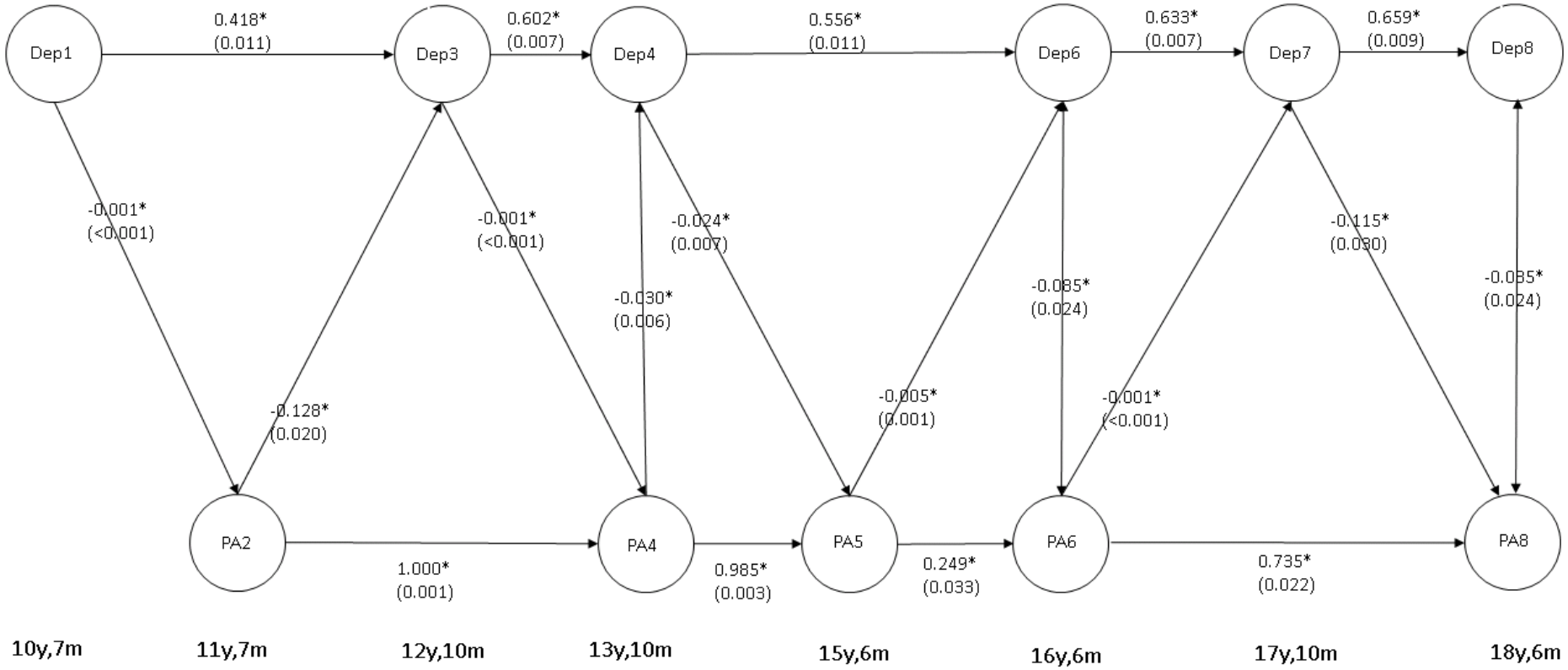
6.1.3. Cross-lagged Structural Equation Modelling

A cross-lagged SEM was fitted to the data from each of the cohorts in order to test the potential bi-directionality of the relationship between physical activity and depression in adolescents.

ALSPAC

When the cross-lagged SEM approach was used in the ALSPAC cohort there was evidence of an autoregressive relationship for both PA and depression (Figure 6.2). There was also evidence of a bi-directional cross-lagged relationship; there was evidence that adolescents who were more active (had a higher PA latent trait score) were less depressed (had a decrease in depression latent trait score) at the next follow up, and vice-versa those who were more depressed (had a higher depression latent trait score) were less active (had a decrease in PA latent trait score) at the next follow up. The magnitudes of these cross-lagged effects changed through time, the strength of the depression to PA inverse association increased over time whereas the strength of the PA to depression inverse association decreased over time. When the SEM analysis was carried out separately in males and females the same autoregressive and cross-lagged relationships were observed in both males (Figure 6.3) and females (Figure 6.4).

Figure 6.2 - Cross lagged SEM in the ALSPAC cohort describing the relationship between physical activity (PA) and depressive (Dep) latent traits

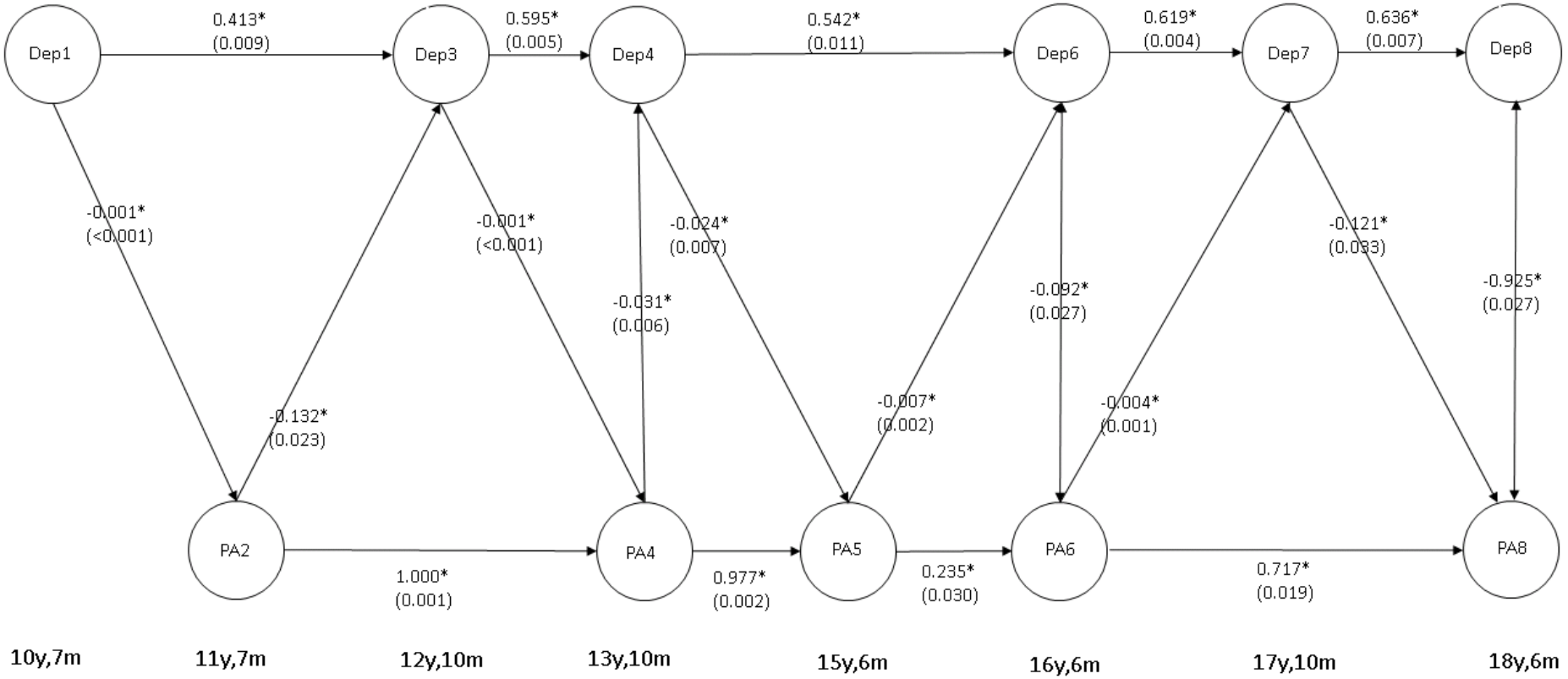


Model adjusted for age, sex, socio-economic position (maternal education and maternal profession) and maternal depression.

Coefficients represent standard deviation change, figures in brackets represent standard error

**represents p-value <0.05*

Figure 6.3 - Cross lagged SEM in the ALSPAC cohort describing the relationship between physical activity (PA) and depressive (Dep) latent traits in males

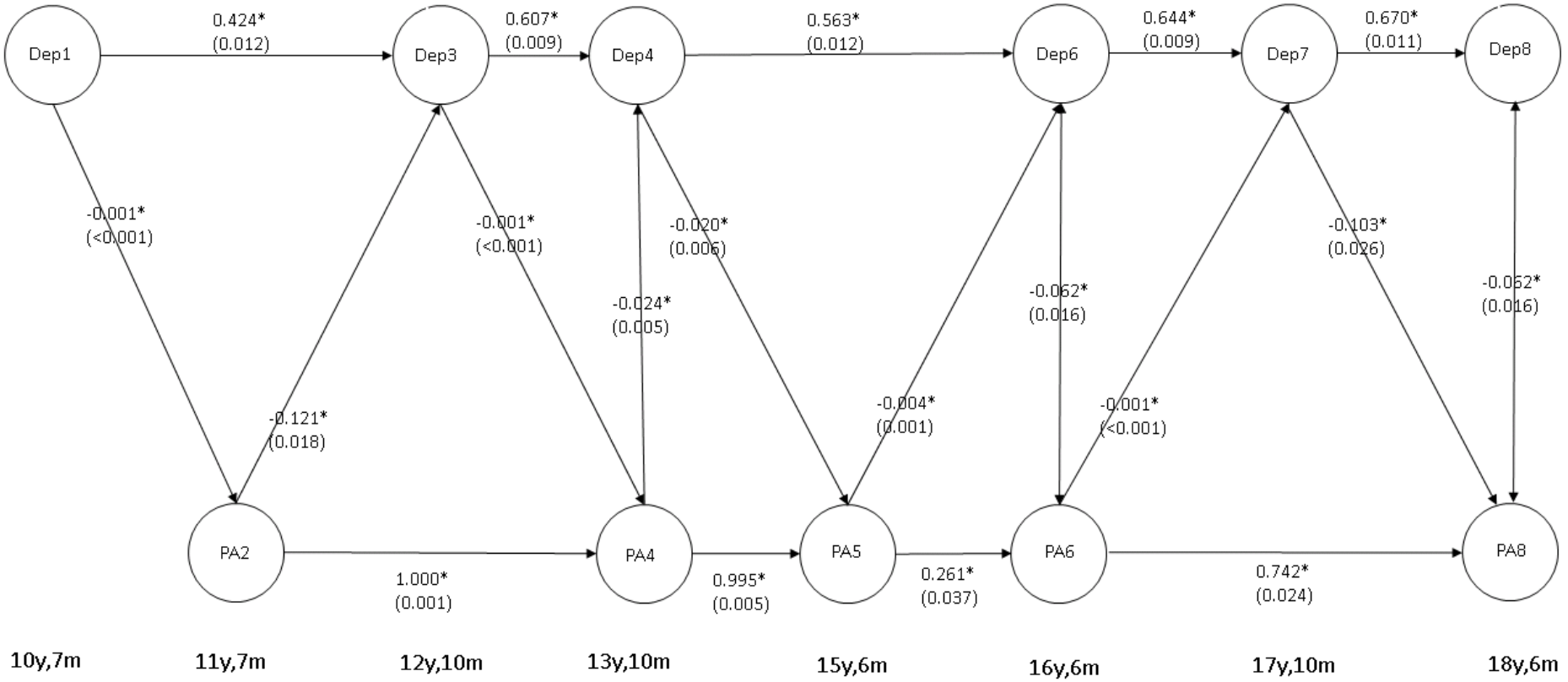


Model adjusted for age, socio-economic position (maternal education and maternal profession) and maternal depression.

Coefficients represent standard deviation change, figures in brackets represent standard error

**represents p-value <0.05*

Figure 6.4 - Cross lagged SEM in the ALSPAC cohort describing the relationship between physical activity (PA) and depressive (Dep) latent traits in females



Model adjusted for age, socio-economic position (maternal education and maternal profession) and maternal depression.

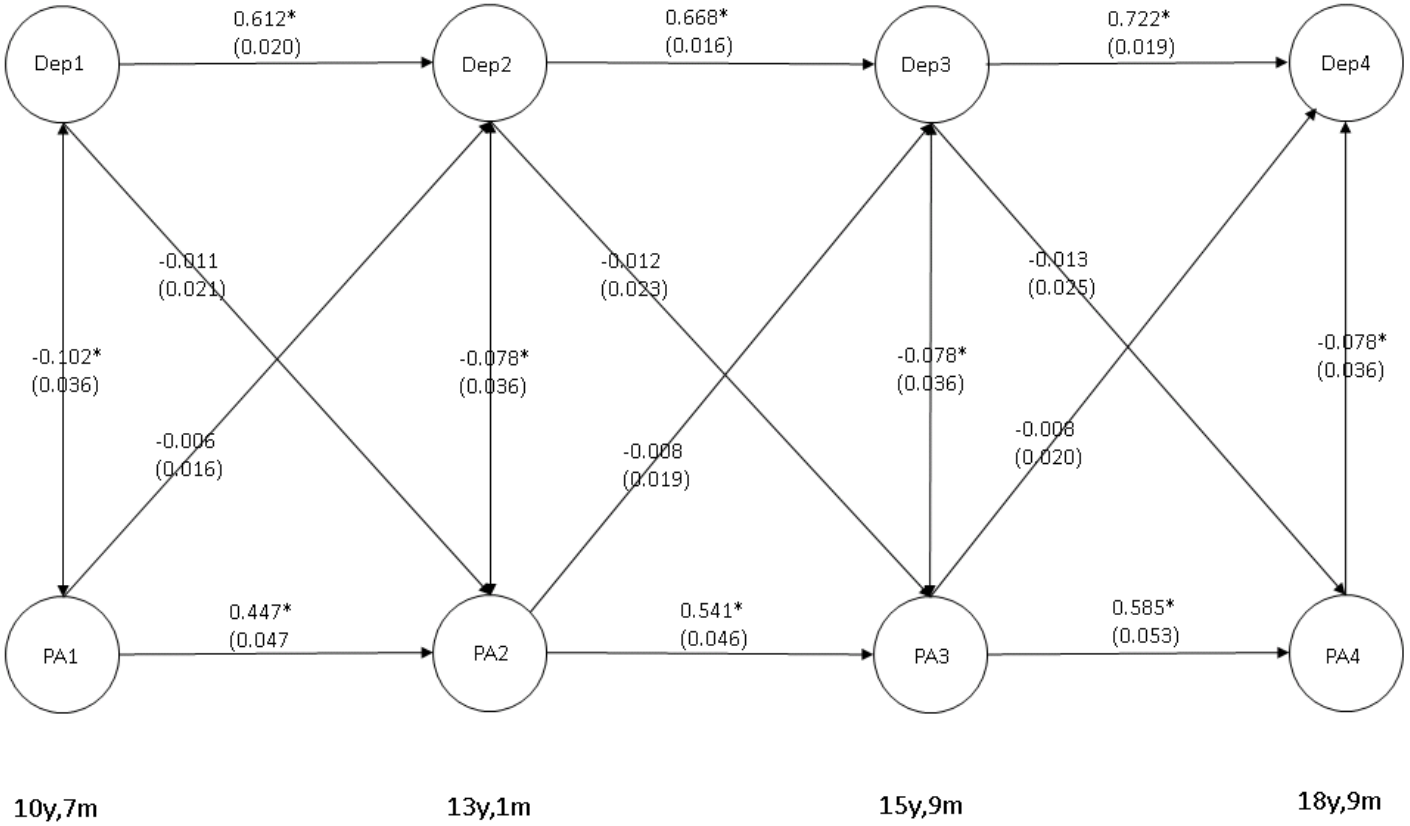
Coefficients represent standard deviation change, figures in brackets represent standard error

**represents p-value <0.05*

TRAILS

When the cross-lagged SEM was applied to the TRAILS cohort there was evidence of an autoregressive association in both level of depression and physical activity (i.e. PA at one time point was associated with PA at the previous time point, and depression at one time point was associated with depression at the previous time point). However there was no evidence of a cross-lagged association between PA and depression in either direction (i.e. no evidence of an association between PA latent trait score and depression latent trait score at the next time point or vice-versa) (Figure 6.5). The same pattern of results were found when the analysis was stratified by sex (Figure 6.6 and Figure 6.7).

Figure 6.5 - Cross lagged SEM in the TRAILS cohort describing the relationship between physical activity (PA) and depressive (Dep) latent traits

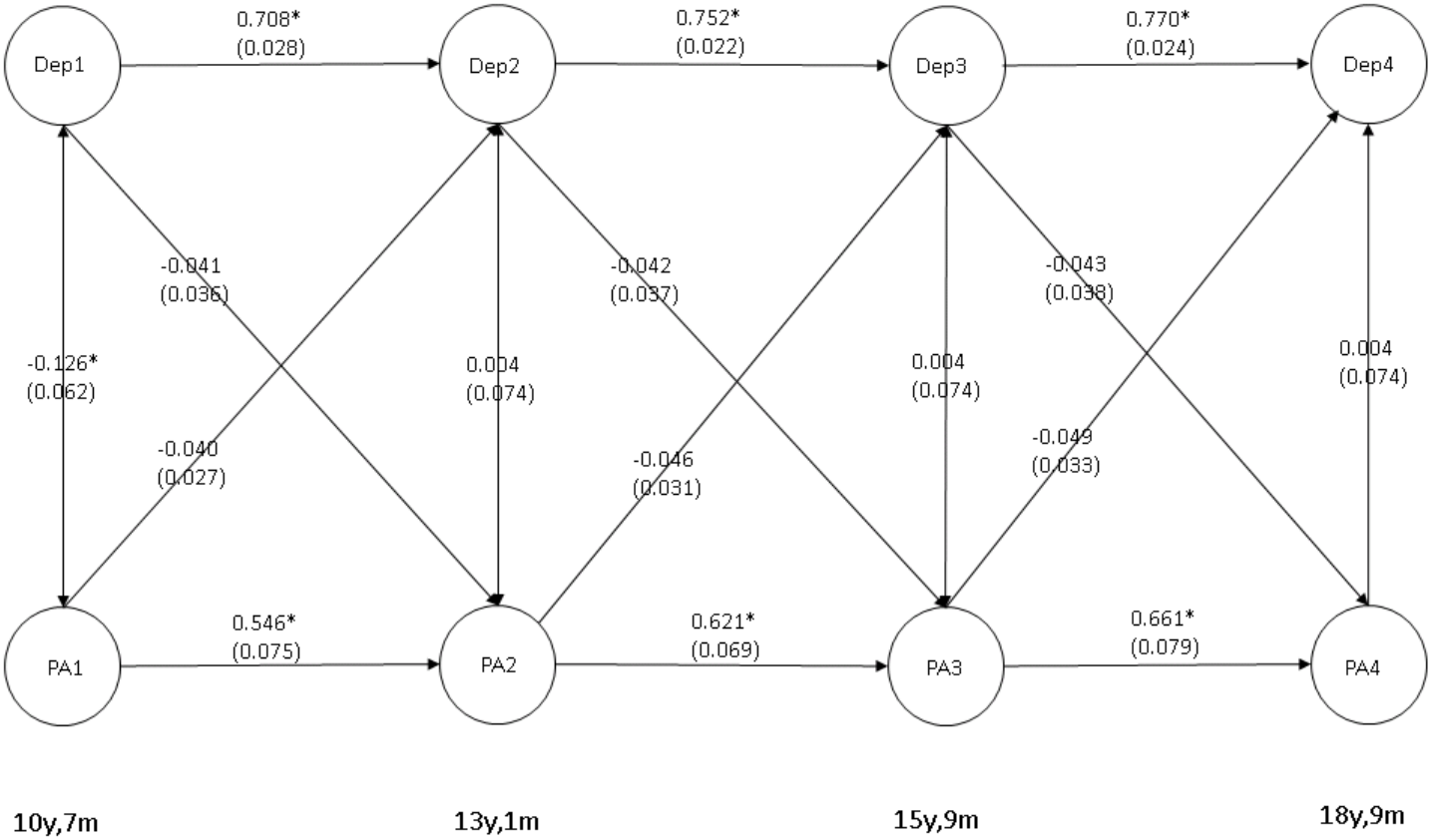


Model adjusted for age, sex, socio-economic position, maternal depression, smoking and alcohol use.

Coefficients represent standard deviation change, figures in brackets represent standard error

**represents p-value <0.05*

Figure 6.6 - Cross lagged SEM in the TRAILS cohort describing the relationship between physical activity (PA) and depressive (Dep) latent traits in males

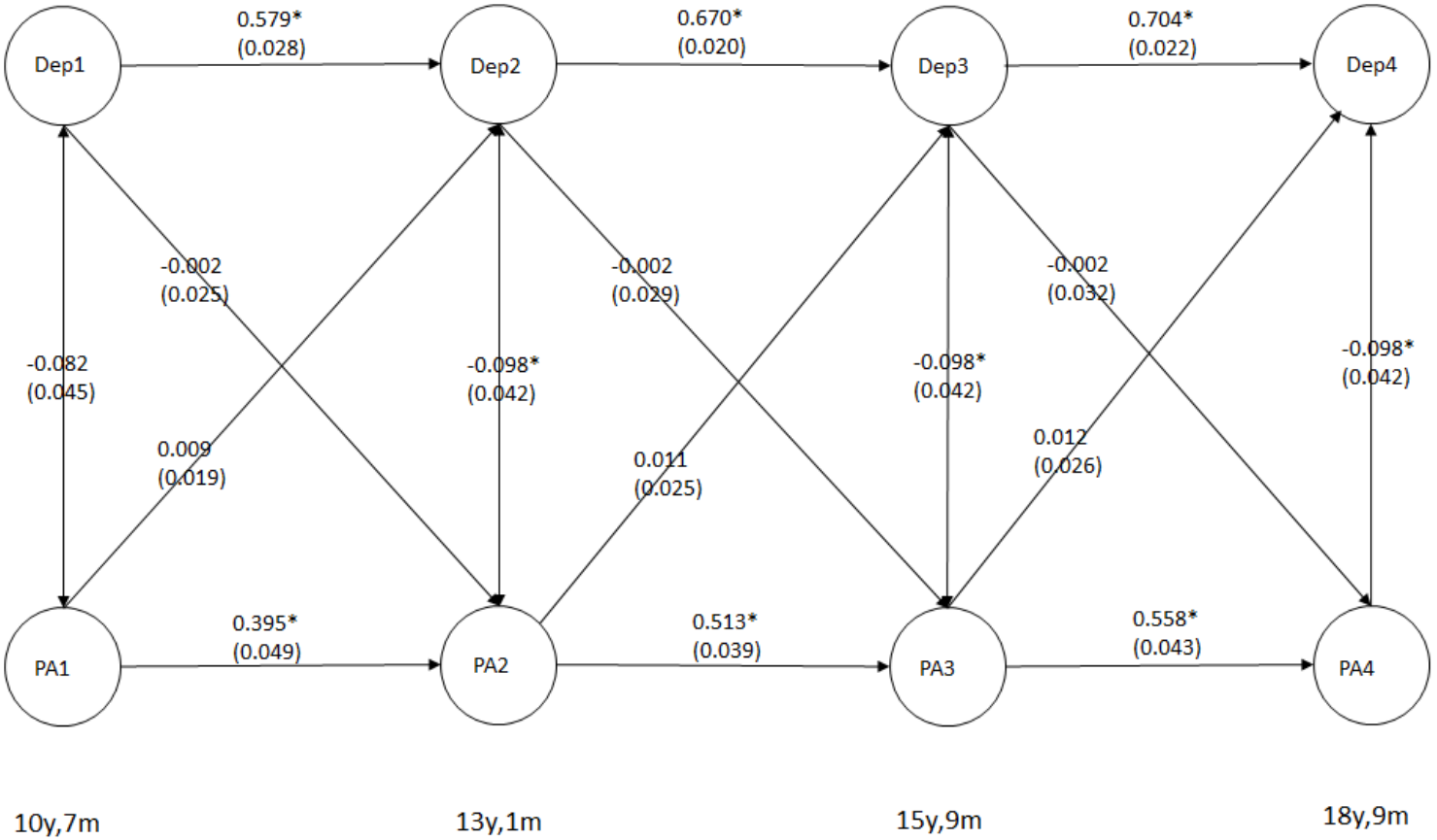


Model adjusted for age, socio-economic position, maternal depression, smoking and alcohol use.

Coefficients represent standard deviation change, figures in brackets represent standard error

**represents p-value <0.05*

Figure 6.7 - Cross lagged SEM in the TRAILS cohort describing the relationship between physical activity (PA) and depressive (Dep) latent traits in females



Model adjusted for age, socio-economic position, maternal depression, smoking and alcohol use.

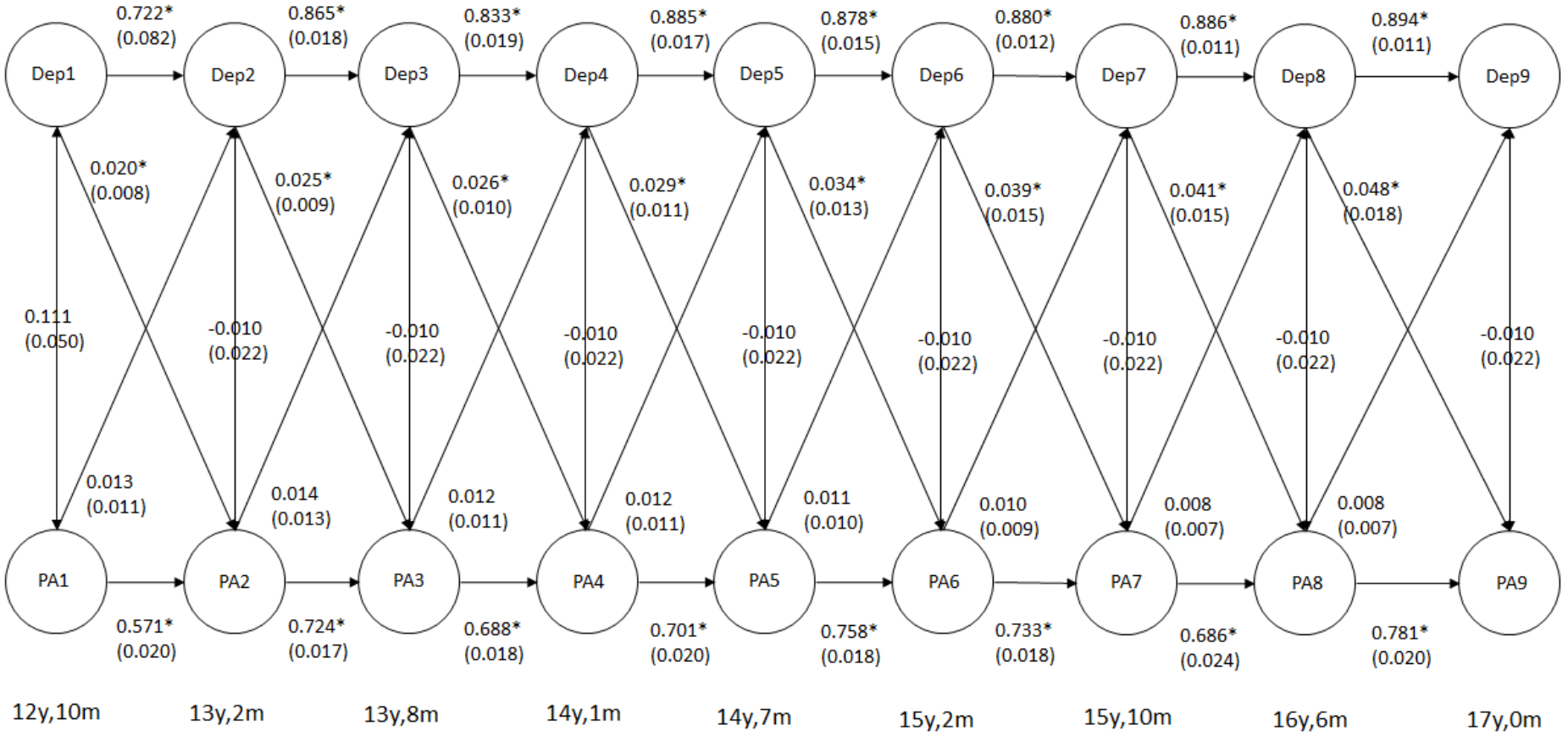
Coefficients represent standard deviation change, figures in brackets represent standard error

**represents p-value <0.05*

NDIT

When the cross-lagged SEM approach was used in the NDIT cohort again there was evidence of an autoregressive association in both depression and physical activity (Figure 6.8). There was also evidence of a cross-lagged association between depression latent trait and later physical activity. Those who were more depressed (higher depression latent trait score) were more active at the next time point (higher PA latent trait score). When the analysis was carried out stratified by sex there was no evidence of a cross-lagged association in either males or females (Figure 6.9 and Figure 6.10), this may be due to an issue of statistical power; when the analysis is stratified by sex there is a reduction in sample size and therefore a reduction in the ability to detect an effect.

Figure 6.8 - Cross lagged SEM in the NDIT cohort describing the relationship between physical activity (PA) and depressive (Dep) latent traits

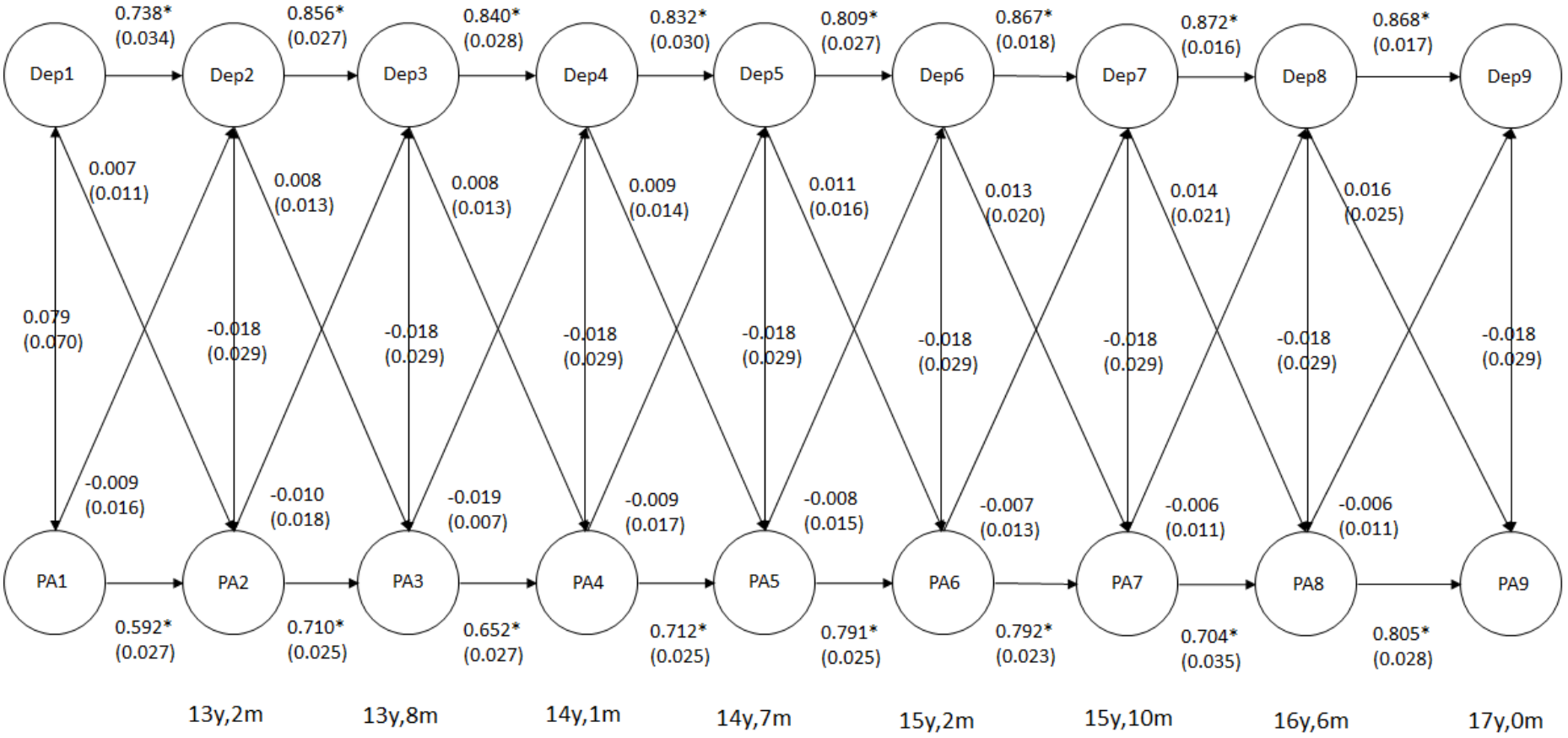


Model adjusted for age, sex, socio-economic position and alcohol use.

Coefficients represent standard deviation change, figures in brackets represent standard error

**represents p-value <0.05*

Figure 6.9 - Cross lagged SEM in the NDIT cohort describing the relationship between physical activity (PA) and depressive (Dep) latent traits in females

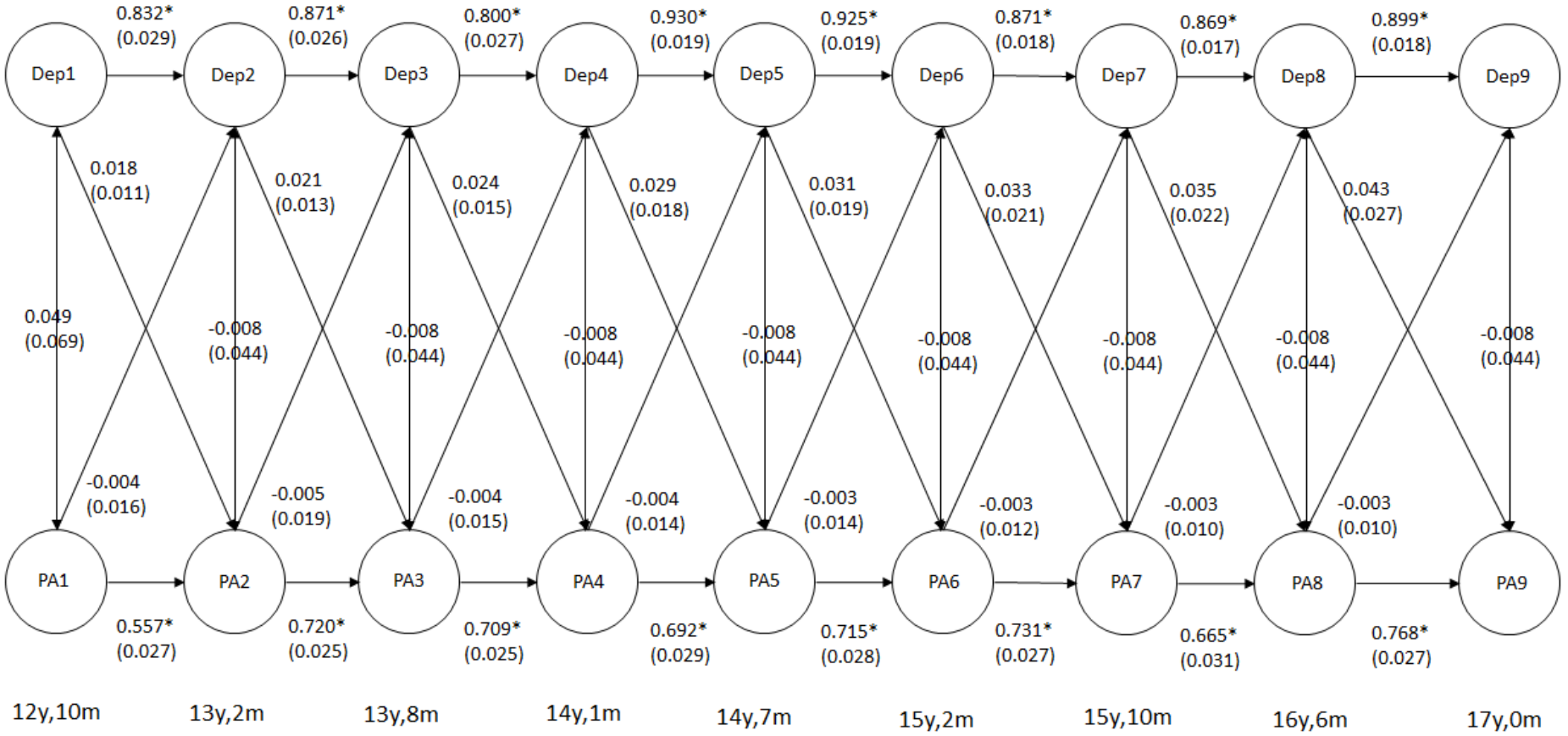


Model adjusted for age, socio-economic position and alcohol use.

Coefficients represent standard deviation change, figures in brackets represent standard error

**represents p-value <0.05*

Figure 6.10 - Cross lagged SEM in the NDIT cohort describing the relationship between physical activity (PA) and depressive (Dep) latent traits in males



Model adjusted for age, socio-economic position and alcohol use.

Coefficients represent standard deviation change, figures in brackets represent standard error

**represents p-value <0.05*

6.1.4. Mendelian Randomization

A split sample Mendelian randomization analysis (see Section 3.7.2) was carried out to examine the causal effect of PA on depression. SNPs for physical activity were available only in the ALSPAC cohort and as such the MR analysis was restricted to participants from ALSPAC.

The ALSPAC sample was first split into two groups, the physical activity prediction scores generated for the first sub-group were applied to the participants in the other sub-group and vice versa. The prediction scores were then used in 2SLS regression IV analysis (analysing the two sub-groups separately). The results from each sub-group suggested no evidence of an association between daily number of minutes spent in MVPA and depressive symptoms (Table 6.8). The results from the two sub-groups were then pooled and the results again suggested no evidence of an association between MVPA and depressive symptoms (pooled coefficient; -0.122, 95% CI; -0.551, 0.308, p-value; 0.579) (Table 6.8). It should be noted that the F statistic (a marker of the quality of the genetic instrument) from the first stage of the two stage least squares regression of subgroups one and two were both very low (0.02 and 0.35 respectively), and substantially lower than the widely accepted minimum value of 10 (as outlined previously in Section 3.7.2 used to define a "strong" instrument) so the results should be interpreted with caution.

Table 6.8 Results of MR analysis investigating the association between MVPA and depression in the ALSPAC cohort

	n	F statistic	Coefficient	95% CI	p-value
Subgroup 1					
MVPA	1225	0.02	-0.064	(-1.334, 1.206)	0.922
Subgroup 2					
MVPA	1225	0.35	-0.129	(-0.585, 0.327)	0.579
Meta-Analysis					
MVPA	2450	NA	-0.122	(-0.551, 0.308)	0.579

6.2. Summary of findings

In the analyses investigating a potential relationship between physical activity and depression in adolescence there was no consistent evidence of an association. The majority of analyses produced effect estimates with wide confidence intervals (which spanned the null), and there was inconsistency in direction of the association, with some model coefficients being positive such that an increase in PA was associated with an increase in depressive symptoms, whereas in other models, the coefficient was in the opposite direction (negative) such that an increase in PA was associated with a decrease in depressive symptoms. There was no evidence that the results differed by gender.

When comparing the results of the analyses using the objective and self-report measures of PA collected in the ALSPAC cohort, the regression analyses with the objective measures all provided no evidence of an association (95% confidence intervals were fairly symmetrical about the null). However, in the regression analyses of the self-report data, although the confidence intervals crossed the null the direction of the coefficients indicated that increased frequency of PA may be associated with a reduction in later depressive symptoms.

Using PLS-R groups of physical activities with different associations with depression were identified. However, there was no clear pattern in the grouping of these activities. From this analysis it was not possible to disentangle what aspect(s) of physical activity may be important in the relationship with depression in adolescence.

In the most robust analysis that was conducted, the cross-lagged SEM which makes full use of the repeated measurements available, there was evidence of an inverse association between physical activity and depression latent traits in the ALSPAC

cohort but not in the TRAILS or NDIT cohorts. Although there was no evidence of a cross-lagged relationship in the TRAILS or NDIT cohorts the direction of the coefficients were also inverse. However, the coefficients were very small, therefore there may be an association between PA and later depression in the TRAILS and NDIT cohorts but the analysis was not powered to detect it.

MR analysis is a potentially useful tool in terms of understanding more about causal relationships as it does not suffer from the problems of reverse causation and residual confounding associated with traditional epidemiological techniques. In the current study an MR approach was used to investigate the relationship between PA and depression in adolescence. Unfortunately the genetic instrument used in the analysis was very poor and as such the MR analysis was unable to shed any further light on the relationship between PA and depression in adolescence.

6.3. Strengths and Limitations

A limitation of this study is in the measurement of PA. Measurement error in the exposure will bias results towards the null. Although objective data on PA was available from accelerometers used in one cohort (ALSPAC) the majority of PA data was collected using self-report questionnaires. This is likely to be less precise than an objective measure and is potentially open to bias, participants may over report their level of activity due to social desirability, which (as highlighted above) would bias analyses towards the null (see Section 3.4.1). It also difficult to capture information on adolescent activity on a self-report questionnaire as young people are often active in short bursts [93, 94]. Furthermore, it is difficult to capture information regarding the intensity of activity on a self-report questionnaire compared to a measure such as accelerometer. The findings of the investigation were inconsistent between the results of the linear regression models for the objective versus self-report measures

of PA. This may be due to issues of measurement or it may reflect that the self-report and objective data represent slightly different aspects of PA. The self-report question asks about frequency of activity whereas the accelerometry variables are based on total amount and intensity of activity. Relying purely on an objective measure like accelerometery does not capture the context of any physical activity, and therefore not permit investigation of whether the context of activity is important. The measurement of physical activity is difficult - we can use an objective measure (e.g. accelerometry) to capture data on amount and intensity of activity but alongside these we need to record information about what the individual was doing. This would give us more information on the context and help us disentangle the role of different aspects of physical activity.

An issue that should be considered is the potential involvement of sedentary behaviour, as a distinct behaviour which is different from lack of physical activity. Sedentary behaviour is a group of behaviours that are carried out whilst sitting or lying down that require only a very low energy expenditure (e.g. watching television) [169-171]. For example it may be possible for someone to do enough physical activity to reach the recommended amount but to also spend a lot of time sedentary if they spend a lot of time sitting down (perhaps at a computer for work/school). Recently there has been a growing body of evidence suggesting that sedentary behaviour is associated with increased risk of depression [172]. Sedentary behaviour may be a confounder in the relationship between physical activity and depression, as such it needs to be accounted for. A recent methodological development known as compositional analysis [173] may offer a useful way to include data on both physical activity and sedentary behaviour in the same model to investigate their association with adolescent depression. An individual's total time can be thought of as the total time spent in vigorous activity, moderate activity, light activity, sedentary and asleep. This composition of daily time can be expressed as ratios of these individual aspects and it is these ratios (and changes in these ratios)

that may be relevant in the investigation with a health outcome (such as depression) and this is what can be explored in a compositional analysis [173].

Some of the strengths of this investigation were the same as described for the earlier analyses that examined the relationship between obesity and depression (see section 5.3) i.e. that longitudinal data was available, as such allowing investigation into direction of causality between PA and depression in adolescence rather than a cross-sectional association between these variables. Data was available on a large number of participants across multiple time points in the three cohorts, allowing analysis to be replicated in different cohorts to look for consistency in findings. There was also a large amount of information on important confounders, including confounders that have often been ignored by previous studies (e.g. maternal depression) (see section 2.2.1).

A variety of analytical methods were used in this study. A strength of this approach is that it allows investigation into whether findings are robust against different methods which make different assumptions. The SEM analysis also allowed investigation into a potential bi-directional relationship between physical activity and depression, the MR analysis attempted to investigate the relationship between physical activity and depression free from the issues of reverse causation and confounding associated with standard observational epidemiological approaches, and the PLS-R analysis allowed investigation into what aspects of PA may be important in the relationship with depression.

6.4. Comparison with previous studies

The literature surrounding the association between physical activity and depression in adolescence is sparse, a recent systematic review was only able to identify six longitudinal studies to include in the review [41]. All of the six previous studies in the systematic review used self-report measures of PA [86-91]. Of the six longitudinal studies that are available five found evidence of an inverse relationship between PA and depression in adolescence [86-89, 91], whilst one study found no evidence of an association [90], compared to the current study where the findings were inconsistent.

The self-report measures of PA used in the previous literature varied greatly, in the current study two of the three cohorts used self-report of frequency of PA in a week, the most comparable to this in the previous literature is Motl et al [89] (who asked if participants take regular exercise, answers were on a Likert scale from which a latent trait was derived) and Rothon et al [90] (who asked how many hours a week does the participant exercise in their free time). In a latent trait analysis Motl et al [89] found that an increase in PA latent trait was associated with a decrease in depression latent trait, this is similar to what was observed in the current analysis in the ALSPAC cohort, whereas Rothon et al [90] found no evidence of an association between PA and depression. The different findings in the two studies may be due (at least in part) to that fact that Motl et al [89] used a continuous measure of depression as the outcome variable whereas Rothon et al [90] used a binary "depressed/not depressed" outcome. This may also be a reason why the results of Motl et al [89] are similar to the current study whereas the results of Rothon et al [90] are not.

Importantly, none of the previous studies adjusted for previous depression meaning that any association observed may reflect a persistence of symptoms rather than a causal association between PA and depression. It should also be noted that other important confounders were not always adjusted for, for example maternal

depression was not adjusted for in any of the previous studies, only one of the studies adjusted for participant age [87], and one study did not adjust for social-economic status [88]. Hence the association observed in previous studies could be attributable to residual confounding. Although none of the previous studies carried out a formal test for an interaction by sex, two of the studies carried out analyses separately in males and females; there was no evidence that the relationship between PA and depression in adolescence differed by sex, this was the same as was found in the current study. None of the previous studies have attempted to understand the potential importance in differences between frequency, intensity and context of activity (which was highlighted as an issue by the earlier systematic review in this area [41]). The analyses presented in this thesis attempted to explore the importance of these different aspects of PA, however there was no evidence for either amount or intensity of PA being associated with later depression in adolescents.

6.5. Implications and Future work

There was no robust evidence of an association between physical activity and depression in adolescence. However, this may be due to challenges in the measurement of PA, both in terms of accuracy of measurement and also understanding the wider context of that activity (for example, if the activity is undertaken alone or with others etc.). Therefore future studies need to collect longitudinal repeated measures data on amount, frequency, intensity and context of physical activity. This future work could, for example, use a combined objective and self-report measure of physical activity, e.g. accelerometry in combination with a structured activity diary where participants record the type of activity that they are carrying out which can then be matched to the accelerometry data. This would provide more in-depth information on PA, recording data about the amount, frequency, intensity and context of activity and allow the use of methods such as compositional analysis and partial least squares regression to try and identify which aspects of PA relate to depression.

Any future analyses also need to account for the potential involvement of sedentary behaviour and its' potential interplay with levels of physical activity and depression. One way to do this would be to use a compositional analysis [173]. To carry out this analysis data would need to be collected on time participants spent sedentary and in light, moderate and vigorous activity (which could all be captured by accelerometry data).

Currently MR analysis is not appropriate for investigating the relationship between PA and depression in adolescence due to the lack of a robust genetic instrument for PA. This does not however necessarily rule out an IV approach to aid in causal inference but would require a non-genetic instrumental variable to be identified. This non-genetic IV could perhaps be change in adolescent PA in response to a policy intervention or the Olympics taking place. It may be possible for future studies to explore the use of an IV approach if an appropriate instrument could be identified.

Further work needs to be carried out to disentangle the relationship between physical activity and depression in adolescents. A cohort where all the required measures are already being collected in the way that is needed may not exist. However, rather than setting up an entirely new cohort study if a cohort where participants are of the appropriate age could be identified, through collaboration with the cohort research team it may be possible to insert the required measures into future data collection. For example the Born in Bradford [174] cohort has collected longitudinal data on study children who are now approximately 9 years old, whilst the participants of the Growing up in Wales [175] cohort are younger but may still present an opportunity for collaboration. Forming collaborations with these (and other) cohorts may allow the required measurements needed for further investigation into the relationship between PA and depression to be inserted and

collected on participants in future follow up occasions throughout adolescence (and even into adulthood).

CHAPTER 7. RESULTS AND DISCUSSION - OBJECTIVE 3; MEDIATION OF THE RELATIONSHIP BETWEEN OBESITY AND DEPRESSION

The cross-lagged SEM models that were used to investigate the potential bi-directional relationship between obesity and depression in adolescence were extended to examine a number of potential mediators of the relationship between obesity and depression. Potential mediators included both biological (cortisol, CRP and IgE) and psychosocial factors (body image and self-esteem). As outlined earlier (Section 3.5), analyses were restricted to the ALSPAC and TRAILS cohort which had available data on potential mediators of interest.

ALSPAC

There was evidence that body image mediated the association between obesity and depression in the ALSPAC cohort (Table 7.1 row 5 and Figure 7.1). The indirect effect of obesity on depression via body image was estimated as 0.065 (SE 0.010), hence a one standard deviation change in the obesity latent trait score at 15 years 6 months was associated with a 0.065 standard deviation increase in depression latent trait at 16 years 6 months due to the effect of obesity on body image. When the analysis was carried out separately in females and males, there was evidence of mediation via body image (Figure 7.2) in females but not in males (Figure 7.3). In females the indirect effect obesity on depression via body image was estimated as 0.070 (SE 0.011), hence a one standard deviation increase in obesity latent trait score

was associated with a 0.070 standard deviation increase in depression latent trait due to the effect of obesity on body image.

There was no evidence that C-reactive protein mediated the relationship between obesity and depression in adolescents in the ALSPAC cohort (Table 7.1 row 6). The estimate of the indirect effect was 0.010 with a standard error of 0.007. The mediation analysis was repeated investigating males and females separately. There was no evidence that CRP mediated the relationship between obesity and depression in either males or females (see Appendix 8).

TRAILS

When cortisol was included as a potential mediator in the SEM model examining the association between obesity and depression in the TRAILS cohort there was no evidence of an indirect effect of obesity on depression through cortisol (indirect effect 0.005, SE 0.003), i.e. no evidence of mediation (Table 7.1 row 9). When the analysis was carried out separately in male and females there was no evidence that cortisol mediated the association between obesity and depression in either males or females (see Appendix 8).

When C-reactive protein was included in the SEM model as a potential mediator there was no evidence of an indirect effect of obesity on depression through CRP (indirect effect 0.006, SE 0.005), i.e. no evidence of mediation (Table 7.1 row 10). When the analysis was carried out separately in male and females there was no evidence of mediation via CRP in either males or females (see Appendix 8).

When IgE was included in the SEM model as a potential mediator there was no evidence of an indirect effect of obesity on depression through IgE (indirect effect 0.011, SE 0.009), i.e. no evidence of mediation (Table 7.1 row 11). When the analysis

was carried out separately in male and females there was no evidence of mediation via IgE in either males or females (see Appendix 8).

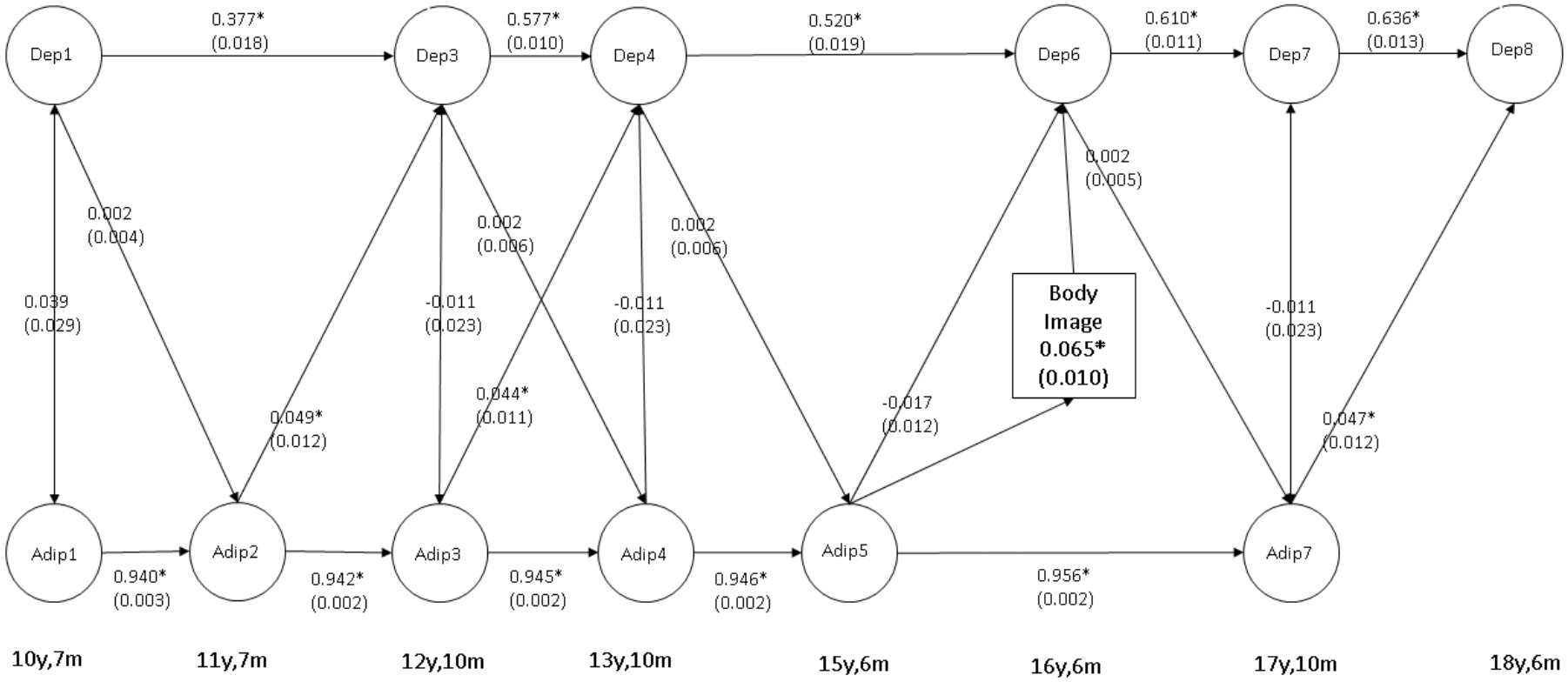
When body image was included in the SEM model as a potential mediator there was no evidence of an indirect effect of obesity on depression through body image (indirect effect 0.006, SE 0.004), i.e. no evidence of mediation (Table 7.1 row 12). When the analysis was carried out separately in male and females there was no evidence of mediation via body image in either males or females (see Appendix 8).

When self-esteem was included in the SEM model as a potential mediator there was no evidence of an indirect effect of obesity on depression through self-esteem (indirect effect 0.007, SE 0.006), i.e. no evidence of mediation (Table 7.1 row 13). When the analysis was carried out separately in male and females there was no evidence of mediation via self-esteem in either males or females (see Appendix 8).

Table 7.1 - Results of the analyses investigating potential mediators on the causal pathway between obesity and depression

Mediator	Age at obesity measurement	Age at mediator measurement	Age at depression measurement	Indirect effect coeff. (SE)	Direct effect coeff. (SE)
ALSPAC					
Body Image	15y 6m	16y 6m	16y 6m	0.065 (0.010)	-0.017 (0.012)
CRP	15y 6m	15y 6m	16y 6m	0.010 (0.007)	0.038 (0.010)
TRAILS					
Cortisol	13y 1m	15y 9m	15y 9m	0.005 (0.003)	0.010 (0.009)
CRP	13y 1m	15y 9m	15y 9m	0.006 (0.005)	0.008 (0.006)
IgE	13y 1m	15y 9m	15y 9m	0.011 (0.009)	0.006 (0.005)
Body Image	13y 1m	15y 9m	15y 9m	0.006 (0.004)	0.007 (0.007)
Self-esteem	15y 9m	18y 9m	18y 9m	0.007 (0.006)	0.009 (0.007)

Figure 7.1 – Structural equation model to investigate the role of body image as a mediator between obesity and depression in the ALSPAC cohort

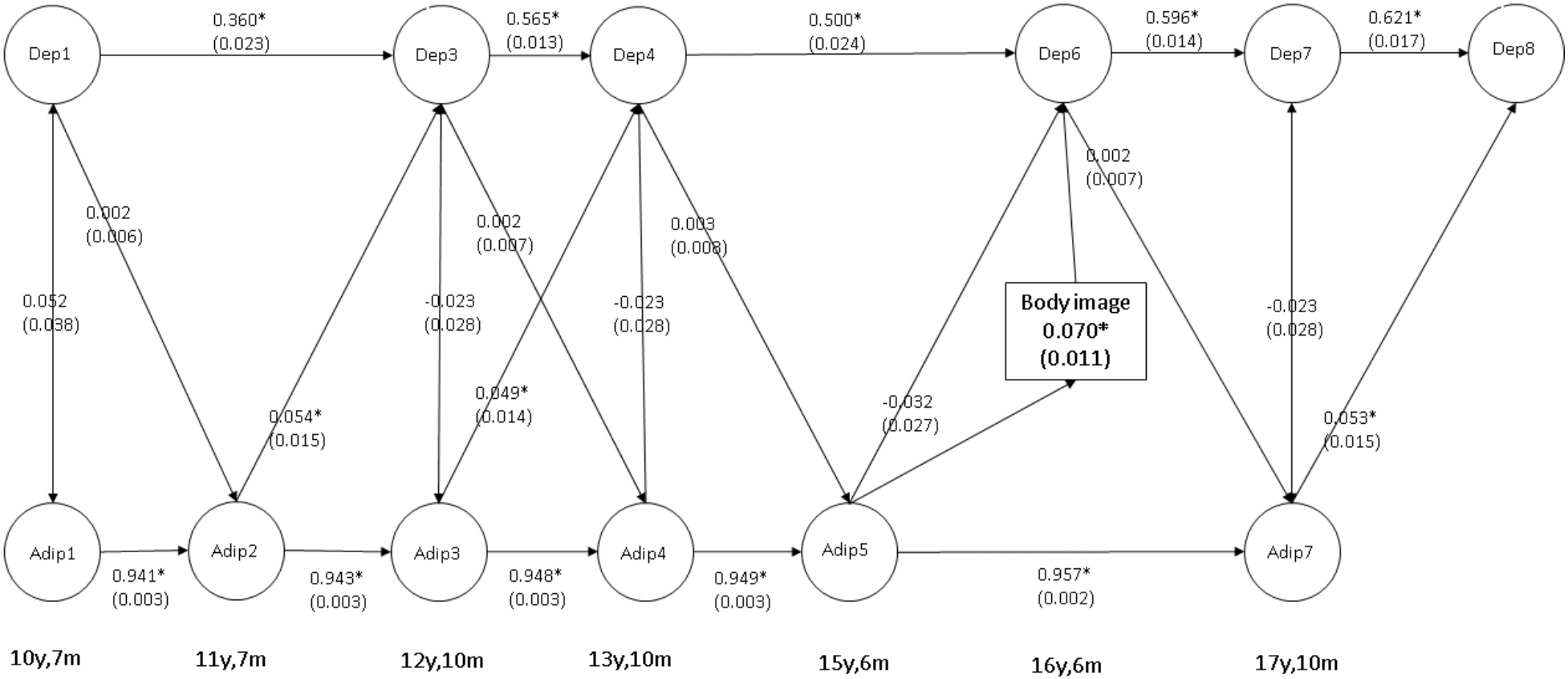


Model adjusted for age, sex, socio-economic position (maternal education and maternal profession) and maternal depression.

Coefficients represent standard deviation change, figures in brackets represent standard error

**represents p-value <0.05*

Figure 7.2 – Structural equation model to investigate the role of body image as a mediator between obesity and depression in females in the ALSPAC cohort

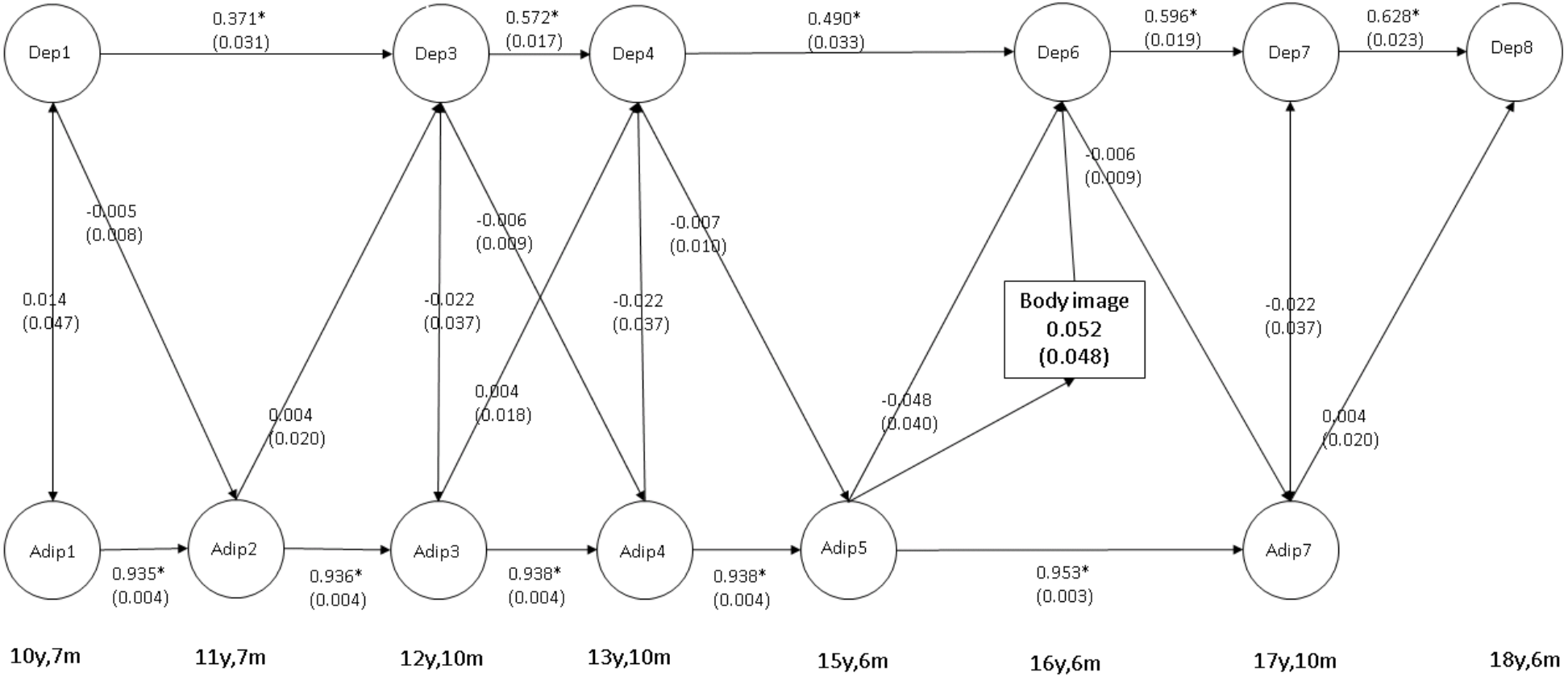


Model adjusted for age, socio-economic position (maternal education and maternal profession) and maternal depression.

Coefficients represent standard deviation change, figures in brackets represent standard error

**represents p-value <0.05*

Figure 7.3 – Structural equation model to investigate the role of body image as a mediator between obesity and depression in males in the ALSPAC cohort



Model adjusted for age, socio-economic position (maternal education and maternal profession) and maternal depression.

Coefficients represent standard deviation change, figures in brackets represent standard error

**represents p-value <0.05*

7.1. Summary of findings

There was no evidence that any of the biological factors investigated; cortisol, CRP or IgE, mediated the relationship between obesity and depression in adolescence. When potential psychosocial mediators were investigated, there was no evidence of mediation via self-esteem. There was however evidence that body image mediated the relationship between obesity and future depression in the ALSPAC cohort, but not in the TRAILS cohort (effect was in the same direction but much smaller and CI spanned the null).

7.2. Strengths and Limitations

As discussed in section 5.3 there are a number of issues that may be relevant to the overall association between obesity and depression; such as the potential influence of puberty and social context. These factors need also to be accounted for when investigating potential mediators. For example it is possible that social context could play an important role when considering perception of body image as potential mediator. If body image mediates the relationship between obesity and depression but different groups value certain body types differently then this would need to be accounted for in the mediation analysis [160, 161]. It would be similarly plausible to suggest that body image may be affected by puberty [176], and therefore puberty should be included in mediation analyses (as an interaction effect).

Another potential limitation of the study is related to the measurement of the mediator variables. In this investigation the mediation analysis was carried out at only one time point (due to the availability of data). The fact that data on the potential mediators was only collected at one time point meant that previous levels of the mediator could not be adjusted for. Therefore the mediator could be continuing at the same level and not causally associated (i.e. the level of the mediator

has not changed in response to changes in the explanatory variable). The mediator variables were also measured at the same follow up occasion as the outcome. Therefore although this is an improvement on fully cross-sectional data, it is not possible to determine the temporal order between the mediator and outcome variables.

Psychosocial constructs that we cannot directly observe, such as perception of body image and self-esteem, are difficult to measure. As such these variables are likely to suffer from problems of measurement error, this error will result in lower power to detect mediation via these variables, underestimate the mediated effect and overestimate the direct effect. This may, at least in part, explain the inconsistency in the mediation analysis findings regarding body image between the ALSPAC and TRAILS cohorts.

A strength of this study was the method used in the investigation into potential mediators. The use of the SEM (with bootstrapping) approach allows the direct quantification of the indirect effect via a mediator (i.e. the mediated effect). The use of SEM also reduces the problems of measurement error in the measurement of the exposure and outcome variables through the use of latent traits for obesity and depression. In the analysis of the mediator data several assumptions are being made however; firstly that there is no unmeasured confounding, that there is no interaction between the exposure (obesity) and the mediator variables, there are no confounders of the mediator-outcome association that are influenced by the exposure, and that we are testing for a linear association. With regards to the assumption of unmeasured confounding, this is an issue in all observational epidemiology and is not testable given only the observed data. Interactions between an exposure variable and a mediator and the related influence on an outcome are difficult to conceptualise. A method is available that allows for interaction on mediation by decomposing indirect effects into different components [177], but

currently this method is only available for one exposure, mediator and outcome measurement, rather than repeated measures.

7.3. Comparison with previous literature

Although there have been studies investigating the relationship between obesity and variables such as body image [178] that have been proposed as mediators of the obesity and depression relationship, the literature formally investigating potential mediators of an obesity depression relationship in adolescence is sparse (see Section 2.1.4). Only three studies were identified investigating mediators of an obesity-depression relationship in an adolescent population [81-83].

The three studies that are available have all used the Baron and Kenny approach to mediation, this method has been criticised for a number of reasons (see Section 3.7.3) [84]. Two of the three studies in the previous literature were also cross-sectional in nature making it impossible to establish any temporal order between the exposure, mediator and outcome variables. Only one of the previous studies investigated mediation in adolescence whereby obesity was the exposure and depression the outcome. This previous study [82] was a cross-sectional sample from a prospective cohort study that used a binary obese/not obese variable for the exposure (based on self-report BMI), a continuous depression score as the outcome and a continuous body image score as the mediator (see section 2.1.4). In comparison, our current study used continuous latent traits for the obesity and depression variables, whilst a continuous body image score was investigated as a mediator. The previous study [82] found an increase in depressive mood was fully mediated by an increase in body dissatisfaction in obese males but not females. In contrast, although the results of my project were inconsistent between cohorts, there was evidence in the ALSPAC

cohort that body image acted as a partial mediator between obesity and depression in females but not in males.

7.4. Implications for future work

Understanding the factors that may lie on the causal pathway from obesity to depression in adolescence is important as it may help in the identification of novel intervention targets. There was some evidence that body image may mediate the obesity to depression relationship in adolescence and hence this may be worthy of further investigation. However, before further mediation analyses are carried out, the precise relationship between obesity and depression in adolescence requires further disentangling as discussed in section 5.5. To further investigate mediation via body image (or other potential mediators) then longitudinal data with repeated measures of obesity, the mediator(s) and depression should be used to help establish the direction of causality. The results of Objective 1 suggest that future work into the relationship between obesity and depression in adolescence should focus on females (see Section 5.2), therefore further mediation analyses should be carried out stratified by sex, with particular interest given to mediation of the obesity-depression relationship in females.

To improve on the issue of measurement error in mediator variables such as body image, which are constructs that we cannot directly observe, then future mediation studies could use latent variables (i.e. a latent body image trait) rather than self-report scores for these variables. A latent trait will of course have to be based on collected observed data, however this observed data could come from a variety of different questions that may relate to the underlying latent trait and not simply a self-report score for the mediator of interest. Future mediation studies will also need to consider the potential impact of other variables such as pubertal stage and social context as interactions and or in sub-group analyses. To do this future studies will

need to measure these factors, including a wide range of pubertal stages and social contexts and then carry out appropriate analyses. To help with the issue of reverse causality then repeated longitudinal measures of the exposure, mediator and outcome variables should be collected.

A further way to help establish causality in future mediation studies could be to use Network Mendelian Randomization (NMR) [179]. In a NMR analysis an instrumental variable is used in place of the observed exposure variable (as in a standard MR analysis) whilst another instrumental variable is used in place of the observed mediator variable. Mediation can then be investigated in NMR by extending the standard MR 2SLS regression or within an SEM framework. This type of analysis would help with the problems of measurement, confounding and direction of effects common to mediation analysis of observational data, but would require the identification of instrumental variables for the appropriate mediators. This may prove difficult however when considering potential psychosocial mediators such as body image.

Collecting more robust data on potential mediators of the obesity-depression relationship together with application of sophisticated statistical approaches to addressing this question, will hopefully permit greater insight into this important association. This may lead to the identification of novel intervention targets in the prevention of adolescent depression.

CHAPTER 8. CONCLUSION

A detailed discussion of each objective of the project was given at the end of each results chapter. This chapter will give a brief summary of the key findings of the project along with the implications of these findings in terms of preventative strategies for depression in adolescence and suggestions for future work.

8.1. Key findings

There was evidence (albeit inconsistent between cohorts) of a positive relationship between obesity and depressive symptoms in adolescent females; an increase in the measure of obesity was associated with an increase in depressive symptom score at the next follow up occasion (Figure 8.1). When potential mediators of the relationship between obesity and depression in adolescence were investigated there was evidence (in one cohort) of partial mediation via participant perception of body image (Figure 8.1). An increase in obesity was associated with an increase in negative self-perception of body image which in turn was associated with an increase in depression. There was no evidence of mediation via the other psychosocial or biological variables investigated (self-esteem, cortisol, CRP and IgE).

When investigating the potential relationship between physical activity and depression in adolescence there was no consistent evidence of any association (in either objectively measured or self-report physical activity) (Figure 8.1). The direction of effect was often suggestive of an inverse relationship but with the confidence intervals spanning the null. This may however be due to issues in the measurement of PA and the fact that we are investigating potentially very small effects.

Figure 8.1 – Summary of key findings

Objective 1: Obesity and depression in adolescence

In the linear regression and GEE analyses there was inconsistent evidence of a positive relationship between obesity and later depressive symptoms in females and the results were fairly consistent in finding no evidence of an association in males. Below are the results from the most robust analysis: cross-lagged SEM investigating the relationship between obesity and depression at the next time point in females.

ALSPAC

- Females: Evidence of a positive association between obesity and later depressive symptoms. A 1 SD increase in obesity latent trait score was associated with a 0.054 SD (SE 0.015) increase in depression latent trait score at the next time point

TRAILS

- Females: No evidence of an association. A 1 SD increase in obesity latent trait score was associated with a 0.017 SD (SE 0.018) increase in depression latent trait score at the next time point

NDIT

- Females: No evidence of an association. A 1 SD increase in obesity latent trait score was associated with a -0.047 SD (SE 0.064) decrease in depression latent trait score at the next time point

Meta-analysis

- Females: Evidence of a positive association between obesity and later depressive symptoms. When estimates were pooled a 1 SD increase in obesity latent trait score was associated with a 0.035 SD (95% CI 0.003, 0.067) increase in depression latent trait score at the next time point.

Objective 2: Physical Activity and Depression

In the linear regression and GEE analyses there was no consistent evidence of an association between PA and later depressive symptoms. However, the direction of the coefficients often suggested a potential inverse relationship. Below are the results from the most robust analysis: cross-lagged SEM investigating the relationship between PA and depression at the next time point.

ALSPAC

- Evidence of an inverse relationship. An increase in physical activity latent trait score was associated with a (small) decrease in depression latent trait score at the next time point (e.g. a 1 SD increase in PA latent trait score was associated with a -0.005 SD (SE 0.001) decrease in depression latent trait score at the next time point).

TRAILS

- No evidence of an association between physical activity and later depressive symptoms. Although coefficients are negative and of a similar magnitude to those observed in the ALSPAC cohort (e.g. a 1 SD increase in PA latent trait score was associated with a -0.006 SD (SE 0.016) decrease in depression latent trait score at the next time point).

NDIT

- No evidence of an association between physical activity and later depressive symptoms (e.g. a 1 SD increase in PA latent trait score was associated with a 0.012 SD (SE 0.011) increase in depression latent trait score at the next time point).

Objective 3: Mediation

- Evidence of mediation of obesity to later depressive symptoms via body image in females (in the ALSPAC cohort). A 1 SD increase in obesity latent trait score was associated with a 0.070 (SE 0.011) SD increase in depression latent trait score due to the effect of obesity on body image.

8.2. Limitations

Discussion of specific limitations relevant to the analyses of the three objectives of this study can be found in sections 5.3, 6.3 and 7.2. In this section more general limitations of the study as a whole will be discussed.

There are problems that are inherent to any study based on observational data, namely; selection bias, confounding and reverse causation.

Selection bias can occur when the actual study sample is not a truly random sample of the intended study population. For example, depending on the nature of the study, an observational cohort may attempt to sample individuals from various different regions or certain schools, but some individuals may be less likely to participate in a study, or more likely to drop out. Selection bias occurs if the exposure and the outcome both affect selection into (or dropout from) a study, as collider bias will then induce an association between them in the observed study sample [180]. Even a well planned study will face problems due to the nature of requiring individuals to consent to take part in a study. Not all individuals who are selected to take part in the study will consent. It is possible that certain groups may be more or less likely to consent to take part in a study, potentially introducing bias. Similarly those individuals who consent to take part in a study may differ in certain aspects from those who do not, but it is usually only possible to investigate this superficially, by comparing study data to routinely collected data [181].

A related issue is that of generalisability – where the results of a study may be unbiased, but may not be generalizable to other populations. For example, if there are interactions/modifiers/moderators then the study effect estimates will differ across levels of the modifier/moderator (e.g. if a risk factor interacts with the effect of obesity on depression, then the estimated effect of obesity on depression will not

generalise to populations with different risk factor distributions). A well planned study will attempt to include participants from a wide range of areas (i.e. from both high and low income areas, a wide range of cultural backgrounds) but it is still very difficult to make sure the study is truly representative.

In this thesis attempts have been made to minimise the potential issue described above (i.e. selection bias). Selection bias can occur if the exposure and the outcome both affect selection into (or dropout from) a study. For example, in the investigation into the relationship between obesity and depression, if there is a positive relationship between obesity and depression but individuals with high levels of obesity and/or depression do not consent to take part in the study then the results of the observed data would be biased towards the null and not reflect the "true" relationship that exists in the population. The ALSPAC cohort is a prospective birth cohort, as such it is not possible that selection into the study could be affected by adolescent obesity, physical activity or depression of the child. In the TRAILS and NDIT cohorts the recruitment strategy attempted to include a representative sample, however it is possible that due to individuals needing to consent to participate the actual sample recruited into the study may differ (with respect to the exposure and outcome) from the wider population. As there is no information on those individuals who did not consent to participate it is not possible to test this. However, future work could use external information (e.g. expected proportion of males and females in the general population) to investigate variables that may be related with study participation, differences between the study population and the wider population and to derive bias adjusted estimates. The three cohorts all attempted to recruit samples that are generalisable and representative of the wider population, however as would be expected for longitudinal cohorts which have collected repeated measures data there was drop out from the three cohorts (ALSPAC, TRAILS and NDIT) which provided data for this thesis. Missing data will always reduce the precision of an analysis and may introduce bias if missingness is related to the exposure and outcome variables. To address the potential issue of missing data the

method of cross-lagged SEM was utilised in the thesis. Cross-lagged SEM uses a maximum-likelihood approach to estimation, as such both individuals with complete and incomplete data contribute information to the analysis. However, this method does rely on untestable assumptions about the mechanism of missingness. See Appendix 9 for a further discussion of the potential bias introduced by missing data.

Observational data is also susceptible to the problem of confounding. Although known and measured confounding factors may be adjusted for in statistical analyses, investigations are still limited by what confounders have been measured in the data. Observational studies will attempt to collect information on a wide range of confounding variables, however if a specific investigation into a particular relationship was not considered when the study was planned then not all of the relevant confounders may have been measured. Residual confounding may also be present if the confounders that have been measured were done so with a large amount of error. Even if all known confounders of a relationship are measured accurately and adjusted for appropriately in an analysis there is still the problem of potentially unknown confounders. There may be variables that we are not aware of that confound a relationship and as such we do not measure and adjust for them. This is one of the motivations for the use of MR analysis which helps to overcome these issues of residual confounding [60].

In an attempt to minimise the potential issue of confounding in analyses based on observational data, important confounders that were measured by the cohorts have been adjusted for in the analyses conducted as part of the thesis, whilst the influence of certain other potential confounding factors (such as puberty) have also been investigated as part of sensitivity analyses. There is however still the potential for residual confounding, for example, the measures of smoking and alcohol used as confounders were self-report in nature. Therefore these factors may be measured with error and potentially biased. For example, if there is a positive association between alcohol use and obesity, and alcohol use and depression but individuals

under-report their alcohol use (e.g. potentially due to social desirability) then this systematic error in the confounder would bias the results of an analysis of the relationship between obesity and depression away from the null (i.e. overestimate any effect). It is also possible that there are other variables not collected by the cohorts that would have been useful to include as confounders in the analyses, a potential example of this could be stressful life events; where it would be plausible to hypothesise that this factor may be related to both the exposure and outcome. To further address the issue of confounding MR analyses have been carried out. Due to the random way in which genes are inherited this type of analysis is a method that is useful in overcoming the problem of confounding that more traditional epidemiological techniques are susceptible to.

Studies based on observational data, even those which have collected longitudinal measures, cannot rule out reverse causation. For example if the first measurement taken in a cohort study is when participants are age 10, it is not possible to know what happened before this time point. When investigating the relationship between two variables, A (exposure) and B (outcome), it is not possible to identify whether B caused A before the first measurements were taken without further information. This is further motivation for studies to collect information on instrumental variables (both genetic and non-genetic).

The cross-lagged SEM analysis goes some way to addressing reverse causality. In this analysis the directionality of the association over time between the variables of interest (e.g. obesity and depression) is explicitly investigated. However, there is still the potential issue that we are only able to analyse data at the time points at which measurement took place. MR analysis is a useful tool in addressing the potential problem of reverse causation. As genes are determined at conception an analysis using a genetic instrument as an exposure variable cannot be susceptible to reverse causation.

A further potential issue in this study is in the length of time between follow up measures. If obesity or physical activity is causally related to later depressive symptoms in adolescence it is not known whether this effect will take place very quickly or whether it will take a long time to manifest. Therefore it's not known how long a time gap there should be between follow up measurements, and as such the time points used in this study could be either too close or too far apart to be able to investigate a potential effect. This means that analyses investigating lagged effects could be structurally misspecified. However, given the nature of the data collected it is difficult to explore this further.

Another potential issue is that of multiple testing. Multiple testing refers to the situation whereby as part of an analysis many statistical tests are carried out. This is potentially problematic as the greater the number of statistical tests performed the greater the likelihood of obtaining a statistically significant result by chance. As part of this thesis several hypothesis tests have been performed, therefore it is possible that some findings with an associated p-value of less than 0.05 may in fact represent false positives. In this situation, where multiple analyses are used to investigate the relationship between an exposure and outcome, it is more appropriate to look for consistency in direction and magnitude of estimated effects and the width of the associated confidence intervals coupled with a more stringent p-value, rather than reliance on the "standard" significance cut-point of a p-value of 0.05 to provide evidence of an association. This is the approach that has been taken when considering the overall summary of the findings in relation to the key questions of the thesis.

8.3. Implications of findings

Reducing levels of obesity in the population has many benefits in terms of physical health [182, 183]. Findings presented in this thesis highlight the potential importance of reducing obesity in order to improve depressive symptoms in adolescent females. Perception of body image may play an important role in mediating this relationship and, if replicated in future studies, this may be a target for future intervention. It is therefore important that efforts to reduce levels of obesity continue to be part of a preventative strategy to improving adolescent health. This may have wider benefits in terms of improving the mental health of young females.

There is currently a large focus on physical activity for improving mental health in adolescents. The current government guidance is for children/adolescents to carry out at least one hour of MVPA every day, as a way of improving adolescent mental health and preventing depression [24]. This recommendation however is not currently backed up by scientific evidence, the National Institute for Health and Clinical Excellence (NICE) [184] rate the evidence for physical activity in the prevention and treatment of depression in adolescence as grade "C" which is defined as "expert committee reports or opinions and/or clinical experiences of respected authorities". In other words the guidance is based on clinical opinion rather than being supported by high quality scientific evidence.

The current study found no consistent evidence of an association between physical activity and depressive symptoms in adolescence. As such, the recommendations from this project would be that:

- physical activity continues to be recommended to adolescents, as physical activity is known to be important for physical health
- government policy recognises that further research is required to clarify the relationship between physical activity and mental health

- further research should have a particular focus on what aspect(s) of physical activity may be important (i.e. amount, frequency, intensity and/or context)

Moreover, it is important that the uncertainty surrounding the role of PA in terms of adolescent mental health is acknowledged and that other strategies to prevent adolescent depression should be explored.

8.4. Future work

To build on the current study, to address the limitations as discussed in sections 5.3, 6.3 and 7.2 and to further improve preventative strategies for depression in adolescence further work should be carried out into the relationship between obesity, physical activity and depression in adolescence.

Key to such future work is the collection of high quality data to be used in future analyses. Longitudinal repeated measures of obesity, physical activity, depression, potential mediators, important confounders, potential interaction and moderator variables and genetic data need to be collected on adolescents. As well as genetic data for use as instrumental variables future studies should also explore the use of other non-genetic instruments, for example policy changes such as the "sugar tax" that may impact obesity, policy interventions aimed at depression, or the impact of the Olympics on PA.

In order to ensure that potential interaction effects, mediation and MR analyses are sufficiently powered and that potentially small effect sizes could be investigated, data from a very large cohort would be required. It may be difficult however to find all of this data collected as part of an existing study and it may be difficult and costly to set up a new cohort study collecting all of this information. Therefore, the evidence base could be improved upon using an evidence synthesis approach. For

example, rather than trying to find or set up a large cohort that addresses all the potential questions, future work could focus on synthesising the evidence from existing studies that address certain aspects of the overall picture that we are trying to understand. Similarly, small focussed studies could be carried out into the different questions trying to be answered that each concentrate on the optimum measurement of the key factors related to the different questions. This could perhaps be achieved by identifying cohorts of younger children and establishing collaborations with the cohort research team to enable the inclusion of additional measures at appropriate ages in coming years. For example, participants in both the Born In Bradford [174] and Growing up in Wales [175] cohorts are currently still under 10 years old and hence there is the potential for future work in this area to be embedded within these cohorts.

The measures of obesity collected and investigated in future studies should be objective in nature; BMI from measured height and weight and direct measures of adiposity such as DXA fat percentage. The future obesity investigations should focus on adolescent females and investigate how best to conceptualize obesity as a risk factor for depression in adolescence, for example investigating a threshold effect, obesity as a chronic stressor and size of change in measure of obesity as exposure variables. Any further work should also investigate the impact of puberty and social context on the relationship between obesity and depression in adolescence. To help strengthen the evidence of a causal relationship between obesity and depression then further MR work should also be carried out, the first step should be a giant meta-analysis of all GWAS studies of obesity to identify the best possible genetic instrument for use in MR analyses. To help elucidate the causal pathway between obesity and depression in adolescence then further mediation analyses should also be carried out, with particular focus on body image. If possible identifying cohorts where the cultures may differ with regard to body shape preference would be an advantage. For example if in a certain cohort adolescent females feel negatively about being slim then the role of body image as a mediator could be strikingly

different to a cohort where being slim is seen as desirable. Taking advantage of these potential cross-cultural comparisons would be very useful in furthering our understanding of the relationship between obesity and depression in adolescence. For example cohorts from South America (e.g. Pelotas cohort [185, 186]), Africa (e.g. Birth to 20 cohort [187]), or Eastern Europe (e.g. Krakow cohort [188]) are likely to have different confounding structures to cohorts in Western Europe and North America.

The physical activity data collected in future studies should be both objective and self-report data; accelerometer data coupled with a structured activity diary so that participants can record the nature of the activity carried out and this can be matched to the accelerometer data on amount and intensity of activity. This would allow investigation into the importance of amount, frequency, intensity and context of physical activity in the association with adolescent depression. Detailed repeated measures data would also allow investigation into short versus long-term effects. Future studies should also account for the importance of sedentary behaviour using compositional analysis that would allow investigation into the effect and relative importance of (and changes in) different daily activity behaviours, and utilise techniques such as PLS-R to identify those aspects of PA (which is a broad construct) that are most related to depression.

8.5. Closing Remarks

Depression during adolescence is a major public health problem that confers significant burden to both individuals and the healthcare system. A preventative approach targeting modifiable risk factors of adolescent depression could improve the mental health of the nation and reduce these associated burdens. In this study obesity and physical activity were investigated as potentially modifiable risk factors

of depression in adolescence. There was some evidence that higher levels of obesity were related to increased depressive symptoms in females and that this relationship may be mediated by body image. However there was no robust evidence of an association between physical activity and depression.

Public health strategies to reduce obesity amongst adolescents may therefore not only have important benefits in terms of physical health, but may also improve the mental health of adolescent females. However, based on this study, there is little evidence that advice encouraging greater levels of physical activity confer beneficial effects in terms of adolescent mental health, but clearly such efforts will have wider health benefits that it is important to acknowledge. As outlined, much work is still to be done to better understand the complex relationship between obesity, physical activity and depression in order to better inform preventative strategies and improve mental health outcomes for future generations.

REFERENCES

1. World Health Organization., *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. 1992: WHO.
2. American Psychiatric Association., *Diagnostic and statistical manual of mental disorders: DSM-5*. 2013, Washington: American Psychiatric Association.
3. Pickles, A., et al., *Child psychiatric symptoms and psychosocial impairment: relationship and prognostic significance.* Br J Psychiatry, 2001. **179**: p. 230-5.
4. Mathers, C.D. and D. Loncar, *Projections of global mortality and burden of disease from 2002 to 2030.* PLoS Med, 2006. **3**(11): p. e442.
5. Schulz, R., et al., *Association between depression and mortality in older adults: the Cardiovascular Health Study.* Arch Intern Med, 2000. **160**(12): p. 1761-8.
6. Anderson, R.J., et al., *The prevalence of comorbid depression in adults with diabetes: a meta-analysis.* Diabetes Care, 2001. **24**(6): p. 1069-78.
7. Pohjasvaara, T., et al., *Depression is an independent predictor of poor long-term functional outcome post-stroke.* Eur J Neurol, 2001. **8**(4): p. 315-9.
8. Fergusson, D.M. and L.J. Woodward, *Mental health, educational, and social role outcomes of adolescents with depression.* Arch Gen Psychiatry, 2002. **59**(3): p. 225-31.
9. Kessler, R.C., S. Avenevoli, and K. Ries Merikangas, *Mood disorders in children and adolescents: an epidemiologic perspective.* Biol Psychiatry, 2001. **49**(12): p. 1002-14.
10. Collishaw, S., *Annual research review: Secular trends in child and adolescent mental health.*J Child Psychol Psychiatry, 2015. **56**(3): p 370-93.
11. Patton, G.C., et al., *Predicting female depression across puberty: a two-nation longitudinal study.* J Am Acad Child Adolesc Psychiatry, 2008. **47**(12): p. 1424-32.
12. Ogden, C.L., et al., *Trends in Obesity Prevalence Among Children and Adolescents in the United States, 1988-1994 Through 2013-2014.*JAMA, 2016. **315**(21): p. 2292-9.
13. Windfuhr, K., et al., *Suicide in juveniles and adolescents in the United Kingdom.* Journal of Child Psychology and Psychiatry, 2008. **49**(11): p. 1155-1165.
14. Keenan-Miller, D., C.L. Hammen, and P.A. Brennan, *Health outcomes related to early adolescent depression.* Journal of Adolescent Health, 2007. **41**(3): p. 256-262.
15. Fletcher, J.M., *Adolescent Depression: Diagnosis, Treatment, and Educational Attainment.* Health Economics, 2008. **17**(11): p. 1215-1235.
16. Lewinsohn, P.M., et al., *Natural course of adolescent major depressive disorder: I. Continuity into young adulthood.* J Am Acad Child Adolesc Psychiatry, 1999. **38**(1): p. 56-63.
17. Fombonne, E., et al., *The Maudsley long-term follow-up of child and adolescent depression. 1. Psychiatric outcomes in adulthood.* Br J Psychiatry, 2001. **179**: p. 210-7.
18. McCrone, P.D., S. Patel, A. Knapp, M. Lawton-Smith, *Paying the price: The cost of mental health care in England to 2026*. 2008, King's Fund.
19. Forsight Mental Capital and Wellbeing Project., *Final Project Report*. 2008, The Government Office for Science: London.
20. Department of Health., *Obesity and healthy eating*. 2014.
21. World Health Organisation. *Obesity and Overweight*. 2016 [cited 2017 8/8/17].
22. Reilly, J.J. and D. Wilson, *ABC of obesity - Childhood obesity.* British Medical Journal, 2006. **333**(7580): p. 1207-1210.
23. Guh, D.P., et al., *The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis.* BMC Public Health, 2009. **9**: p. 88.

24. Bull, F.a.t.E.W.G., *Physical Activity Guidelines in the U.K.: Review and Recommendations*, D.o. Health, Editor. 2010.
25. Twisk, J.W.R., *Physical activity guidelines for children and adolescents - A critical review.* Sports Medicine, 2001. **31**(8): p. 617-627.
26. Department of Health., *National service framework for mental health: modern standards and service models*. 1999, Department of Health: London.
27. Department of Health., *Choosing health: making healthy choices easier*. 2004, Department of Health: London.
28. National Institute for Health and ClinicalExcellence., *Promoting young people's social and emotional well being in secondary education*. 2009, National Institute for Health and Clinical Excellence: London.
29. Reilly, J.J. and D. Wilson, *ABC of obesity. Childhood obesity.* BMJ, 2006. **333**(7580): p. 1207-10.
30. Wardle, J. and L. Cooke, *The impact of obesity on psychological well-being.* Best Pract Res Clin Endocrinol Metab, 2005. **19**(3): p. 421-40.
31. Erermis, S., et al., *Is obesity a risk factor for psychopathology among adolescents?* Pediatr Int, 2004. **46**(3): p. 296-301.
32. Erickson, S.J., et al., *Are overweight children unhappy?: Body mass index, depressive symptoms, and overweight concerns in elementary school children.* Arch Pediatr Adolesc Med, 2000. **154**(9): p. 931-5.
33. Falkner, N.H., et al., *Social, educational, and psychological correlates of weight status in adolescents.* Obesity Research, 2001. **9**(1): p. 32-42.
34. Lamertz, C.M., et al., *Are obese adolescents and young adults at higher risk for mental disorders? A community survey.* Obesity Research, 2002. **10**(11): p. 1152-1160.
35. Eisenberg, M.E., D. Neumark-Sztainer, and M. Story, *Associations of weight-based teasing and emotional well-being among adolescents.* Arch Pediatr Adolesc Med, 2003. **157**(8): p. 733-8.
36. Brewis, A., *Biocultural aspects of obesity in young Mexican schoolchildren.* Am J Hum Biol, 2003. **15**(3): p. 446-60.
37. Herva, A., et al., *Obesity and depression: results from the longitudinal Northern Finland 1966 Birth Cohort Study.* International Journal of Obesity, 2006. **30**(3): p. 520-527.
38. Boutelle, K.N., et al., *Obesity as a Prospective Predictor of Depression in Adolescent Females.* Health Psychology, 2010. **29**(3): p. 293-298.
39. Luppino, F.S., et al., *Overweight, obesity, and depression: a systematic review and meta-analysis of longitudinal studies.* Arch Gen Psychiatry, 2010. **67**(3): p. 220-9.
40. Korczak, D.J., et al., *Are children and adolescents with psychiatric illness at risk for increased future body weight? A systematic review.* Dev Med Child Neurol, 2013. **55**(11): p. 980-7.
41. Hoare, E., et al., *Associations between obesogenic risk factors and depression among adolescents: a systematic review.* Obes Rev, 2014. **15**(1): p. 40-51.
42. Mustillo, S., et al., *Obesity and psychiatric disorder: developmental trajectories.* Pediatrics, 2003. **111**(4 Pt 1): p. 851-9.
43. Anderson, S.E., et al., *Adolescent obesity and risk for subsequent major depressive disorder and anxiety disorder: Prospective evidence.* Psychosom Med, 2007. **69**(8): p. 740-747.
44. Anderson, S.E., et al., *Obesity and depressed mood associations differ by race/ethnicity in adolescent girls.* Int J Pediatr Obes, 2011. **6**(1): p. 69-78.
45. Jokela, M., et al., *Body mass index and depressive symptoms: instrumental-variables regression with genetic risk score.* Genes Brain Behav, 2012.
46. Kivimaki, M., et al., *Examining overweight and obesity as risk factors for common mental disorders using fat mass and obesity-associated (FTO) genotype-instrumented analysis: The Whitehall II Study, 1985-2004.* Am J Epidemiol, 2011. **173**(4): p. 421-9.

47. Lawlor, D.A., et al., *Using genetic loci to understand the relationship between adiposity and psychological distress: a Mendelian Randomization study in the Copenhagen General Population Study of 53,221 adults.* J Intern Med, 2011. **269**(5): p. 525-37.
48. Samaan, Z., et al., *The protective effect of the obesity-associated rs9939609 A variant in fat mass- and obesity-associated gene on depression.* Mol Psychiatry, 2012.
49. Goodman, E. and R.C. Whitaker, *A prospective study of the role of depression in the development and persistence of adolescent obesity.* Pediatrics, 2002. **110**(3): p. 497-504.
50. Roberts, R.E. and H.T. Duong, *Obese youths are not more likely to become depressed, but depressed youths are more likely to become obese.* Psychol Med, 2013. **43**(10): p. 2143-51.
51. Pine, D.S., et al., *Psychiatric symptoms in adolescence as predictors of obesity in early adulthood: a longitudinal study.* Am J Public Health, 1997. **87**(8): p. 1303-10.
52. Anderson, S.E., et al., *Association of depression and anxiety disorders with weight change in a prospective community-based study of children followed up into adulthood.* Arch Pediatr Adolesc Med, 2006. **160**(3): p. 285-91.
53. Elgar, F.J. and J.M. Stewart, *Validity of self-report screening for overweight and obesity. Evidence from the Canadian Community Health Survey.* Can J Public Health, 2008. **99**(5): p. 423-7.
54. Maynard, L.M., et al., *Childhood body composition in relation to body mass index.* Pediatrics, 2001. **107**(2): p. 344-350.
55. Freedman, D.S., et al., *Relation of BMI to fat and fat-free mass among children and adolescents.* Int J Obes (Lond), 2005. **29**(1): p. 1-8.
56. Lawlor, D.A., et al., *Association between general and central adiposity in childhood, and change in these, with cardiovascular risk factors in adolescence: prospective cohort study.* BMJ, 2010. **341**: p. c6224.
57. Needham, B.L. and R. Crosnoe, *Overweight status and depressive symptoms during adolescence.* J Adolesc Health, 2005. **36**(1): p. 48-55.
58. Herva, A., et al., *Obesity and depression: results from the longitudinal Northern Finland 1966 Birth Cohort Study.* Int J Obes (Lond), 2006. **30**(3): p. 520-7.
59. Lamertz, C.M., et al., *Are obese adolescents and young adults at higher risk for mental disorders? A community survey.* Obes Res, 2002. **10**(11): p. 1152-60.
60. Smith, G.D. and S. Ebrahim, *'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease?* Int J Epidemiol, 2003. **32**(1): p. 1-22.
61. Lawlor, D.A., et al., *Mendelian randomization: using genes as instruments for making causal inferences in epidemiology.* Stat Med, 2008. **27**(8): p. 1133-63.
62. Palmer, T.M., et al., *Using multiple genetic variants as instrumental variables for modifiable risk factors.* Stat Methods Med Res, 2012. **21**(3): p. 223-42.
63. Xu, H., et al., *Chronic inflammation in fat plays a crucial role in the development of obesity-related insulin resistance.* J Clin Invest, 2003. **112**(12): p. 1821-30.
64. Maury, E. and S.M. Brichard, *Adipokine dysregulation, adipose tissue inflammation and metabolic syndrome.* Mol Cell Endocrinol, 2010. **314**(1): p. 1-16.
65. Berg, A.H. and P.E. Scherer, *Adipose tissue, inflammation, and cardiovascular disease.* Circ Res, 2005. **96**(9): p. 939-49.
66. Capuron, L. and A.H. Miller, *Immune system to brain signaling: neuropsychopharmacological implications.* Pharmacol Ther, 2011. **130**(2): p. 226-38.
67. Howren, M.B., D.M. Lamkin, and J. Suls, *Associations of depression with C-reactive protein, IL-1, and IL-6: a meta-analysis.* Psychosom Med, 2009. **71**(2): p. 171-86.
68. Dowlati, Y., et al., *A meta-analysis of cytokines in major depression.* Biol Psychiatry, 2010. **67**(5): p. 446-57.
69. Pasquali, R. and V. Vicennati, *Activity of the hypothalamic-pituitary-adrenal axis in different obesity phenotypes.* Int J Obes Relat Metab Disord, 2000. **24 Suppl 2**: p. S47-9.

70. Walker, B.R., *Activation of the hypothalamic-pituitary-adrenal axis in obesity: cause or consequence?* Growth Horm IGF Res, 2001. **11 Suppl A**: p. S91-5.
71. Holsboer, F., *The corticosteroid receptor hypothesis of depression.* Neuropsychopharmacology, 2000. **23**(5): p. 477-501.
72. Nemiary, D., et al., *The Relationship Between Obesity and Depression Among Adolescents.* Psychiatr Ann, 2012. **42**(8): p. 305-308.
73. Friedman, K.E., et al., *Body image partially mediates the relationship between obesity and psychological distress.* Obesity Research, 2002. **10**(1): p. 33-41.
74. Daly, M., *The relationship of C-reactive protein to obesity-related depressive symptoms: a longitudinal study.* Obesity (Silver Spring), 2013. **21**(2): p. 248-50.
75. Peterson, R.E., et al., *Binge Eating Disorder Mediates Links between Symptoms of Depression, Anxiety, and Caloric Intake in Overweight and Obese Women.* J Obes, 2012. **2012**: p. 407103.
76. Konttinen, H., et al., *Emotional eating, depressive symptoms and self-reported food consumption. A population-based study.* Appetite, 2010. **54**(3): p. 473-479.
77. Gavin, A.R., G.E. Simon, and E.J. Ludman, *The association between obesity, depression, and educational attainment in women: the mediating role of body image dissatisfaction.* J Psychosom Res, 2010. **69**(6): p. 573-81.
78. Friedman, K.E., et al., *Body image partially mediates the relationship between obesity and psychological distress.* Obes Res, 2002. **10**(1): p. 33-41.
79. Mond, J., et al., *Obesity, Body Dissatisfaction, and Emotional Well-Being in Early and Late Adolescence: Findings From the Project EAT Study.* Journal of Adolescent Health, 2011. **48**(4): p. 373-378.
80. Dockray, S., E.J. Susman, and L.D. Dorn, *Depression, Cortisol Reactivity, and Obesity in Childhood and Adolescence.* Journal of Adolescent Health, 2009. **45**(4): p. 344-350.
81. Roberts, R.E. and H.T. Duong, *Does major depression affect risk for adolescent obesity?* J Affect Disord, 2015. **186**: p. 162-7.
82. Mond, J., et al., *Obesity, body dissatisfaction, and emotional well-being in early and late adolescence: findings from the project EAT study.* J Adolesc Health, 2011. **48**(4): p. 373-8.
83. Dockray, S., E.J. Susman, and L.D. Dorn, *Depression, cortisol reactivity, and obesity in childhood and adolescence.* J Adolesc Health, 2009. **45**(4): p. 344-50.
84. Hayes, A.F., *Beyond Baron and Kenny: Statistical Mediation Analysis in the New Millennium.* Communication Monographs, 2009. **76**(4): p. 408-420.
85. Twisk, J.W., *Physical activity guidelines for children and adolescents: a critical review.* Sports Medicine, 2001. **31**(8): p. 617-27.
86. Fredricks, J.A. and J.S. Eccles, *Is extracurricular participation associated with beneficial outcomes? Concurrent and longitudinal relations.* Dev Psychol, 2006. **42**(4): p. 698-713.
87. Gore, S., Farrel, F., Gordon, J., *Sports involvement as protection against depressed mood.* Journal of Research on Adolescence, 2001(11): p. 119-13-.
88. Jerstad, S.J., et al., *Prospective Reciprocal Relations Between Physical Activity and Depression in Female Adolescents.* Journal of Consulting and Clinical Psychology, 2010. **78**(2): p. 268-272.
89. Motl, R.W., et al., *Naturally occurring changes in physical activity are inversely related to depressive symptoms during early adolescence.* Psychosom Med, 2004. **66**(3): p. 336-42.
90. Rothon, C., et al., *Physical activity and depressive symptoms in adolescents: a prospective study.* BMC Med, 2010. **8**: p. 32.
91. Sund, A.M., B. Larsson, and L. Wichstrom, *Role of physical and sedentary activities in the development of depressive symptoms in early adolescence.* Soc Psychiatry Psychiatr Epidemiol, 2011. **46**(5): p. 431-41.

92. Stavrakakis, N., et al., *Bidirectional Prospective Associations Between Physical Activity and Depressive Symptoms. The TRAILS Study.* Journal of Adolescent Health, 2012. **50**(5): p. 503-508.
93. Sallis, J.F., *Self-Report Measures of Childrens Physical-Activity.* Journal of School Health, 1991. **61**(5): p. 215-219.
94. Bailey, R.C., et al., *The Level and Tempo of Childrens Physical Activities - an Observational Study.* Medicine and Science in Sports and Exercise, 1995. **27**(7): p. 1033-1041.
95. Lewinsohn, P.M., et al., *Adolescent psychopathology: II. Psychosocial risk factors for depression.* J Abnorm Psychol, 1994. **103**(2): p. 302-15.
96. Birkeland, M.S., T. Torsheim, and B. Wold, *A longitudinal study of the relationship between leisure-time physical activity and depressed mood among adolescents.* Psychology of Sport and Exercise, 2009. **10**(1): p. 25-34.
97. Steptoe, A. and N. Butler, *Sports participation and emotional wellbeing in adolescents.* Lancet, 1996. **347**(9018): p. 1789-92.
98. Haarasilta, L.M., et al., *Correlates of depression in a representative nationwide sample of adolescents (15-19 years) and young adults (20-24 years).* Eur J Public Health, 2004. **14**(3): p. 280-5.
99. Allison, K.R., et al., *Relationship of vigorous physical activity to psychologic distress among adolescents.* J Adolesc Health, 2005. **37**(2): p. 164-6.
100. De Moor, M.H., et al., *Genome-wide association study of exercise behavior in Dutch and American adults.* Med Sci Sports Exerc, 2009. **41**(10): p. 1887-95.
101. Kim, J., et al., *Practical issues in genome-wide association studies for physical activity.* Ann N Y Acad Sci, 2011. **1229**: p. 38-44.
102. Zoladz, J.A. and A. Pilc, *The effect of physical activity on the brain derived neurotrophic factor: from animal to human studies.* J Physiol Pharmacol, 2010. **61**(5): p. 533-41.
103. Puterman, E., et al., *Physical activity moderates effects of stressor-induced rumination on cortisol reactivity.* Psychosom Med, 2011. **73**(7): p. 604-11.
104. Fox, K.R., *The influence of physical activity on mental well-being.* Public Health Nutr, 1999. **2**(3A): p. 411-8.
105. Camacho, T.C., et al., *Physical activity and depression: evidence from the Alameda County Study.* Am J Epidemiol, 1991. **134**(2): p. 220-31.
106. Golding, J., *Children of the nineties. A longitudinal study of pregnancy and childhood based on the population of Avon (ALSPAC).* West Engl Med J, 1990. **105**(3): p. 80-2.
107. Golding, J., et al., *ALSPAC--the Avon Longitudinal Study of Parents and Children. I. Study methodology.* Paediatr Perinat Epidemiol, 2001. **15**(1): p. 74-87.
108. Huisman, M., et al., *Cohort profile: the Dutch 'TRacking Adolescents' Individual Lives' Survey'; TRAILS.* Int J Epidemiol, 2008. **37**(6): p. 1227-35.
109. Oldehinkel, A.J., et al., *Cohort Profile Update: the TRacking Adolescents' Individual Lives Survey (TRAILS).* Int J Epidemiol, 2015. **44**(1): p. 76-76n.
110. O'Loughlin, J., et al., *Cohort Profile: The Nicotine Dependence in Teens (NDIT) Study.* Int J Epidemiol, 2015. **44**(5): p. 1537-46.
111. Turner, N., et al., *Validity of the Short Mood and Feelings Questionnaire in late adolescence.* Psychol Assess, 2014. **26**(3): p. 752-62.
112. Goodman, R., et al., *Rating child psychiatric caseness from detailed case histories.* J Child Psychol Psychiatry, 1996. **37**(4): p. 369-79.
113. Cole, D.A., et al., *Structural differences in parent and child reports of children's symptoms of depression and anxiety.* Psychol Assess, 2000. **12**(2): p. 174-85.
114. Hunt, M., J. Auriemma, and A.C. Cashaw, *Self-report bias and underreporting of depression on the BDI-II.* J Pers Assess, 2003. **80**(1): p. 26-30.

115. Angold, A., Costello, E. J., Messer, S. C., Pickles, A., Winder, F., & Silver, D, *The development of a short questionnaire for use in epidemiological studies of depression in children and adolescents.* International Journal of Methods in Psychiatric Research, 1995. **5**: p. 237 - 249.
116. Achenbach, TM., *Manual for the ASEBA School-Age Forms & Profiles*. 2001, Burlingtion, VT: University of Vermont.
117. Kandel, D.B. and M. Davies, *Epidemiology of depressive mood in adolescents: an empirical study.* Arch Gen Psychiatry, 1982. **39**(10): p. 1205-12.
118. Elgar, F.J. and J.M. Stewart, *Validity of Self-report Screening for Overweight and Obesity Evidence from the Canadian Community Health Survey.* Canadian Journal of Public Health-Revue Canadienne De Sante Publique, 2008. **99**(5): p. 423-427.
119. Freedman, D.S., et al., *Relation of BMI to fat and fat-free mass among children and adolescents.* International Journal of Obesity, 2005. **29**(1): p. 1-8.
120. Brener, N.D., et al., *Reliability and validity of self-reported height and weight among high school students.* J Adolesc Health, 2003. **32**(4): p. 281-7.
121. Elgar, F.J., et al., *Validity of self-reported height and weight and predictors of bias in adolescents.* J Adolesc Health, 2005. **37**(5): p. 371-5.
122. Goodman, E., B.R. Hinden, and S. Khandelwal, *Accuracy of teen and parental reports of obesity and body mass index.* Pediatrics, 2000. **106**(1 Pt 1): p. 52-8.
123. Troiano, R.P. and K.M. Flegal, *Overweight prevalence among youth in the United States: why so many different numbers?* Int J Obes Relat Metab Disord, 1999. **23 Suppl 2**: p. S22-7.
124. John, U., et al., *Validity of overweight and obesity in a nation based on self-report versus measurement device data.* Eur J Clin Nutr, 2006. **60**(3): p. 372-7.
125. Rothman, K.J., *BMI-related errors in the measurement of obesity.* Int J Obes (Lond), 2008. **32 Suppl 3**: p. S56-9.
126. Ellis, K.J., *Human body composition: in vivo methods.* Physiol Rev, 2000. **80**(2): p. 649-80.
127. Addo, O.Y. and J.H. Himes, *Reference curves for triceps and subscapular skinfold thicknesses in US children and adolescents.* Am J Clin Nutr, 2010. **91**(3): p. 635-42.
128. Locke, A.E., et al., *Genetic studies of body mass index yield new insights for obesity biology.* Nature, 2015. **518**(7538): p. 197-206.
129. Burgess, S. and S.G. Thompson, *Use of allele scores as instrumental variables for Mendelian randomization.* Int J Epidemiol, 2013. **42**(4): p. 1134-44.
130. Mattocks, C., et al., *Calibration of an accelerometer during free-living activities in children.* Int J Pediatr Obes, 2007. **2**(4): p. 218-26.
131. Mattocks, C., et al., *Early life determinants of physical activity in 11 to 12 year olds: cohort study.* BMJ, 2008. **336**(7634): p. 26-9.
132. Mattocks, C., et al., *Use of accelerometers in a large field-based study of children: protocols, design issues, and effects on precision.* J Phys Act Health, 2008. **5 Suppl 1**: p. S98-111.
133. Sallis, J.F., et al., *The development of self-administered physical activity surveys for 4th grade students.* Res Q Exerc Sport, 1993. **64**(1): p. 25-31.
134. Barnett, T.A., et al., *Physical activity growth curves relate to adiposity in adolescents.* Ann Epidemiol, 2013. **23**(9): p. 529-33.
135. Booij, S.H., et al., *Chronicity of depressive problems and the cortisol response to psychosocial stress in adolescents: the TRAILS study.* Psychoneuroendocrinology, 2013. **38**(5): p. 659-66.
136. Carnegie, R., et al., *Cortisol awakening response and subsequent depression: prospective longitudinal study.* Br J Psychiatry, 2014. **204**(2): p. 137-43.
137. Khandaker, G.M., et al., *Association of serum interleukin 6 and C-reactive protein in childhood with depression and psychosis in young adult life: a population-based longitudinal study.* JAMA Psychiatry, 2014. **71**(10): p. 1121-8.
138. Khandaker, G.M., et al., *Association between serum C-reactive protein and DSM-IV generalized anxiety disorder in adolescence: Findings from the ALSPAC cohort.* Neurobiol Stress, 2016. **4**: p. 55-61.

139. Jonker, I., J.G.M. Rosmalen, and R.A. Schoevers, *Childhood life events, immune activation and the development of mood and anxiety disorders: the TRAILS study.* Transl Psychiatry, 2017. **7**(5): p. e1112.
140. Cox, J.L., J.M. Holden, and R. Sagovsky, *Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale.* Br J Psychiatry, 1987. **150**: p. 782-6.
141. Lovibond, P.F. and S.H. Lovibond, *The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories.* Behav Res Ther, 1995. **33**(3): p. 335-43.
142. Joinson, C., et al., *Early menarche and depressive symptoms from adolescence to young adulthood in a UK cohort.* J Am Acad Child Adolesc Psychiatry, 2013. **52**(6): p. 591-8 e2.
143. Joinson, C., et al., *Association between pubertal development and depressive symptoms in girls from a UK cohort.* Psychol Med, 2012. **42**(12): p. 2579-89.
144. Brookes, S.T., et al., *Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test.* J Clin Epidemiol, 2004. **57**(3): p. 229-36.
145. Hanley, J.A., et al., *Statistical analysis of correlated data using generalized estimating equations: an orientation.* Am J Epidemiol, 2003. **157**(4): p. 364-75.
146. Akaike, H., *A new look at the statistical model identification.* IEEE Transactions on Automatic Control, 1974. **19**: p. 716-723.
147. Pan, W., *Akaike's information criterion in generalized estimating equations.* Biometrics, 2001. **57**(1): p. 120-5.
148. White, H., *A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity.* Econometrica, 1980. **48**(4): p. 817-838.
149. StataCorp, *2015*, in *Stata Statistical Software: Release 14*. StataCorp LP: College Station, TX.
150. Little, T.D., Cunningham, W.A., Shahar, G., Widaman, K.F., *To Parcel or Not to Parcel: Exploring the Question, Weighing the Merits.* Structural Equation Modelling, 2002. **9**(2): p. 151 - 173.
151. Little, T.D., Preacher, K.J., Card, N.A, *New developments in latent variable panel analyses of longitudinal data.* International Journal of Behavioral Development, 2007. **31**(4): p. 357 - 365.
152. Davis, W.R., *The FC1 rule of identification for confirmatory factor analysis: A general sufficient condition.* Sociological Methods & Research, 1993. **21**(4): p. 403 - 437.
153. Reilly, T., *A necessary and sufficient condition for identification of confirmatory factor analysis models of complexity one.* Sociological Methods & Research, 1995. **23**(4): p. 421 - 441.
154. Burgess, S., S.G. Thompson, and C.C.G. Collaboration, *Avoiding bias from weak instruments in Mendelian randomization studies.* Int J Epidemiol, 2011. **40**(3): p. 755-64.
155. Pierce, B.L., H. Ahsan, and T.J. Vanderweele, *Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants.* Int J Epidemiol, 2011. **40**(3): p. 740-52.
156. Bowden, J., G. Davey Smith, and S. Burgess, *Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression.* Int J Epidemiol, 2015. **44**(2): p. 512-25.
157. Abdi, H. and L.J. Williams, *Partial least squares methods: partial least squares correlation and partial least square regression.* Methods Mol Biol, 2013. **930**: p. 549-79.
158. Evans, D.M., P.M. Visscher, and N.R. Wray, *Harnessing the information contained within genome-wide association studies to improve individual prediction of complex disease risk.* Hum Mol Genet, 2009. **18**(18): p. 3525-31.
159. Richmond, R.C., et al., *Assessing causality in the association between child adiposity and physical activity levels: a Mendelian randomization analysis.* PLoS Med, 2014. **11**(3): p. e1001618.

160. Pinhey, T.K., Rubinstein, D.H., Colfax, R.S., *Overweight and happiness: the reflected self-appraisal hypothesis reconsidered.* Social Science Quaterly, 1997. **78**(3): p. 747-755.
161. Needham, B.L. and R. Crosnoe, *Overweight status and depressive symptoms during adolescence.* Journal of Adolescent Health, 2005. **36**(1): p. 48-55.
162. de Wit, L.M., et al., *Depression and body mass index, a u-shaped association.* BMC Public Health, 2009. **9**: p. 14.
163. Southwick, S.M., M. Vythilingam, and D.S. Charney, *The psychobiology of depression and resilience to stress: implications for prevention and treatment.* Annu Rev Clin Psychol, 2005. **1**: p. 255-91.
164. Skrove, M., P. Romundstad, and M.S. Indredavik, *Resilience, lifestyle and symptoms of anxiety and depression in adolescence: the Young-HUNT study.* Soc Psychiatry Psychiatr Epidemiol, 2013. **48**(3): p. 407-16.
165. Siervogel, R.M., et al., *Puberty and body composition.* Horm Res, 2003. **60**(Suppl 1): p. 36-45.
166. Loomba-Albrecht, L.A. and D.M. Styne, *Effect of puberty on body composition.* Curr Opin Endocrinol Diabetes Obes, 2009. **16**(1): p. 10-5.
167. Ebrahim, S. and G. Davey Smith, *Mendelian randomization: can genetic epidemiology help redress the failures of observational epidemiology?* Hum Genet, 2008. **123**(1): p. 15-33.
168. Boutelle, K.N., et al., *Obesity as a prospective predictor of depression in adolescent females.* Health Psychol, 2010. **29**(3): p. 293-8.
169. Ainsworth, B.E., et al., *Compendium of physical activities: an update of activity codes and MET intensities.* Med Sci Sports Exerc, 2000. **32**(9 Suppl): p. S498-504.
170. Owen, N., et al., *Environmental determinants of physical activity and sedentary behavior.* Exerc Sport Sci Rev, 2000. **28**(4): p. 153-8.
171. Clark, B.K., et al., *Validity and reliability of measures of television viewing time and other non-occupational sedentary behaviour of adults: a review.* Obes Rev, 2009. **10**(1): p. 7-16.
172. Teychenne, M., K. Ball, and J. Salmon, *Sedentary behavior and depression among adults: a review.* Int J Behav Med, 2010. **17**(4): p. 246-54.
173. Dumuid, D., et al., *Compositional data analysis for physical activity, sedentary time and sleep research.* Stat Methods Med Res, 2017: p. 962280217710835.
174. Wright, J., et al., *Cohort Profile: the Born in Bradford multi-ethnic family cohort study.* Int J Epidemiol, 2013. **42**(4): p. 978-91.
175. Morgan, K.L., et al., *Cohort Profile: Growing Up in Wales: The Environments for Healthy Living study.* Int J Epidemiol, 2016. **45**(2): p. 364-73.
176. Voelker, D.K., J.J. Reel, and C. Greenleaf, *Weight status and body image perceptions in adolescents: current perspectives.* Adolesc Health Med Ther, 2015. **6**: p. 149-58.
177. VanderWeele, T.J., *A unification of mediation and interaction: a 4-way decomposition.* Epidemiology, 2014. **25**(5): p. 749-61.
178. Ferrari, E.P., et al., *Body image dissatisfaction and anthropometric indicators in male children and adolescents.* Eur J Clin Nutr, 2015. **69**(10): p. 1140-4.
179. Burgess, S., et al., *Network Mendelian randomization: using genetic variants as instrumental variables to investigate mediation in causal pathways.* Int J Epidemiol, 2015. **44**(2): p. 484-95.
180. Munafo, M., R., Tilling, K., Taylor, A., E., Evans, D., M., Davey Smith, G.,, *Collider Scope: When selection bias can substantially influence observed associations.* International Journal of Epidemiology, 2017.
181. Fry, A., et al., *Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants with the General Population.* Am J Epidemiol, 2017.
182. Goldstein, D.J., *Beneficial health effects of modest weight loss.* Int J Obes Relat Metab Disord, 1992. **16**(6): p. 397-415.
183. Vidal, J., *Updated review on the benefits of weight loss.* Int J Obes Relat Metab Disord, 2002. **26 Suppl 4**: p. S25-8.

184. National Institute for Health and Clinical Excellence., *Depression in children and young people: identification and management*. 2015, NICE: London.
185. Santos, I.S., et al., *Cohort profile update: 2004 Pelotas (Brazil) Birth Cohort Study. Body composition, mental health and genetic assessment at the 6 years follow-up.* Int J Epidemiol, 2014. **43**(5): p. 1437-1437a-f.
186. Horta, B.L., et al., *Cohort Profile Update: The 1982 Pelotas (Brazil) Birth Cohort Study.* Int J Epidemiol, 2015. **44**(2): p. 441, 441a-441e.
187. Richter, L., et al., *Cohort Profile: Mandela's children: the 1990 Birth to Twenty study in South Africa.* Int J Epidemiol, 2007. **36**(3): p. 504-11.
188. Jedrychowski, W., E. Flak, and E. Mroz, *The adverse effect of low levels of ambient air pollutants on lung function growth in preadolescent children.* Environ Health Perspect, 1999. **107**(8): p. 669-74.

APPENDICES

Appendix 1: List of Abbreviations

2SLS	Two Stage Least Squares
AIC	Akaike Information Criterion
AIDs	Acquired Immune Deficiency syndrome
ALSPAC	Avon Longitudinal Study of Parents and Children
APS	Affective Problems Scale
BMI	Body Mass Index
CI	Confidence Interval
CPM	Counts Per Minute
CRP	C-Reactive Protein
DASS	Depression Anxiety Stress Scales
DSM	Diagnostic and Statistical Manual of mental disorders
DXA	Dual-energy X-ray Absorptiometry
EPDS	Edinburgh Postnatal Depression Scale
GEE	Generalized Estimating Equations
GLM	Generalised Linear Model
GWAS	Genome Wide Association Study
HPA	Hypothalamic-Pituitary-Adrenal
HIV	Human Immunodeficiency Virus
ICD	International Classification of Diseases
IgE	Immunoglobulin E
IQR	Interquartile Range
IV	Instrumental Variable
KDSS	Kandel Depressive Symptom Score
MDD	Major Depressive Disorder
MFQ	Mood and Feelings Questionnaire
MR	Mendelian Randomization
MVPA	Moderate to Vigorous Physical Activity
NDIT	Nicotine Dependence in Teens
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
NMR	Network Mendelian Randomization
OLS	Ordinary Least Squares
OR	Odds Ratio
PA	Physical Activity
PCA	Principal Components Analysis
PLS-R	Partial Least Squares Regression

QIC	Quasi-Information Criterion
RMSEP	Root Mean Squared Error of Prediction
SD	Standard Deviation
SE	Standard Error
SEM	Structural Equation Modelling
SEP	Socio-Economic Position
SES	Socio-Economic-Status
SMFQ	Short Moods and Feelings Questionnaire
SNP	Single Nucleotide Polymorphism
TRAILS	Tracking Adolescents' Individual Lives Survey
YSR	Youth Self Report

Appendix 2: Impact of alcohol on the investigation into the relationship between obesity and depression in adolescence in the ALSPAC cohort

Table A2.1 Results of linear regression analysis investigating the association between BMI (exposure) and depression (outcome) in the ALSPAC cohort, including alcohol use as a confounder

Timepoint	Exposure	Outcome	Model 1 - main model				Model 2 - Including interaction (coefficient for interaction effect)			
			n	coeff.	95%CI	p-value	n	coeff[#].	95%CI	p-value
F10 to TF1 10y8m to 12y10m	BMI	Depression	4161	0.017	(0.008, 0.027)	<0.001	4161	-0.001	(-0.020, 0.018)	0.903
TF1 to TF2 12y10m to 13y10m	BMI	Depression	798	-0.008	(-0.025, 0.009)	0.380	798	0.011	(-0.023, 0.045)	0.536
TF2 to CCS 13y10m to 16y8m	BMI	Depression	2820	0.017	(0.006, 0.028)	0.002	2820	0.026	(0.004, 0.047)	0.019
TF4 to CCT 17y10m to 18y8m	BMI	Depression	1444	0.004	(-0.008, 0.016)	0.467	1444	0.024	(0.001, 0.048)	0.045

Model 1 is adjusted for age (at outcome), sex, alcohol, previous depression, maternal depression, maternal education and maternal profession

*Model 2 is Model 1 plus the inclusion of a BMI*Sex interaction term*

Timepoint	Exposure	Outcome	Model 3 - stratified (females)				Model 4 - stratified (males)			
			n	coeff.	95%CI	p-value	n	coeff.	95%CI	p-value
F10 to TF1 10y8m to 12y10m	BMI	Depression	2137	0.018	(0.004, 0.031)	0.010	2024	0.017	(0.004, 0.031)	0.011
TF1 to TF2 12y10m to 13y10m	BMI	Depression	408	-0.003	(-0.029, 0.022)	0.795	390	-0.012	(-0.035, 0.010)	0.282
TF2 to CCS 13y10m to 16y8m	BMI	Depression	1596	0.026	(0.011, 0.041)	0.001	1224	0.003	(-0.012, 0.018)	0.716
TF4 to CCT 17y10m to 18y8m	BMI	Depression	914	0.012	(-0.003, 0.027)	0.118	530	-0.014	(-0.033, 0.005)	0.139

Model 3 is Model 1 stratified by gender - analysing females only and without including sex as a confounder

Model 4 is Model 1 stratified by gender - analysing males only and without including sex as a confounder

Appendix 3: The impact of smoking on the relationship between obesity and depression in adolescence in the ALSPAC cohort

Table A3.1 Results of linear regression analysis investigating the association between BMI (exposure) and depression (outcome) in the ALSPAC cohort, including smoking as a confounder

Timepoint	Exposure	Outcome	Model 1 - main model				Model 2 - Including interaction (coefficient for interaction effect)			
			n	coeff.	95%CI	p-value	n	coeff[#].	95%CI	p-value
F10 to TF1 10y8m to 12y10m	BMI	Depression	4186	0.016	(0.007, 0.026)	0.001	4186	0.001	(-0.018, 0.019)	0.932
TF1 to TF2 12y10m to 13y10m	BMI	Depression	3572	0.009	(0.0001, 0.018)	0.046	3572	0.020	(0.002, 0.038)	0.027
TF2 to CCS 13y10m to 16y8m	BMI	Depression	2852	0.017	(0.006, 0.028)	0.003	2852	0.025	(0.004, 0.047)	0.021
TF4 to CCT 17y10m to 18y8m	BMI	Depression	680	0.008	(-0.009, 0.026)	0.341	680	0.043	(0.008, 0.078)	0.016

Model 1 is adjusted for age (at outcome), sex, smoking, previous depression, maternal depression, maternal education and maternal profession

*Model 2 is Model 1 plus the inclusion of a BMI*Sex interaction term*

[#] *Results presented are for the BMI* Sex interaction term*

Timepoint	Exposure	Outcome	Model 3 - stratified (females)				Model 4 - stratified (males)			
			n	coeff.	95%CI	p-value	n	coeff.	95%CI	p-value
F10 to TF1 10y8m to 12y10m	BMI	Depression	2139	0.018	(0.004, 0.031)	0.010	2047	0.015	(0.002, 0.028)	0.023
TF1 to TF2 12y10m to 13y10m	BMI	Depression	1795	0.019	(0.006, 0.032)	0.005	1777	-0.001	(-0.013, 0.011)	0.887
TF2 to CCS 13y10m to 16y8m	BMI	Depression	1613	0.025	(0.010, 0.040)	0.001	1239	0.004	(-0.012, 0.019)	0.624
TF4 to CCT 17y10m to 18y8m	BMI	Depression	467	0.020	(-0.001, 0.040)	0.059	213	-0.026	(-0.056, 0.003)	0.079

Model 3 is Model 1 stratified by gender - analysing females only and without including sex as a confounder

Model 4 is Model 1 stratified by gender - analysing males only and without including sex as a confounder

Appendix 4: Inclusion of a BMI squared term into ALSPAC linear regression models to test for a "U" shaped relationship between obesity and depression

Table A4.1 Results of linear regression analysis investigating the association between BMI (exposure) and depression (outcome) in the ALSPAC cohort, including a BMI squared term in the model

Timepoint	Exposure	Outcome	Model 1 - main model				Model 2 - Including interaction (coefficient for interaction effect)			
			n	coeff.	95%CI	p-value	n	coeff[#].	95%CI	p-value
F10 to TF1 10y8m to 12y10m	BMI^2	Depression	4264	-0.0003	(-0.002, 0.003)	0.792	4264	-0.00003	(-0.0005, 0.004)	0.895
TF1 to TF2 12y10m to 13y10m	BMI^2	Depression	3964	0.0003	(-0.001, 0.002)	0.705	3964	0.0004	(-0.00004, 0.0008)	0.074
TF2 to CCS 13y10m to 16y8m	BMI^2	Depression	2864	0.0007	(-0.001, 0.003)	0.510	2864	0.0006	(0.00004, 0.001)	0.033
TF4 to CCT 17y10m to 18y8m	BMI^2	Depression	1723	-0.0002	(-0.002, 0.001)	0.828	1723	0.0002	(-0.0002, 0.0006)	0.328

Model 1 is adjusted for age (at outcome), sex, previous depression, maternal depression, maternal education and maternal profession

*Model 2 is Model 1 plus the inclusion of a BMI^2 *Sex interaction term*

[#] *Results presented are for the BMI^2 * Sex interaction term*

Timepoint	Exposure	Outcome	Model 3 - stratified (females)				Model 4 - stratified (males)			
			n	coeff.	95%CI	p-value	n	coeff.	95%CI	p-value
F10 to TF1 10y8m to 12y10m	BMI^2	Depression	2172	-0.0005	(-0.003, 0.002)	0.725	2092	0.0001	(-0.003, 0.003)	0.942
TF1 to TF2 12y10m to 13y10m	BMI^2	Depression	2013	-0.0007	(-0.003, 0.001)	0.519	1951	0.001	(-0.001, 0.004)	0.254
TF2 to CCS 13y10m to 16y8m	BMI^2	Depression	1621	0.001	(-0.001, 0.004)	0.376	1243	-0.001	(-0.005, 0.003)	0.583
TF4 to CCT 17y10m to 18y8m	BMI^2	Depression	1091	-0.0006	(-0.003, 0.001)	0.577	632	0.0008	(-0.001, 0.003)	0.544

Model 3 is Model 1 stratified by gender - analysing females only and without including sex as a confounder

Model 4 is Model 1 stratified by gender - analysing males only and without including sex as a confounder

Appendix 5: Impact of puberty on the relationship between obesity and depression in adolescent females

Table A5.1 Results of linear regression analyses investigating the relationship between obesity and depression in adolescent females in the ALSPAC cohort, model is additionally adjusted for puberty

Timepoint	Exposure	Outcome	Results		
			coeff.	95%CI	p-value
F10 to TF1					
10y8m to 12y10m	BMI	Depression	0.007	(-0.011, 0.026)	0.432
n=1172	Experienced 1st Menarche		0.060	(-0.053, 0.174)	0.296
TF1 to TF2					
12y10m to 13y10m	BMI	Depression	0.016	(-0.0004, 0.032)	0.056
n=1128	Experienced 1st Menarche		0.050	(-0.065, 0.165)	0.393
TF2 to CCS					
13y10m to 16y8m	BMI	Depression	0.024	(0.009, 0.039)	0.002
n=1621	Experienced 1st Menarche		-0.090	(-0.189, 0.009)	0.073
TF4 to CCT					
17y10m to 18y8m	BMI	Depression	0.005	(-0.009, 0.019)	0.649
n=1091	Experienced 1st Menarche		-0.088	(-0.195, 0.018)	0.104

Model is adjusted for age (at outcome), previous depression, maternal depression, maternal education and maternal profession

Appendix 6: Impact of the inclusion of physical activity and BMI in the same model investigating the relationship with depression

Throughout the thesis obesity and physical activity have been considered and investigated separately in their potential relationships with adolescent depression. However, as obesity and physical activity are likely to be interrelated, it is possible that these variables may confound one another. Therefore the linear regression analyses have been repeated where both a measure of obesity and physical activity have been included in the same (where possible). The conclusions drawn from the analyses which included both obesity and physical activity did not differ from the models where these variables were considered separately.

Table A6.1 Results of linear regression analyses including a measure of both obesity and physical activity in the relationship with depression in the ALSPAC cohort

Timepoint	Exposure	Coeff	95% CI	p-value
TF2 to CCS	BMI	0.016	(0.002, 0.029)	0.020
13y10m to 16y6m	Total daily minutes of PA	0.0003	(-0.0004, 0.0009)	0.379
n=2024				
TF2 to CCS	BMI	0.016	(0.002, 0.029)	0.020
13y10m to 16y6m	Accelerometer counts per minute	0.0001	(-0.0002, 0.0003)	0.531
n=2024				
TF2 to CCS	BMI	0.016	(0.003, 0.029)	0.019
13y10m to 16y6m	Daily minutes of MVPA	0.001	(-0.001, 0.003)	0.455
n=2024				
TF2 to CCS	BMI	0.016	(0.003, 0.029)	0.020
13y10m to 16y6m	Percentage of time spent in MVPA	0.005	(-0.013, 0.023)	0.563
n=2024				
TF2 to CCS	BMI	0.015	(0.002, 0.028)	0.022
13y10m to 16y6m	At least 1 hour of MVPA a day	0.056	(-0.121, 0.233	0.538
n=2024				
TF2 to CCS	BMI	0.015	(0.003, 0.026)	0.014
13y10m to 16y6m	Self reported frequency of PA in past year			
n=2558	*Never*	ref		
	Less than once a month	0.048	(-0.431, 0.527)	
	1-3 times a month	-0.260	(-0.599, 0.078)	
	1-4 times a week	-0.206	(-0.524, 0.112)	
	5 or more times a week	-0.281	(-0.598, 0.037)	0.053

Models adjusted for age, sex, previous depression, maternal depression, maternal education, maternal education and accelerometer weartime (where appropriate)

Table A6.2 Results of linear regression analyses including a measure of both obesity and physical activity in the relationship with depression TRAILS cohort

Timepoint	Exposure	Coeff	95% CI	p-value
T1 to T2	BMI	0.015	(0.001, 0.030)	0.037
10y7m to 13y1m	No. of days of PA a week			
n=1828	*Never*	ref		
	Once a week	0.050	(-0.107, 0.207)	
	2 or 3 days a week	0.011	(-0.136, 0.157)	
	4 or 5 days a week	-0.006	(-0.166, 0.154)	
	6 or 7 days a week	-0.045	(-0.205, 0.115)	0.720
T2 to T3	BMI	0.017	(0.003, 0.031)	0.019
13y1m to 15y10m	No. of days of PA a week			
n=1467	*Never*	ref		
	Once a week	-0.077	(-0.279, 0.125)	
	2 or 3 days a week	-0.059	(-0.239, 0.120)	
	4 or 5 days a week	-0.066	(-0.250, 0.117)	
	6 or 7 days a week	0.031	(-0.175, 0.238)	0.635
T3 to T4	BMI	-0.013	(-0.026, 0.001)	0.071
15y10m to 18y7m	No. of days of PA a week			
n=1267	*Never*	ref		
	Once a week	-0.090	(-0.290, 0.109)	
	2 or 3 days a week	-0.033	(-0.211, 0.145)	
	4 or 5 days a week	0.031	(-0.156, 0.219)	
	6 or 7 days a week	-0.130	(-0.311, 0.052)	0.118

Model adjusted for age, sex, previous depression, maternal depression, SEP, alcohol and smoking

Table A6.3 Results of linear regression analyses including a measure of both obesity and physical activity in the relationship with depression in the NDIT cohort

Timepoint	Exposure	Coeff	95% CI	p-value
T1 to T2	BMI	-0.008	(-0.024, 0.009)	0.356
12y9m to 13y0m	No. bouts of MVPA	-0.00004	(-0.004, 0.004)	0.987
n=496				
T12 to T13	BMI	0.023	(0.005, 0.041)	0.011
15y2m to 15y7m	No. bouts of MVPA	-0.001	(-0.006, 0.005)	0.828
n=433				
T19 to T20	BMI	-0.004	(-0.020, 0.011)	0.586
17y0m to 17y1m	No. bouts of MVPA	-0.002	(-0.009, 0.005)	0.673
n=358				

Model is adjusted for age, sex, previous depression, maternal education, maternal profession and alcohol use

Appendix 7: Investigation into the association between genetic instrument for obesity with BMI and confounders

Table A6.1 Results of the investigation into the association between genetic instrument for obesity and BMI

Time Point	n	coefficient	95% CI	p-value
F10: 10y8m	5584	0.105	(0.093, 0.118)	<0.001
TF1: 12y10m	5087	0.118	(0.102, 0.133)	<0.001
TF2: 13y10m	4719	0.113	(0.098, 0.129)	<0.001
TF4: 17y10m	3724	0.124	(0.103, 0.144)	<0.001

Table A6.2 Results of the investigation into the association between genetic instrument for obesity and confounding variables

	n	coefficient	95% CI	p-value
Maternal Depression	5884	-0.006	(-0.026, 0.015)	0.584
Sex: Female	8313	1.000*	(0.994, 1.008)	0.785
Maternal Education	7363	0.994#	(0.987, 1.000)	0.066
Maternal Profession	6219	1.007#	(1.000, 1.015)	0.059

**coefficient is an Odds Ratio from logistic regression*

 # coefficient is an Odds Ratio from ordinal regression

Appendix 8: Results of mediation analysis stratified by sex

Table A8.1 Results of the investigation into potential mediators of the obesity – depression relationship in adolescence, stratified by sex

Mediator	Age at obesity measurement	Age at mediator measurement	Age at depression measurement	Indirect effect coeff. (SE)	Direct effect of obesity coeff. (SE)
ALSPAC					
CRP (Males)	15y6m	15y6m	16y6m	0.008 (0.007)	-0.004 (0.003)
CRP (Females)	15y6m	15y6m	16y6m	0.011 (0.009)	0.032* (0.010)
TRAILS					
Cortisol (Males)	13y1m	15y9m	15y9m	0.007 (0.005)	0.011 (0.010)
Cortisol (Females)	13y1m	15y9m	15y9m	0.004 (0.003)	0.014 (0.014)
CRP (Males)	13y1m	15y9m	15y9m	0.005 (0.004)	0.015 (0.013)
CRP (Females)	13y1m	15y9m	15y9m	0.006 (0.005)	0.011 (0.010)
IgE (Males)	13y1m	15y9m	15y9m	0.011 (0.010)	0.006 (0.005)
IgE (Females)	13y1m	15y9m	15y9m	0.010 (0.008)	0.008 (0.008)
Body Image (Males)	13y1m	15y9m	15y9m	0.005 (0.003)	0.012 (0.011)
Body Image (Females)	13y1m	15y9m	15y9m	0.007 (0.006)	0.011 (0.009)
Self Esteem (Males)	15y9m	18y9m	18y9m	0.006 (0.005)	0.008 (0.007)
Self Esteem (Females)	15y9m	18y9m	18y9m	0.007 (0.005)	0.007 (0.006)

**p-value <0.05*

Appendix 9: Brief discussion of missing data

In longitudinal cohort studies where repeated waves of follow up take place it is often difficult to collect complete data on all participants. This is potentially problematic for researchers as missing data results in a loss of precision and power, and may also introduce bias to a study. As would be expected there is missing data in the three longitudinal cohorts that provided data for this thesis (ALSPAC, TRAILS and NDIT).

The first time point from which data was used in this thesis from the TRAILS and NDIT cohorts was also the first wave of data collection in these cohorts. The ALSPAC cohort however is a prospective birth cohort, the first wave of data used in this thesis was collected at a (mean) age of approximately 10 years 8 months. Therefore, there was already drop out between the first measurement occasion and the first measurement of obesity, physical activity and depression used in the thesis (data was collected at this time point on 7457 of the 14701 individuals who provided data at the first measurement occasion (51%)). In the linear regression analysis investigating the relationship between obesity and depression utilising outcome data from the final wave of follow up, the number of individuals who could be included in this analysis was only 1723 (12% of the individuals who provided data at the first measurement occasion, or 23% of the individuals who provided data at the first measurement occasion used in the thesis).

In the TRAILS cohort the amount of missingness was much smaller than in the ALSPAC cohort, for example at the final time point used in the thesis 1881 individuals (84% of the original sample size) provided at least some data, with 1696 individuals (76% of the original sample size) providing data on the depression outcome. Although the overall retention rate in the TRAILS cohort was relatively

high this does not necessarily mean that missing data was not a potential issue in the analyses. For example, in the linear regression analysis investigating the relationship between obesity and depression utilising outcome data from the final wave of follow up, the number of individuals who could be included in this analysis was 1276 (57% of the individuals who provided data at the first measurement occasion). Similarly, the number of individuals who could be included in the analysis of physical activity and depression at this time point was 1352 (60% of the starting sample).

The situation in the NDIT cohort regarding missing data was similar to that observed in the TRAILS cohort. Of the 1294 individuals that provided data at the first measurement occasion, 840 (65%) provided at least some data at follow up wave 20 (the final wave of data collection used in this thesis). This seems like a reasonable overall retention rate, however, in the linear regression analysis investigating the relationship between obesity and depression utilising outcome data from the final wave of follow up, the number of individuals who could be included in this analysis was 416 (32% of the individuals who provided data at the first measurement occasion).

The missing data in the three cohorts discussed above reduces the precision and power of analyses. The missingness may also introduce bias to the study. If individuals who are missing and those individuals who have complete data differ with respect to the exposure and outcome variables then bias may be introduced to complete-case analyses investigating the association between exposure and outcome if missingness is related to the outcome given the covariates in the model. Three "types" of missingness are often used to describe the missing data mechanism: missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR). The MCAR refers to the situation where missingness is independent of both the observed and unobserved data, MAR refers to the situation when missingness is independent of unobserved data but related to observed data, whilst MNAR is when missingness is related to the unobserved data. A complete-

case analysis is only unbiased (with respect to missing data) under the assumption of MCAR.

As a brief exploration of this potential issue, differences between those individuals who were and were not included in the first linear regression model investigating the association between obesity and later depressive symptoms (in each cohort) have been examined. In the TRAILS cohort there was no evidence of a difference in baseline depression, maternal depression, sex, alcohol or smoking frequency, however there was evidence of a difference in baseline BMI, SES and age between those who were and were not included in the regression model (Table A9.1). Those who were not included had a higher BMI, were of lower SES and were older at baseline than those who were included. In the NDIT cohort there was no evidence of a difference in any baseline characteristics between those who were and were not included in the regression model except for age and weak evidence for a difference in mean BMI (Table A9.2). Those who were missing were older and had a higher baseline BMI (Table A9.2). In the ALSPAC cohort those who were and were not included in the regression model differed with respect to BMI, depression, maternal depression, socioeconomic status, sex and age (Table A9.3). Those who were not included in the regression model had a higher BMI, higher depressive symptoms score, had lower socioeconomic status, were older, and were a greater proportion of males (Table A9.3).

Table A9.1 – Comparison of characteristics of individuals who were and were not included in the linear regression model investigating the association between BMI and depression in the TRAILS cohort

	Not included in regression model			Included in regression model			
	N	Mean (SD)	Median (IQR)	N	Mean (SD)	Median (IQR)	p-value*
BMI	394	18.7 (3.58)	17.9 (16.0, 20.6)	1836	17.9 (2.94)	17.3 (15.9, 19.3)	0.002
Depression	355	0.29 (0.25)	0.23 (0.08, 0.46)	1836	0.29 (0.25)	0.23 (0.08, 0.46)	0.683
Maternal depression	203	0.29 (0.35)	0.14 (0, 0.43)	1836	0.25 (0.34)	0.14 (0, 0.43)	0.114
SES:	352			1836			
Lowest 25%	155 (44%)			398 (22%)			
Middle 50%	145 (41%)			939 (51%)			
Highest 25%	52 (15%)			499 (27%)			<0.001
Sex:	394			1836			
Male	209 (53%)			889 (48%)			
Female	185 (47%)			947 (52%)			0.096
Age	394	10.7 (0.68)	11 (10, 12)	1836	10.6 (0.64)	11 (10, 12)	0.010
Lifetime Alcohol Use:	363			1836			
Never	247 (68%)			1271 (69%)			
Once	58 (16%)			283 (15%)			
2-3 times	29 (8%)			148 (8%)			
4-6 times	7 (2%)			64 (3%)			
7 times or more	22 (6%)			70 (4%)			0.194
Smoking frequency:	343			1836			
Not at all	335 (98%)			1791 (98%)			
Sometimes	7 (2%)			40 (2%)			
Often	1 (<1%)			5 (<1%)			0.985

**p-value from t-test (or Mann-whitney if data not normally distributed) if variable is continuous or Chi squared if categorical*

Table A9.2 – Comparison of characteristics of individuals who were and were not included in the linear regression model investigating the association between BMI and depression in the NDIT cohort

	Not included in regression model			Included in regression model			
	N	Mean (SD)	Median (IQR)	N	Mean (SD)	Median (IQR)	p-value*
BMI	699	20.3 (3.93)	19.4 (17.5, 22.1)	496	19.8 (3.69)	19.1 (17.2, 21.7)	0.047
Depression	720	2.1 (0.63)	2.0 (1.7, 2.5)	496	2.1 (0.59)	2.0 (1.7, 2.5)	0.982
Maternal depression:	82			496			
No	61 (75%)			402 (81%)			
Yes	21 (26%)			94 (19%)			0.162
Maternal education:	95			496			
High School - Attended	7 (7%)			41 (8%)			
High School - Graduated	24 (25%)			84 (17%)			
CEGEP - Attended	9 (9%)			41 (8%)			
CEGEP - Graduated	15 (16%)			77 (16%)			
University - Attended	9 (9%)			48 (10%)			
University - Graduated BSc	20 (21%)			121 (24%)			
University - Graduated MSc	6 (6%)			39 (8%)			
University - Graduated PhD	0 (0%)			4 (1%)			
Other	5 (5%)			41 (8%)			0.696
Maternal profession:	93			496			
Full-time job	54 (58%)			291 (59%)			
Part-time job	15 (16%)			97 (20%)			
Full-time student	1 (1%)			1 (<1%)			
Part-time student	0 (0%			1 (<1%)			
Homemaker	9 (10%)			43 (9%)			
Not working for health reasons	1 (1%)			9 (2%)			
Unemployed	3 (3%)			14 (3%)			
On welfare	2 (2%)			1 (<1%)			
Other	8 (9%)			39 (8%)			0.367
Sex:	798			496			
Male	390 (49%)			233 (47%)			
Female	408 (51%)			263 (53%)			0.507
Age	771	12.8 (0.60)	12.7 (12.4, 13.0)	496	12.7 (0.45)	12.6 (12.4, 12.9)	<0.001
Alcohol Use:	718			496			
Never	404 (56%)			288 (58%)			
A bit to try	215 (30%)			158 (32%)			
Once or a couple of times a month	74 (10%)			41 (8%)			
Once or a couple of times a week	19 (3%)			8 (2%)			
Everyday	6 (1%)			1 (<0.1%)			0.262

**p-value from t-test (or Mann-whitney if data not normally distributed) if variable is continuous or Chi squared if categorical*

Table A9.3 – Comparison of characteristics of individuals who were and were not included in the linear regression model investigating the association between BMI and depression in the ALSPAC cohort

	Not included in regression model			Included in regression model			
	N	Mean (SD)	Median (IQR)	N	Mean (SD)	Median (IQR)	p-value*
BMI	3110	18.5 (3.38)	17.7 (16.1, 20.2)	4264	18.1 (2.99)	17.4 (16.0, 19.6)	<0.001
Depression	3008	4.2 (3.63)	3 (2, 6)	4264	3.9 (3.41)	3 (1, 6)	0.001
Maternal depression	4581	6.4 (5.22)	5 (2, 9)	4264	5.7 (4.79)	5 (2, 8)	<0.001
Maternal education:	8076			4264			
CSE	2092 (26%)			402 (9%)			
Vocational	904 (11%)			311 (7%)			
O Level	2781 (34%)			1492 (35%)			
A Level	1515 (19%)			1256 (29%)			
Degree	784 (10%)			803 (19%)			<0.001
Maternal social class:	5773			4264			
I	261 (5%)			330 (8%)			
II	1598 (28%)			1544 (36%)			
III (non-manual)	2502 (44%)			1772 (42%)			
III (manual)	533 (9%)			250 (6%)			
IV	670 (12%)			313 (7%)			
V	166 (3%)			54 (1%)			
Armed Forces	3 (<1%)			1 (<1%)			<0.001
Sex:	14948			4264			
Male	7845 (52%)			2092 (49%)			
Female	7103 (48%)			2172 (51%)			<0.001
Age	3193	10.7 (0.30)	10.7 (10.5, 10.8)	4264	10.6 (0.23)	10.6 (10.4, 10.8)	<0.001

**p-value from t-test (or Mann-whitney if data not normally distributed) if variable is continuous or Chi squared if categorical*

There was a considerable amount of dropout in the three cohorts which provided data for this thesis. Missing data will always reduce the precision and power of the analyses which have been carried out, and may also cause bias if missingness is related to the outcome given the covariates in the model. In all three cohorts those who were not included in the analyses had a higher (mean) BMI than those who were included. If there is a positive association between obesity and later depressive symptoms, and missingness is associated with higher BMI, then this could bias findings towards the null. The cross-lagged SEM approach used as part of the analysis attempts to address the potential issue of loss of precision and bias introduced to complete-case analysis due to missing data. Cross-lagged SEM uses a maximum-likelihood approach to estimation, and as such both individuals with complete and incomplete data contribute information to the analysis. However, this method does rely on the MAR assumption, if missingness is related to the unobserved data (MNAR) then bias would still be introduced. For example, if those individuals who have a higher depression symptom score at the follow up time point are more likely to be missing (therefore MNAR) then this could bias results towards the null when investigating the relationship between obesity and later depressive symptoms. The MNAR assumption is not testable, but future work could investigate this further using statistical models that allow for MNAR data structures.