Using Scotland-wide record linkage to investigate the educational and health outcomes of children treated for chronic conditions

Michael Fleming

BSc (honours), MSc

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy (PhD)

Institute of Health and Wellbeing

College of Medical, Veterinary and Life Sciences (MVLS)

University of Glasgow

June 2017

Abstract

Objectives: This study linked Scottish education data to a number of administrative health datasets to explore associations between childhood chronic ill health and subsequent educational and health outcomes. Chronic conditions investigated were diabetes, asthma, epilepsy, attention deficit hyperactivity disorder (ADHD) and depression. Educational outcomes were number of days absent from school, number of school exclusions, special educational need (SEN), academic attainment and unemployment. Health outcomes were all-cause and cause-specific hospital admission, total number of hospital admissions, total length of hospital admission and all-cause mortality.

Approach: Pupil census data and associated education records for all children attending primary and secondary schools in Scotland between 2009 and 2013 were linked to national prescribing data, hospital admissions, death records and retrospective maternity records enabling outcomes to be studied whilst controlling for socioeconomic, demographic and obstetric factors including birth outcomes and maternal antecedents. Specific medications are prescribed for some particular chronic conditions; therefore, children identified as receiving these medications whilst at school were assumed to have these conditions.

Results: Children treated for each of the five conditions had more frequent absenteeism from school and were more likely than their peers to have SEN. However, only children treated for depression, epilepsy or ADHD experienced poorer academic attainment and increased odds of unemployment. Furthermore, children treated for depression or ADHD were significantly more likely to be excluded from school. Children treated for asthma experienced poorer academic attainment but no increased odds of unemployment and the association with attainment disappeared after adjusting for their increased absenteeism. Children treated for each of the five conditions had an increased risk of hospital admission and children treated for depression or epilepsy also had an increased risk of recurrent hospitalisation and longer stays in hospital. All of the chronic conditions, with the exception of ADHD, were associated with increased mortality.

Conclusion: All five of the chronic conditions investigated in this thesis were associated with adverse educational and health outcomes. The number of outcomes affected varied by condition. Treated depression, epilepsy and ADHD were associated with the most wide-ranging impacts. Children treated for depression fared worse than their peers across all nine outcomes, and children treated for epilepsy and ADHD across eight and six

respectively. In contrast, children treated for asthma and diabetes fared worse than their peers in respect of around half the outcomes investigated. Children with these chronic conditions at school appear to experience significant educational and health disadvantage; therefore further work is required to understand the underlying mechanisms and to develop effective interventions to reduce their risk.

Table of Contents

List of Tables

List of Figures

List of Publications

Fleming M, Fitton CA, Steiner MFC, McLay JS, Clark D, King A, Mackay DF, Pell JP. Educational and Health Outcomes of Children Treated for Attention-Deficit/Hyperactivity Disorder. JAMA Pediatrics. 2017;171(7):e170691

Dedication

This thesis is dedicated to my mum Joyce who passed away on 23rd October 2014 during the first year of my PhD. I miss you every single day and without your unwavering love and support at every stage of my life I would not have been in a position to undertake this thesis.

Acknowledgements

I would firstly like to thank my three supervisors: Professor Jill Pell and Dr Daniel Mackay at the University of Glasgow and Dr James McLay at the University of Aberdeen. Together, they provided invaluable help, support, guidance, supervision, reassurance and advice throughout every stage of the PhD. Thanks to all for reviewing the many incomplete drafts of this thesis, to Danny for statistical advice, to James for clinical input and to Jill for initiating and leading the project and for showing faith and confidence in me back in 2013 by giving me the opportunity to undertake this PhD.

I'd also like to thank (in no particular order):

- The Farr Institute for funding this project

- My employers ISD Scotland for allowing me to collaborate with Glasgow University and undertake this PhD

- Dave Clark from the eDRIS team in ISD who expertly undertook the linkage of the pupil census to the CHI database enabling it to be linked to health data.

- Albert King from ScotXed who prepared all of the education datasets and provided ongoing advice and support around the education data.

- The wider eDRIS team within ISD. Specifically Dave Bailey and Katrina Smith, who were the two allocated research coordinators on the project, and Carole Morris, who stepped in to cover on several occasions. Together, they provided expert advice around various governance issues, help and guidance on completing the necessary governance documentation, ongoing support throughout the PhD and an excellent service as regards reviewing and signing out copious amounts of output from the safe haven.

- Stuart McTaggart at ISD Scotland who provided additional help around the PIS prescribing data

- My advisors of study Professor Colin McCowan and Dr James Lewsey who undertook my annual reviews and provided excellent guidance and advice.

- My two ISD managers throughout the duration of the study: Joan Forrest and Jean Sneddon. Both showed support and understanding particularly throughout the writing up period.

- The administrative staff at Glasgow University, particularly Lorraine Waddell, Margaret Ashton and Rachel Allen for help and support throughout.

- Glasgow University library staff. In particular Blair Thompson, for help with preparing the layout of the thesis document in word, and Heather Worlledge-Andrew for early help and advice regarding my literature review.

- Joseph Hilbe for providing valuable Stata code allow me to perform negative binomial modelling

- Dr Daniel Mackay and Dr John McClure for giving me the opportunity to gain some valuable tutoring and lecturing experience during the PhD

- Dr Frank Popham and Professor Frank Sullivan who were my PhD viva examiners and Professor Andrew Jahoda who acted as convenor.

- My family and friends who have supported me over the duration of the PhD. In particular my dad (James) and sister (Hazel) and my mum (Joyce) who passed away nine months into the PhD and whom I miss every day. You have all been there with me from the very beginning supporting me in everything that I have done and helping whenever you can. Thanks to my parents for instilling in me the importance of a good education very early on in my life.

- Last but by no means least my wonderful wife Laura. I feel like we have breathed every single second of this PhD together. You have supported me hugely and continuously from day one and have been my main cheerleader, calming influence, voice of reason and positivity, and have kept me sane. Not to mention your unwavering patience, endless words of encouragement at every stage and the seemingly endless cups of tea, coffee, meals and snacks to get me through writing up. I'm looking forward to spending a lot more time with you once this thesis is finally submitted!

Author's Declaration

I declare that the contents of this thesis are my own work and have not been submitted for any other degree at the University of Glasgow or any other institution. Where the work of others has been used it has been indicated and appropriately referenced

Michael Fleming

Definitions/Abbreviations

ADHD	Attention Deficit Hyperactivity Disorder
AED	Anti-Epileptic Drug
AIC	Akaike Information Criterion
AOA	Atos Origin Alliance
BECTS	Benign Epilepsy with Centrotemporal Spikes
BIC	Bayesian Information Criterion
BNF	British National Formulary
CAHMS	Child and Adolescent Mental Health Services
CAT	California Achievement Test
CBT	Cognitive Behavioural Therapy
CfE	Curriculum for Excellence
CHI	Community Health Index
CHIAG	Community Health Index Advisory Group
CNS	Central Nervous System
COPD	Chronic Obstructive Pulmonary Disorder
DSM	Diagnostic and Statistical Manual
EAT	Educational Achievement Test

ECT	Electroconvulsive therapy
eDRIS	Electronic Data Research and Innovation Service
EET	In Education, Employment or Training
EPCC	Edinburgh Parallel Computing Centre
EPR	Electronic Patient Record
GEE	Generalised Estimating Equation
HR	Hazard Ratio
ICD	International Classification of Diseases
IQ	Intelligence Quotient
IRR	Incidence Rate Ratio
ISD	Information Services Division
ITBS	Iowa Test of Basic Skills
ITED	Iowa Test of Educational Development
LABA	Long Acting Beta Agonist
LD	Learning Difficulty
LTRA	Leukotriene Receptor Antagonists
MRR	Mortality Rate Ratio
NEET	Not in Education, Employment or Training
NHS	National Health Service

NRS	National Records of Scotland
NSS	National Services Scotland
OR	Odds Ratio
PAC	Privacy Advisory Committee
PBPP	Public Benefit and Privacy Panel
PIAT	Peabody Individual Achievement Test
PIS	Prescribing Information System
PSD	Practitioner Services Division
QIC	Quasi-likelihood under the independence model criterion
SABA	Short Acting Beta Agonist
SCN	Scottish Candidate Number
ScotXed	Scottish Exchange of Educational Data
SEN	Special Educational Need
SFTP	Secure File Transfer Protocol
SG	Scottish Government
SMR (01/02/04)	Scottish Morbidity Record (01/02/04)
SMR	Standardised Mortality Ratio
SNRI	Selective Norepinephrine Reuptake Inhibitor
SQA	Scottish Qualifications Authority

SSRI	Selective Serotonin Reuptake Inhibitor
SUDEP	Sudden unexpected deaths in epileptic patients
T1DM/T2DM	Type 1 Diabetes Mellitus/ Type 2 Diabetes Mellitus
VPN	Virtual Private Network
WIAT	Wechsler Individual Achievement Test
WJ	Woodcock Johnson Test of Achievement
WRAT	Wide Range Achievement Test
ZINB	Zero Inflated Negative Binomial
ZIP	Zero Inflated Poisson

1 Introduction

The relationship between education and health is bidirectional. Well established modifiable risk factors for poor health include smoking, alcohol consumption, diet and physical activity whilst non-modifiable risk factors include age, gender and genetic predisposition [1]. However, the determinants of health extend beyond just the health sector and also include a range of social and environmental determinants [2,3]. Education is one such determinant which can impact health both directly and indirectly. Educating people on healthcare and healthy lifestyles can directly improve their health outcomes. However, wider and more general improvements in overall educational performance and attainment can indirectly impact a person's future health, wellbeing and quality of life via subsequent benefits of increased employment, higher income, better working conditions, improved living conditions, greater access to healthcare and leisure activities, and better neighbourhoods. In this respect, school performance is critical to a person's future prospects [4]. However, many factors can impact on a child's educational outcomes. Potential social and environmental barriers to a good education include poverty, a troubled home life or poor parental influence [5]. Important health factors include childhood chronic disease and poor physical or mental health or disability. A chronic condition is defined as one which lasts for longer than three months. Whilst they can be treated by medication, they typically cannot be cured [6]. Indeed children with chronic conditions are among those most at risk of suffering poor education at school and previous studies have reported associations between a range of chronic conditions and poor school performance in addition to long lasting effects on health and wellbeing [7,8].

Improving our understanding of the relationships between health and education is therefore a very important area of public health and cross-sectoral record linkage provides an invaluable resource for epidemiological research. Studies in other countries have identified negative effects of diabetes [9,10], epilepsy [11], asthma [12], attention deficit hyperactivity disorder [13] and depression [14,15] on school outcomes but there is a paucity of literature in Scotland and the wider UK and very few studies worldwide have analysed population-wide data. Population-wide Scottish pupil census data have previously been linked to obstetric records to demonstrate links between gestational age/low birth weight and increased risk of special educational need [16,17]. My thesis built on these foundations by linking education records to a wider range of health datasets including prescribing, hospital admissions, maternity data and death records in order to investigate the impact of five particular

childhood chronic conditions on subsequent educational and health outcomes whilst adjusting for sociodemographic, maternity and health confounding factors.

1.1 Study aims and objectives

The overall aim of this thesis was to study the educational and health outcomes of children treated for chronic conditions compared with their peers. Specific medications are prescribed for some particular chronic conditions; therefore, children identified as receiving these medications whilst at school were assumed to have these conditions. Education and prescribing data were linked together to identify children treated for five specific chronic conditions whilst at school: diabetes; asthma; epilepsy; ADHD; depression. Educational data allowed specific school outcomes to be investigated: number of days absent; number of exclusions; special educational need; educational attainment; unemployment. Linkage to hospital admissions and death records enabled investigation of health outcomes: all-cause hospital admission; cause-specific hospital admission; all-cause mortality. Linkage to maternity data enabled all of the statistical models to be adjusted for maternal confounders in addition to sociodemographic confounders obtained from the education datasets. Children may experience more than one disease; therefore, final statistical models included all diseases together to investigate whether the effects of each disease were independent of the other conditions.

In order to meet the research objectives above, early stages of the thesis included developing appropriate governance and obtaining relevant approvals to enable linkage of the relevant education and health datasets. Further early stages involved developing techniques to identify children with chronic conditions using prescribing data and developing methodology to create various outcome specific cohorts capable of answering the various research questions. An overall strategic aim was to demonstrate the feasibility and utility of linking a series of health sector and educational sector databases using the recent infrastructure and governance developments led by Farr Scotland. This included linking and accessing large volumes of data from more than one governmental sector within the national safe haven.

Specific research objectives were:

1. Develop governance documentation to obtain approval for linkage of health and education data

2. Perform a contextual review of the literature for each chronic condition

3. Develop methods to identify children treated for each chronic condition whilst at school using linked prescribing data

4. Develop methods to manipulate complex linked data and build specific datasets capable of answering the specific research questions.

5. Investigate the association between each chronic condition and a range of educational and health outcomes

Specific research questions were:

1. Do children with diabetes, asthma, epilepsy, ADHD or depression have significantly more days absent per school year compared to peers

2. Are children with diabetes, asthma, epilepsy, ADHD or depression significantly more likely to be excluded from school compared to peers

3. Are children with diabetes, asthma, epilepsy, ADHD or depression significantly more likely to be recorded as having special educational need compared to peers

4. Do children with diabetes, asthma, epilepsy, ADHD or depression have significantly poorer academic attainment compared to peers

5. Are children with diabetes, asthma, epilepsy, ADHD or depression significantly more likely to be unemployed after leaving school compared to peers

6. Do children with diabetes, asthma, epilepsy, ADHD or depression have significantly increased risk of all-cause and cause-specific hospital admission compared to peers

7. Do children with diabetes, asthma, epilepsy, ADHD or depression have significantly increased risk of all-cause mortality compared to peers

Relevant interactions were tested and sub-group analyses were performed where appropriate. Sensitivity and mediation analyses were performed whereby attainment and

unemployment were re-analysed after excluding children with SEN and adjusting for school absenteeism. This thesis investigated the overall impact of having each chronic condition on the full range of outcomes. It did not investigate disease specific factors such as disease severity, disease subtypes, age at diagnosis, or length of diagnosis. The thesis also did not investigate medication specific effects such as type or dose of medication.

1.2 Thesis structure

This thesis contains six chapters. This chapter included a brief introduction and set out the main aims and objectives of the study. Chapter 2 contains a separate literature review for each of the five chronic conditions of interest: diabetes, asthma, epilepsy, ADHD and depression. Each condition specific review critically appraises previous comparable studies investigating the educational and health outcomes relevant to this thesis. Chapter 3 describes the datasets and data providers that were used in this thesis and provides background on the current linkage infrastructure in Scotland. It then describes the information governance and linkage processes that I put in place to enable data to be linked together and made available for analyses. Chapter 4 describes the various methods used in this thesis. Firstly, I describe primary linkage methods used to link education data to health data. I then present techniques used to clean the various datasets, manipulate them and append them together in order to create separate outcome specific cohorts capable of answering the various research questions of interest. I next summarise methods used to identify children with each chronic condition of interest based on prescription medications taken whilst at school. Finally I summarise the statistical methods used to investigate each of the educational and health outcomes of interest. Chapter 5 presents results from the various analyses. Some initial results are presented from the data cleaning stage and this is followed by tables of demographic characteristics. The main results for each outcome of interest are presented within separate sub chapters. Each sub chapter contains five condition-specific analyses relating to a particular outcome of interest. The outcomes of interest, in order, are: absences, exclusions, SEN, attainment, unemployment, hospital admissions and mortality. Finally, chapter 6 discusses the findings of the various analyses, puts these into a wider context, makes recommendations and discusses potential policy implications. I then discuss the strengths and limitations of this thesis, outline potential future work, and briefly discuss challenges and lessons learned.

2 Literature review

A literature search using the Medline database was conducted for each chronic condition of interest to identify original articles reporting any of the educational or health outcomes relevant to my thesis: school absence, exclusion, special educational need, academic attainment, unemployment, hospital admissions or mortality. My thesis analysed a population of schoolchildren and compared outcomes for children with and without each of the five chronic conditions in turn. Therefore, within each condition-specific literature search, only articles which similarly compared cohorts of school aged children with and without that disease were identified and reviewed. Articles that solely analysed a chronically ill cohort or with no comparison group were not included in the review. In keeping with the main thesis outcomes, the review, where possible, identified articles reporting on all-cause admission and all-cause mortality as opposed to cause-specific health outcomes.

With regard to academic attainment, the review only identified articles investigating academic outcomes rather than cognitive ability because the former were more pertinent to this thesis. Cognition refers to a group of specific mental processes, including memory, problem solving, decision-making, learning and attention all of which can contribute to academic ability. Cognitive assessments such as the Wechsler Intelligence Scale for Children (WISC-IV), the Differential Ability Scales (DAS-II), and the Stanford Binet Intelligence Scale (SB-V) provide information on cognitive strengths and weaknesses. Academic ability, by comparison, relies on broad skills developed across a number of cognitive and academic domains and, within a school setting, is most commonly monitored via performance in school exams. In the absence of school grades, however, academic achievement tests such as the Wide Range Achievement Test (WRAT), Wechsler Individual Achievement Test (WIAT), Woodcock Johnson Test of Achievement (WJ) and the Kaufman Test of Educational Achievement define academic ability with respect to reading, spelling, arithmetic and language skills and assess these competencies between children of similar age. Some previous studies reporting academic outcomes have used cognitive tests and vice versa. Articles were only included in this review if the title or abstract indicated investigation of academic rather than cognitive outcomes and final analyses were based on school grades, school performance or academic test scores rather than tests of cognition. Articles describing cognitive outcomes in their title were not included regardless of whether they used academic tests.

2.1 Diabetes

Diabetes mellitus is a chronic condition where the concentration of glucose in the blood is too high. Insulin is a hormone produced by the pancreas and regulates glucose in the bloodstream; however in people with diabetes, the pancreas either stops making insulin completely (known as type 1 diabetes or T1DM) or the body becomes resistant to insulin (known as type 2 diabetes or T2DM) [18,19]. T1DM is an autoimmune condition which most often occurs in childhood whereas T2DM generally occurs much later in life and is linked to poor lifestyle and diet, obesity, lack of exercise and family history [18,20]. Whilst T1DM is rarer than T2DM in the general population, it is by far the most common type found in children and adolescents [20]. Risk factors for T1DM are not well understood; however, ethnicity and a variety of genetic and environmental factors are believed to be important [18,20,21].

2.1.1 Prevalence

Despite a very small increase in T2DM in younger populations due to poor diet, reduced physical activity and rising prevalence of obesity [22], the vast majority of diabetic children have T1DM [18,20,23]. Most people are diagnosed with T1DM in childhood with the incidence peaking between the ages of 10 to 14 and stabilising thereafter [18,20,21,23,24]. The incidence of type 1 diabetes has been increasing worldwide, yet the reasons for this increase are unknown [19,21,24]. The incidence of childhood onset T1DM varies greatly between countries. The highest incidence rates have been found in Finland (40 cases per 100,000) whilst the lowest incidence rates have been observed in Venezuela (0.1 per 100,000). Incidence in people younger than 15 years of age in England and Wales is between 22 and 26 cases per 100,000 children per year [24,25] and prevalence is around 195 per 100,000 children [25]. In 2013 the incidence of T1DM in children under 19 in Scotland was 34 per 100,000. Whilst 270,000 people in Scotland had diabetes, 30,000 had T1DM [23]. Approximately 29,000 children and young people in the UK have diabetes: 26,500 T1DM; 500 T2DM; 2,000 of unknown type [18]. Elsewhere, prevalence of 0.39% within school aged children has been reported [9].

2.1.2 Treatment

Diabetic patients must carefully balance food, medications, and physical activity to keep blood glucose levels as close to normal as possible [26]. Type 1 diabetics receive insulin through either injections or an insulin pump and can experience wide fluctuations in blood glucose levels throughout the day ranging from very low (hypoglycaemia) to very high (hyperglycaemia) depending on the timing, type and dose of insulin administered, food ingested and amount of exercise [27]. Type 2 diabetics may be able to control their disease through diet and exercise alone or may require oral medications and/or insulin injections. Over many years, high blood glucose levels can cause damage to the eyes, kidneys, nerves, heart, and blood vessels and those with diabetes suffer increased morbidity and mortality and increased hospitalisation [19,20]. The long term goal of therapy is to prevent long term complications associated with chronic hyperglycaemia such as cardiovascular disease, nephropathy, retinopathy and neuropathy [19,20]. However an excess of hospitalisation and mortality are also observed in young type 1 diabetic patients with a short duration of illness [28]. Therefore the short-term goal of therapy is to avoid potentially fatal fluctuations in blood sugar levels.

2.1.3 Contextual review of the literature

A Medline search using the search terms in Table 1 returned 524 articles of which 110 (30 education and 80 health) remained after review of title and 55 (19 education and 36 health) remained after review of abstract. After full text review, 17 education papers and 9 health papers from the UK, USA, Sweden and Finland remained relevant to my thesis.

Table 1 Search terms for diabetes literature review

(((Diabet* or IDDM or DM or T1DM)
AND
(college or youth or juvenile or young or schol* or adolescent* or child* or student* or pupil* or school)
AND
(school* or education* or attainment or academic or employment* or hospital* or death* or mortality or "learning difficulty" or "learning difficulties" or "learning disability" or "learning disabilities" or exclu* or "special need" or "special needs" or achievement or expulsion or suspension or detention or GPA or grade* or SEN or "assisted support needs" or attendance* or absen* or injur* or leav* or admission*))
NOT
(trial* or intervention* or cross-sectional or "college hospital" or "college teaching hospital" or "absence of")).ti.

**Filters applied: Title, Age 0-18 years, English language

2.1.3.1 School outcomes

Identified studies either reported that children with diabetes experienced worse academic attainment compared with peers or siblings [9,10,29-33] or reported no difference [34-37]. Academic attainment has been measured in several ways across different studies. Dalquist et al [9] linked a Swedish Childhood Diabetes Register to a Swedish Education Register and a National Birth Register and reported that 5,159 children diagnosed with diabetes before 15 years of age had lower school grades on leaving compulsory school compared with 1,330,968 healthy peers. They were more likely to fail exams in English, Swedish, Maths and Sport and were less likely to obtain high marks in English and Sport. This was the only other population-wide study reporting educational outcomes identified. Diabetes cases and school outcomes, including subject specific grades, were attained from population-wide national registers; therefore the large sample size and utilisation of detailed linked administrative data were significant strengths. Furthermore, linkage to birth records allowed adjustment for maternal confounders. Whilst the analyses were adjusted for year of birth, maternal age, maternal parity and maternal education, the investigators did not appear to adjust for socioeconomic status, gender, ethnicity or other potentially important maternity variables such as birthweight or gestational age. Additionally, school grades were not available for children with special educational needs or for those attending special schools. Diabetes has been associated with cognitive impairment therefore some diabetic children may have increased requirement for special education and poorer grades. Valuable information may have been lost through omission of this subgroup of children.

Persson et al [10] used a subset of the same cohort and reported lower mean final grades in compulsory school (16 years) and secondary school (19 years) for 2,485 diabetic children compared with 9,940 peers matched by age and residence. Diabetic children were also less likely to be employed at age 29 years. Matching of controls and further adjustment for gender, socioeconomic background and parental country of birth were strengths of this study; however the authors failed to incorporate maternity data previously linked by Dalquist et al. The study was further limited by decreased power because the analyses of only numeric grades, as opposed to all grades, led to the loss of over half of the original diabetes cases.

Two other studies analysed school reported grades; however, these were not national in coverage and used very small samples. Meo et al [32] reported that 36 diabetic students had worse mean grades compared with 36 peers matched by age, gender, ethnicity and

socioeconomic status. Mean grades were taken from written exams in English, maths, physics, chemistry, biology and humanities. Erkolahti et al [34] by contrast reported no difference in mean school recorded grades for diabetic children compared with two control groups matched for age, gender, social background and living environment. Only 69 pupils were included; however the study investigated both a matched healthy peer group and a matched group suffering from another chronic illness; in this case rheumatoid arthritis. It is possible that, in addition to the specific physiological effects of having particular chronic conditions, generic daily and mental burden of dealing with any form of ill health can independently impact educational outcomes [7,8]. Comparing groups of chronically ill patients is potentially one way to determine if effects are specific to diabetes or common to other chronic conditions.

Previous studies have compared diabetic children to sibling control groups [29,35,38]. The rationale is that siblings have similar home lives, socioeconomic backgrounds, genetics, and parental influence; therefore matching to siblings can, in theory, control for multiple factors. The downside of this type of matching is that siblings can also be affected by the stress of having a brother or sister with a chronic condition. To circumvent such issues some studies have, therefore, matched diabetic subjects to both non-diabetic siblings and non-diabetic peers [29,35,38]. McCarthy et al [35] used this approach, and a combination of the Iowa Test of Basic Skills (ITBS) and the Iowa Test of Educational Development (ITED) to assess academic ability but observed no significant difference in performance between those with diabetes and either control group. The authors did note, however, that poorer glycaemic control not only reduced performance of the affected child but that of their sibling too. The study was limited because formal school grades were not assessed and it was unclear how the ITBS and ITED correlated to other more well-known academic battery tests.

In the absence of formal school grades, parent[38], teacher [29] and individual [31,36,37] questionnaires have been used by several studies to investigate perceived academic progress; however reported feedback is likely to be subjective and may introduce bias. Parent et al [29] observed lower academic skills reported by teachers, especially in writing, for diabetic pupils when compared with siblings. However, children with learning difficulties or those using special education services were not identified which is significant given that children were rated based on expected levels for their age. A further weakness concerned lack of generalisability because the cohort only included urban and suburban schools. Whilst Robinson et al [37] reported no difference in the number of self-

reported GCSEs and A-levels obtained between 224 diabetic and 147 healthy children, the diabetic group experienced a significantly greater number of perceived health difficulties and problems at school which may have been attributable to lower self-esteem in those chronically ill pupils. Lloyd et al [36] also reported lower levels of self-perceived school competency in 40 children with diabetes; however the authors did not identify any actual differences in the number of vocational, ordinary, advanced, higher or university level awards attained [36] when compared to 40 matched peers. Both of these UK studies matched subjects by age, gender and, in the study by Lloyd et al, geography; however both were hindered by small samples and unadjusted univariate analyses despite gathering an array of useful information on parental confounders in their respective questionnaires.
Additionally, Lloyd et al allowed diabetic children to choose their own controls; a friend or relative of similar age and gender in the same geographical area. Therefore, the results may have been biased by children naturally associating with, and choosing, friends who had educational and employment prospects similar to their own.

Small sample size was a recurring theme within many studies [29,30,32,34,35] with the exception of Wennick et al [31] who observed lower self-reported levels of attained education for 106 diabetic subjects identified within a cohort of over 20,000 individuals. Attainment was defined as reaching: 9th year of compulsory school; upper secondary school; college or university. Diabetic children had 7% lower probability of reaching college or university and 7% higher probability of leaving school at the minimum age allowed compared with non-diabetic peers. The cohort of 19-38 year old subjects was obtained from an existing general population-wide cross-sectional health survey of the Swedish population and subjects were classed as diabetic if they self-reported that they had been diagnosed with diabetes before 19 years of age. The study was limited because diabetes diagnoses were based on self-report and, additionally, those not diagnosed before 19 years of age were classed as non-diabetic regardless of whether they were diagnosed between 20 and 38 years of age. Furthermore, the analyses were adjusted for marital status and whether parents were Scandinavian, but not for other important confounders such as other measures of individual and area socioeconomic status or levels of parental education.

Yu et al [33] assessed a range of academic outcomes for 31 children with early onset diabetes (younger than five years of age) compared with 35 subjects with later onset diabetes and 36 non diabetic controls. Outcomes included grade repetition, absence levels, general school experience and attainment; measured using the Peabody Individual Achievement Test (PIAT) and parental reports of grades in English, Maths, Social Studies and Science.

Children with diabetes had lower grades in English compared with non-diabetic children; however no differences in attainment were reported between those with early or late onset diabetes. Whilst parental reporting of some outcomes may be subjective, their direct reporting of school grades should be more accurate. Indeed the study demonstrated positive correlations between parental reports of attainment and PIAT test results. Whilst investigation of age at onset was an asset, the numbers in each group were small and the study participation rate was only 40%. Further, the main reason for non-participation was travel time to the clinic site; therefore the study may have been biased by including only those children worse affected or those parents more willing or able to travel.

Several of the studies discussed already also reported absenteeism associated with diabetes. As part of their wider study, Yu et al [33] reported more absenteeism for diabetic children and observed poor diabetes control to be a strong factor. Absences, however, were only reported by parents for the previous month and then extrapolated back by the authors over the previous year and so may have lacked accuracy. McCarthy et al [35] and Parent et al [29] reported that diabetic children had increased absence rates compared to peers and siblings based on school absence records. Ryan et al [30] reported that diabetic children missed significantly more school, and that school attendance was the strongest predictor of poor test performance. However, in cross sectional studies, reverse causation may operate whereby pupils are absent because they are already experiencing school difficulty. Absence rates were reported by mothers; however, parents of children with chronic conditions may have better recalled absence histories due to increased vigilance therefore potentially biasing results. Glaab et al [38] and an earlier pilot study by Vetiska et al [39] focussed entirely on absenteeism and demonstrated an increased absence rate for diabetic children compared with siblings and peers using data from school report cards. Glaab et al reported parental attitudes towards school attendance, poorer glycaemic control and shorter duration of diabetes to be associated with increased absenteeism. Diabetic children diagnosed within the last year were excluded to allow an adjustment period to illness and those with the poorest level of glycaemic control also dropped out during the study. Therefore, the most severely affected children with potentially higher absenteeism may not have been represented. The study by Glaab et al was large with almost 120,000 controls used; however, the number of diabetic cases and sibling controls was only 78 and 38 respectively. Vetiska et al [39] observed a close correlation between absence rates for siblings suggesting that family attitudes play a role in absenteeism and that parents of diabetic children may have a tendency to over protect. Unlike other studies, Vetiska et al studied

absence rate once days missed due to diabetic clinic appointments had been subtracted from the total days absent, and an excess still existed in the diabetic group.

To my knowledge, no published studies have exclusively investigated numbers of exclusions or enrolment in special education services for diabetic schoolchildren compared to peers. However, studies have reported adverse effects of diabetes on IQ [40,41] and specific areas of cognition such as spelling, reading and arithmetic [30,42-44], spatial and verbal intelligence [45], memory [40,46-49], and attention [46,47]. As the brain cannot store glucose it requires a continuous supply from the vascular system. Children with diabetes are, therefore, at risk of impaired cognitive function due to constant exposure to fluctuating and abnormal levels of insulin and glucose that impact brain function [46] with cognitive deficits in diabetic children reported across most domains [41,49]. Several meta-analyses have generalised these outcomes into broad domains with each made up of several sub-domains [41,49,50]. McCarthy et al [35] reported diabetic children to have more behavioural problems whilst Parent et al [29] reported increased attention difficulties. Increased grade repetition has also been observed [51,52].

Studies observing school leaver outcomes have reported reduced attainment levels [31], less likelihood to proceed to college or university [31,53], higher dropout rates and fewer years of schooling, less likelihood of full time employment [10,53], more absence from work and lower earnings [53] and self-reported difficulties in the workplace [36,37]. Fletcher et al [53] followed approximately 15,000 pupils over four time points and reported lower self-reported educational and employment outcomes in diabetic pupils.

2.1.3.2 Health outcomes

Diabetes is a chronic life-long condition and many of the health implications are observed later in life. Over many years, high blood glucose concentrations can cause damage to the eyes, kidneys, nerves, heart, and blood vessels and those with diabetes suffer increased morbidity and mortality and increased hospitalisation [19,20]. Much of the excess is attributed to long term renal and cardiovascular complications however an excess of hospitalisation and mortality are also observed in young type 1 diabetic patients with a short duration of illness [28]. A recent systematic review and meta-analysis by Morgan et al [54] identified thirteen population-wide studies observing mortality in patients diagnosed with diabetes before 19 years of age and reported an excess worldwide mortality in these subjects. A systematic review by Angus et al [55] evaluated literature assessing hospitalisation following

childhood diagnosis of diabetes (before 15 years of age) and demonstrated a three-fold risk of hospitalisation and a two-fold length of stay in hospital compared with children without diabetes. However, the majority of articles assessed specific diabetic related admissions and only four of the thirty-two studies reported all-cause admissions.

Most studies investigating mortality associated with child and adolescent diagnosed diabetes compared deaths within a diabetic cohort against expected mortality in the same age and sex categories of the general population and calculated standardised mortality ratios (SMRs). Patterson et al [56] combined population-wide diabetes registers from 12 countries and followed up almost 30,000 diabetic subjects diagnosed before the age of fifteen using record linkage methods and direct contact with doctors. The risk of death was two-fold for diabetic subjects across all countries and one third of the observed deaths were directly attributable to diabetes. However, country-specific SMRs varied and differences in follow-up methodology and accuracy and completeness of data between countries was evident. The absolute number of deaths did not differ between genders; however, the SMR was higher for diabetic girls due to their lower rate of death in the wider population.

SMRs reported elsewhere ranged from 2.2 [28] to 6.9 [57] and these studies observed higher SMRs in girls [28,57,58] or no differences by gender [59,60]. Contrary to findings by Patterson et al [56], declining SMRs over time have been observed and attributed to improvements in diabetic healthcare [57,59,61]. Secrest et al [57] observed a seven-fold risk of death within a Pennsylvania diabetic cohort compared to the wider population and attributed this to US citizens having less access to free healthcare. However, analyses were adjusted for age, race and gender but not for other clinical or socioeconomic variables and 3% of the diabetic cohort were lost to follow up. O'Grady et al [58] observed a three-fold risk of death within 17,000 people diagnosed with childhood diabetes compared to the general population and reported that the risk was greater for girls and those aged between 10 and 14 years of age. Whilst three quarters of deaths were directly related to diabetes, the numbers of non-diabetes related deaths were similar to that found within the general population. Gagnum et al [59] did not investigate causes of death but reported a 3.6 fold risk of all-cause mortality for those diagnosed with diabetes in childhood compared to the general population. Fewer deaths were observed in children most recently diagnosed; however, these children had a shorter follow up period and this may have biased the results.

Cause of death has been assessed by other investigators [28,60]. Dalquist et al [28] demonstrated a two-fold risk of mortality for children with diabetes compared to those without and a greater risk of mortality for children aged between 10 and 14 years of age. The authors reported that children with diabetes were at particular risk of unexplained death during sleep. Skrivarhaug et al [60] observed that, in childhood diagnosed diabetic people, acute metabolic complications accounted for most deaths under 30 years of age and cardiovascular disease accounted for most deaths over 30 years of age.

Icks et al [62] compared a diabetic cohort with the general population rather than a comparison group. The investigators compared hospital admissions for 5,784 children and adolescents younger than 20 years of age who were treated for diabetes in Germany in 1997 with deaths in the same age group in the German general population. Diabetic subjects had a three-fold risk of hospital admission and stayed in hospital three times longer than their non-diabetic peers. Additionally, the cost of their hospital care amounted to 1% of the total amount spent on all hospital care within the same age group. The study included cases from 61 diabetes clinics covering approximately 25% of all diabetic people younger than 20 years of age in Germany; however, this clinic sample may not have been representative of people with diabetes in the general community. Actual costs of healthcare were not available and were estimated using mean costs per hospital day in Germany.

Sayers et al [63] matched 1,577 Welsh children diagnosed with diabetes before 15 years of age against 7,800 unaffected individuals from the Welsh general population. Each diabetic child was matched to five non-diabetic children by age, gender, deprivation and county and the risk of follow up hospital admission, identified via record linkage, was six-fold for those with diabetes. The risk was highest in those diagnosed before the age of five and increased with deprivation. This study was nationally representative and utilised multilevel modelling to account for clustering at population level. By investigating all-cause admissions, these studies provided an insight into the total excess hospital care rather than only that attributable to diabetes related causes.

2.1.3.3 Summary

Schoolchildren with diabetes have added stresses compared to their peers when trying to perform well at school. On top of the physical burden of fluctuating blood sugar, and the mental burden of living with a chronic condition, they must inject insulin daily and vigilantly monitor their diet, level of exercise and blood glucose concentrations. Whilst

some previous studies have reported poorer academic performance compared to peers or siblings [9,10,29-33], others have reported no difference [34-37]. Previous studies have also reported increased school absenteeism [29,30,33,35,38,39] and poorer educational and employment outcomes for diabetic children after leaving school. No previous studies comparing diabetic children to peers have investigated school exclusion or special educational need. Increased all-cause admission [62,63] and all-cause mortality [28,56-60] for diabetic children compared to the general population have been reported.

Only two population-wide studies investigated educational outcomes [9,10] whilst others matched a small sample of diabetic children to one or more comparison cohorts [29,30,32-37,52]. Furthermore all but two studies [9,10] recruited diabetes cases from hospital or research clinics which may not represent the wider population of diabetic individuals within the community. The population-wide studies focussed exclusively on academic attainment and none analysed the range of educational and health outcomes explored in this thesis. Only one study adjusted for maternal confounders [9] and none adjusted for as wide a range of obstetric factors considered in this thesis. Only two small sample studies have been undertaken in the UK [36,37]. Whilst studies investigating health outcomes generally used larger cohorts, the majority contained limitations discussed previously.

2.2 Asthma

Asthma is a condition where airways become chronically inflamed and irritated resulting in recurring bouts of wheezing, breathlessness, coughing and tightness in the chest. Asthma attacks occur when the airways narrow and obstruct airflow often in response to contact with irritants or allergens [64]. Risk factors for onset of asthma include decreased lung function in infancy or presence of allergies, infection, antibiotic use, age, gender and socioeconomic status. Common triggers for an attack include exercise and exposure to irritants such as car fumes, pollen, dust, weather, animals or tobacco smoke. Prenatal risk factors include maternal stress, smoking, diet and use of antibiotics [65,66].

2.2.1 Prevalence

Asthma is one of the most common chronic conditions and is believed to affect around 300 million people worldwide. Prevalence is thought to be 16.1% in the UK and 18.4% in Scotland [67]. Other studies have reported UK prevalence of approximately 12,000 people per 100,000 with increased prevalence in females and highest prevalence in people aged 16-20 years (20,000 per 100,000 people). Most new cases are diagnosed in children under the age of 5 years [68]. Asthma prevalence reported in school aged children worldwide has ranged between 2% and 20% [69-77]; however, methodology used to diagnose asthma varies greatly which accounts for the observed differences.

2.2.2 Treatment

The treatment protocol for asthma is well described in the British Thoracic Society (BTS) /NICE Guidelines [67]. The pharmacological treatment of asthma is determined by symptom severity and moves through six pharmacological steps with each step indicating more severe and specific symptomatology. Step one requires intermittent use of a short acting beta agonist (SABA) for the treatment of mild asthma or wheeze whereas step 2 requires regular use of a SABA and an inhaled corticosteroid (ICS) thus indicating a more severe and more specific asthma. The BTS/NICE guidelines are implemented throughout the UK in both primary and secondary care, and so prescription of asthma treatment is routinely used as a surrogate marker for the presence of asthma and for estimation of asthma severity. By identifying all individuals prescribed step 2 and beyond medication it is possible to identify all children with a likely diagnosis of asthma with symptoms of sufficient severity to be troublesome.

2.2.3 Contextual review of the literature

A Medline literature search using the search terms in Table 2 returned 1,302 articles of which 220 (47 education and 173 health) remained after review of title and 38 (30 education and 8 health) remained after review of abstract. After full text review, 16 education papers and 2 health papers remained relevant to my thesis. The vast majority of these studies were undertaken in the USA and the remainder were from the UK, New Zealand, Australia and Greece.

Table 2 Search terms for asthma literature review

(((asthma*)
AND
(college or youth or juvenile or young or schol* or adolescent* or child* or student* or pupil* or school)
AND
(school* or education* or attainment or academic or employment* or hospital* or death* or mortality or "learning difficulty" or "learning difficulties" or "learning disability" or "learning disabilities" or exclu* or "special need" or "special needs" or achievement or expulsion or suspension or detention or GPA or grade* or SEN or "assisted support needs" or attendance* or absen* or injur* or leav* or admission*))
NOT
(trial* or intervention* or cross-sectional or "college hospital" or "college teaching hospital" or "absence of")).ti.

**Filters applied: Title, Age 0-18 years, English language

2.2.3.1 School outcomes

Taras et al [12] conducted a review of 78 previous studies reporting the impact of asthma and various asthma related interventions on school absenteeism and academic or cognitive ability. Almost all studies evaluating school absence observed greater absenteeism for children with asthma. Around two thirds of studies assessing school performance reported no difference in academic outcomes of asthmatic children compared with non-asthmatic peers. The authors, however, noted varying rates of absenteeism and asthma prevalence across studies due to different definitions and methodology. For example, whilst some studies defined asthmatic children based on parental report, others used clinical diagnoses. The review included studies regardless of whether investigators had used a comparison group and several did not compare children with asthma against non-asthmatic peers. For this reason, my literature review conducted for this thesis identified a smaller number of articles.

Most of the articles identified in my review jointly reported both absenteeism and academic outcomes. With the exception of one study [78], the majority observed more absence among schoolchildren with asthma. Associations with educational achievement were more conflicting as some studies reported that children with asthma performed worse than healthy peers [70,79] and others [80,81] observed no difference. The majority of studies reported relatively small sample size with some exceptions. Moonie et al observed absenteeism and attainment of American children with asthma across three separate studies [79,82,83] between 2006 and 2010. The 2006 study analysed 9,014 elementary, middle and high schoolchildren and reported 9.7% asthma prevalence and 1.3 additional absence days per year for asthmatic children compared with non-asthmatic peers (9.2 days versus 7.9 days) [83] after adjusting for gender, race, grade level and days enrolled. A subset of the same sample in 2008 only included pupils who had completed the Missouri Assessment Program (MAP) standardised test during the sampling year [82]. Ten percent of the sample, (397 of 3,812 pupils), were asthmatic and, unsurprisingly, children with asthma again experienced increased absenteeism compared to non-asthmatic peers (10.2 days versus 8.8 days). However, there were no observed differences in standardised test scores despite children with asthma missing more school. Both of these studies sampled children from a predominately African American district of low socioeconomic status; therefore, generalisation to other populations may have been limited. Absence days were obtained directly from school records; however, asthma was diagnosed by a school nurse with no formal clinical examination. Therefore some children may have been misdiagnosed. Finally, analyses were not adjusted for parental or pregnancy factors.

In 2010 Moonie et al [79] reported school recorded absence and academic grade point average using a much larger sample of 300,881 secondary school pupils across 372 schools. The study compared 206,736 children with no chronic conditions against: 15,084 children with asthma; 12,215 children with asthma and another chronic condition; and 66,846 children with a chronic condition but without asthma. Absenteeism of greater than ten days per year was specifically assessed. Children with asthma (adjusted OR 1.5) or asthma and another comorbid chronic condition (adjusted OR 1.6) had significantly increased odds of experiencing yearly absence totalling greater than ten days compared to non-asthmatic peers. However, non-asthmatic children with another chronic condition experienced no increased odds compared to their non-asthmatic peers. These findings via indirect comparison between asthmatic pupils and those with another chronic condition may dispel the hypothesis that living with any chronic condition, regardless of type, can negatively impact school outcomes and increase absenteeism [7,8]. Children with asthma also

had the lowest grade point average when compared with all of the other groups. Whilst it may be surprising that children with asthma received a significantly lower mean grade point average than those with asthma and a comorbid chronic condition, the differences across all groups were small in magnitude and several study design issues were apparent. Whilst absenteeism was analysed and adjusted for gender, ethnicity and grade level, these covariates did not appear to be included when analysing grade point average. Despite a negative correlation between absenteeism and grade point average, and differing absence rates within each chronic disease group, absenteeism was not included as a mediator. Other chronic conditions included cancer, seizure disorders, diabetes, genetic conditions, and hearing or visual impairment; however a definitive list was not provided and these chronic conditions were grouped together despite differing in type and severity. Therefore the disease specific impact on academic outcomes may have varied greatly for children with one of these other chronic conditions. Finally, all of the chronic conditions were reported by pupils or their parents at the start of the school year but not corroborated.

Kohen et al [70] also compared children with asthma to non-asthmatic peers with and without another chronic condition and classified asthmatic children into mild, moderate and severe groups. The study sampled 8,914 children aged 7 to 15 years from Canada's national longitudinal survey of children and youth (NLSCY) and reported 16% asthma prevalence. Compared to children with no chronic conditions, children with asthma had greater absence, increased use of education services, poorer standardised test scores and poorer maternal ratings of school performance. In particular, children with severe asthma had significantly increased odds (OR 3.53) of missing more than a week of school per year compared to children without a chronic condition. Non-asthmatic children with another chronic condition also experienced increased odds (OR 2.05); however, the effect size was smaller than that found in the severe asthma group. Therefore, consistent with the findings by Moonie et al (2010) [79], asthma had a greater effect on absenteeism than other chronic conditions. Introducing absenteeism as a mediator slightly attenuated the association between asthma and maths scores but not reading scores or mother rated school performance. Confounders included gender, age, household income, maternal age, maternal education, maternal employment status, and single parent household; however analyses were not adjusted for other potentially important maternal factors. Unlike Moonie et al (2010) [79], this study did not identify children with asthma and a comorbid chronic condition and this was a more notable limitation. Some asthmatic children may have had additional comorbid chronic conditions which could have inflated the observed effect size

within the asthma group. Finally, absences were recorded early in the school year and therefore numbers of missed days may have been underestimated.

Not all authors that have identified children with other chronic conditions have used them in their analyses. Fowler et al [84] randomly sampled 10,362 children aged 5 to 17 years from the National Health Interview Survey on Child Health (NHIS-CH) and identified 4.9% asthma prevalence based on parental reporting of symptoms within the last 12 months. Children without asthma who had other chronic illnesses such as epilepsy, congenital heart disease and less severe conditions such as skin allergies or frequent headaches were identified and excluded from the analyses. Children with asthma had a higher mean number of days absent (7.6 days versus 2.5 days) and a 1.7 times higher odds of learning disability compared to the healthy peer group but there were no significant differences in grade failure or exclusion from school. Adjusted confounders included gender, age, race and maternal education level and the association between asthma and learning difficulty was attenuated after further adjusting for absenteeism. This suggested that increased absenteeism was an important contributing factor behind asthmatic pupils experiencing difficulty at school. The study identified and excluded non-asthmatic children with other chronic conditions; however children with asthma may also have had other chronic conditions and these children were not identified. Comorbid conditions may have independently contributed to increased absenteeism and learning difficulty; therefore this limitation in study design may have overestimated the effect sizes in the asthma group. Similar to Kohen et al, this study also relied on maternal report of asthma with no clinical diagnoses; therefore, asthma prevalence may have lacked accuracy or included less severe conditions such as wheeze. All of the academic outcomes were reported by parents. However, whilst asthmatic children were only diagnosed if symptoms occurred in the last year, exclusions, learning difficulty and grade repetition were reported if they had ever occurred. Therefore it is possible that not all of these events post-dated the asthma diagnosis.

Other studies investigating both absenteeism and academic outcomes include those by Tsakiris et al [73], Krenitsky-Korn et al [85] and Silverstein et al [81]. Tsakiris et al sampled 1,539 pupils within 98 schools across 3 regions in Greece and observed 17% asthma prevalence based on parental questionnaire responses as to whether their child had received asthma medication within the last 2 years. The authors assessed grade point average across the last two years of elementary and middle school, number of days absent over the last 2 years, as recorded by parents and teachers, and independent parent and teacher reports of academic

performance graded as excellent, very good, good or bad. Children with asthma had a higher mean number of days absent compared with non-asthmatic peers (6.2 days versus 0.3 days); however, whilst children with asthma had poorer teacher rated performance, there was no significant association using parental ratings. Asthmatic children also experienced significantly lower grade point average in elementary but not middle school. The analyses were adjusted for age, gender, birth weight, gestational age, parental smoking status, education level and occupation of parents, and family size. Adjusting for gestational age and birthweight was a particular strength because, whilst obstetric factors can influence educational outcomes [16,17], they are rarely adjusted for in the literature. However, the authors did not consider absenteeism as a mediator for academic performance and this was a notable limitation.

Silverstein et al [81] compared 92 asthmatic children with 92 age-sex matched controls and reported differences in school recorded absenteeism but no difference in school recorded grade point average, class rank, grade promotion or a range of standardised academic achievement tests measuring ability in maths, reading and language. Asthma was diagnosed via linkage to historic hospital data. Whilst this method was robust, these children may have had more severe asthma compared to children treated solely in the community. Other limitations were apparent in addition to the small sample size. The sample was drawn from a suburban population containing few pupils of low socioeconomic status or ethnic minority thus limiting generalisability. Whilst absences were recorded every school term, and pupils contributed a mean of 6.6 school years of data to the analyses, repeated measures were no taken into account in order to correct for correlations between measures repeated on the same pupil. The analyses were adjusted for gender, grade and decade of school attendance, and attainment outcomes were further adjusted for absenteeism; however, the study did not provide details of specific multivariate analyses performed. Finally, it was unclear whether some of the outcomes analysed pre dated asthma diagnoses.

Krenitsky–Korn et al [85] reported increased absence and poorer grades in maths but not English for 28 pupils with asthma compared with 29 sex-matched pupils without asthma; however the authors used a very small sample. Associations between asthma and pupil's attitudes towards absenteeism, school health services and general support were also investigated. However, the analyses of absenteeism and attainment did not appear to be adjusted for any confounders.

Lindgren et al [80] and Liberty et al [71] both reported academic standardised test scores but not absenteeism. Lindgren at al [80] reported no significant differences in test scores between 255 pupils, treated at an asthma clinic, and their non-asthmatic siblings who were not medicated and were within four years of age. Academic outcomes were assessed using the Iowa Test of Basic Skills (ITBS) for children in grades 1-8 and the Iowa Test of Educational Development (ITED) for children in grades 9-12. Matching cases to sibling controls has the unique advantage of controlling for family background and even some genetic factors. For example, family income, socioeconomic status, parental education and employment, parental influence, parenting style, home experiences and neighbourhood characteristics are common between young siblings living together. However, absenteeism was not explored as a potential mediator and siblings were not matched by age or gender resulting in an excess of boys in the asthma group which could have biased the results. It was unclear whether these were adjusted in the resulting analyses. Lack of power due to small sample size may have contributed to the non-significant findings. Finally, the study analysed a sample of children attending an asthma clinic who may have had more severe asthma than those in the wider population.

Liberty et al [71] identified 18.5% (55 out of 298 children) asthma prevalence and measured academic achievement across two time points in a longitudinal design using the Wechsler Individual Achievement Test (WIAT) to explore baseline school readiness and achievement one year later. Asthmatic children lagged 6 months behind non-asthmatic peers in reading but not maths ability one year after starting school after adjusting for the potential mediating effect of absenteeism, and potential confounders including gender, socioeconomic status, initial academic ability, ethnicity and whether the child came from a single parent family. The sample was small and limited to children in their first year of primary school; therefore, these findings cannot be generalised to older children. Other studies reporting no difference in academic outcomes between children with asthma and healthy peers have included Rietveld et al [86] and Roder et al [87]; however, both of these studies used small samples.

Bonilla et al [88], Collins et al [69] and Millard et al [78] all reported associations between asthma and absenteeism. Bonilla et al [88] classified 1,490 children as known asthma if a parent reported past symptoms and the child had a previous clinical diagnosis, high probability asthma (HPA) if symptoms were reported but no previous clinical diagnosis had been made, or low probability asthma (LPA) if less severe symptoms were reported and no previous clinical diagnosis had been made. Those with known asthma missed two

additional days of school over the previous year compared with children in the other two groups and approximately 1.4 additional days were specifically attributable to respiratory issues. Investigating absences specifically due to respiratory symptoms set this study apart from others and was a noteworthy strength. Absences were obtained directly from school records; however, reasons for absence were initially reported to schools by parents and therefore had potential to be inaccurate. Further, generalisability was limited because the cohort contained mainly Hispanic subjects of low socioeconomic status. Finally, the survey response was only 34% and therefore the majority of the original sample was lost introducing the potential for further bias based on the demographics of responders and non-responders to the survey. For example, it might be the case that more proactive parents who took more of an interest in their child's wellbeing responded to the questionnaire. These same parents may also have been more proactive in ensuring that their child attended school regardless of asthma status, thus decreasing the effect size between groups.

Collins et al [69] conducted a telephone interview in households that contained at least one child aged 5-15 years. Of the 4,611 parents contacted, 18.6% reported their child to have asthma and these children were more likely to have missed at least one day of school in the previous month when compared with their non-asthmatic peers. Adjusted covariates included family income and child specific variables such as whether they had received any mental health treatment, were happy at school or whether they had friends to play with. However, analyses were not adjusted for age, ethnicity or socioeconomic status and this was a limitation. Parents who were ex-directory would not have been contactable for inclusion and this may have been a source of selection bias. Parents may also have reported asthma when in fact their child was suffering from coughing or wheezing which may explain the high percentage of asthmatic children identified in the study. Finally, absence was analysed as no days absent versus one or more days absent in the previous month. This provided a less accurate measure of absenteeism than the actual numbers of days absent over a longer duration.

Contrary to other studies, Millard et al [78] did not observe any difference in absenteeism between children with and without asthma on univariate analyses; however some study design issues were apparent. Pupils in 4^{th} to 6^{th} grade across 19 schools, who were thought to have asthma, returned a questionnaire agreeing to an exercise challenge in order to formally diagnose the condition. However only 31% of 3,100 contacted pupils returned the questionnaire and only 80% of eligible children subsequently accepted further testing. This may have resulted in responder bias because the most severely affected asthmatic pupils

may have been less likely to accept an exercise challenge. Compared with overall absence rates across the 19 schools included in the project, and to overall absence rates of all 4th to 6th grade pupils throughout the district, children in both groups were reported to have a slightly lower rate of absence. It is also possible, therefore, that responders were generally more conscientious pupils who valued academic performance, class attendance and being informed about their health and had less absenteeism.

Although no articles specifically assessing the risk of special educational need or exclusion from school for asthmatic children compared to peers were identified, Fowler et al [84] and Kohen et al [70] reported that children with asthma experienced increased use of educational services and increased learning difficulties compared to peers. No differences have been reported for grade repetition [80], grade failure [84], or risk of exclusion from school [84]; however, a variety of other difficulties have been observed. For example, Bonilla et al [88] demonstrated increased levels of distress and anxiety in children with asthma whilst Carpentier et al [89] observed poorer wellbeing, lower levels of happiness and increased mental health issues. Poorer health [70] and less participation in activities [70,85] have also been reported

Previous studies investigating school leaver outcomes include Carpentier et al [89] and Sibbald et al [90]. Carpentier et al [89] reported increased absenteeism but no difference in grade point average for 121 university students between the ages of 18 and 22 years and diagnosed with asthma in childhood compared with 121 age-sex matched non-asthmatic students. However, the results were limited because the analyses were not adjusted for any potential confounders. Additionally, the study may have experienced selection bias by only including higher achieving children who had entered university. These children may have had less severe asthma compared to the asthmatic children in the general population. Sibbald et al [90] reported an increased risk of unemployment at age 23 years for individuals with either current or previous asthma or wheeze compared to peers without. The previous asthma group included people who had not suffered any symptoms within the previous twelve months. Analyses were adjusted for gender, paternal social class, birth region and the subjects' level of education which was particularly important because level of education strongly influences a person's likelihood of employment. However, two main limitations were apparent. Firstly, this study was not limited to people with asthma as it also included those diagnosed with wheeze. Secondly, the final asthma assessment occurred at age 23, therefore some subjects may have been diagnosed with asthma after the age of 18 years. Contrary to these findings, Mazurek et al [72] reported children diagnosed

with asthma in childhood to have slightly higher levels of education but observed them to be more likely to work in particular types of occupation.

2.2.3.2 Health outcomes

Children with asthma can experience adverse health outcomes [91-93] and there is a wealth of literature describing temporal trends and regional variability in asthma hospitalisation and mortality [94-98]. Several studies have investigated asthma diagnoses within hospitalised or deceased cohorts. Others have assessed risk of admission or death within an asthma only cohort with no comparison to healthy peers. Very few studies, however, have reported on risk of subsequent hospitalisation or mortality for children with asthma compared to non-asthmatic peers [99 100]. Schwebel et al [99] obtained data on 878 children aged 4.5 years of age (54 months) and assessed subsequent risk of injury admission 2.5 years later. Sixty-seven children were diagnosed with asthma and they experienced an increased risk of parent reported injury compared to children without an asthma diagnosis. Outcomes were modelled using Poisson regression adjusted for a range of confounders including gender, delinquency or hyperactivity, home environment, household smoking, family income and maternal depression. In contrast with previous cross-sectional studies, Schwebel et al used a longitudinal study. This was a notable strength because it ensured that all injuries occurred after diagnosis of asthma. However, the study only included children aged 4.5 years old, used a relatively short follow up period, and did not assess type of injury or severity of asthma.

Kuo et al [100] investigated all-cause mortality and suicide in a 12 year follow up study comparing children with self and parent reported current or previous asthma against peers who had never had asthma. The asthma status of over 160,000 pupils aged between 11 and 16 years was assessed via parental and self-report. Children were classed as having previous asthma if they had not experienced any symptoms within the previous year. The investigators reported no significant difference in risk of all-cause death or death due to natural causes between any of the three groups in the follow up period. However, a significantly increased risk of suicide was reported for children with asthma compared to children without asthma after adjusting for age, gender, allergic rhinitis and smoking status of individuals and their families. The risk increased with asthma severity; however, other confounders such as socioeconomic status and ethnicity did not appear to be considered. An increased risk was also observed for children with previous asthma however this disappeared after adjustment for confounders. The authors used a novel approach to pupil

self-reporting. Having not previously discussed asthma, children were asked to watch five separate video clips of an asthmatic person suffering a particular symptom of asthma and they were then asked if they had ever experienced this symptom themselves. Students who answered yes to any of the five questions were classed as asthmatic. Using visual methods rather than more complicated written descriptions may have enabled students to more easily identify their symptoms and therefore provide more accurate answers. The study further included baseline levels of depression within the three groups and this did not attenuate any of the observed results.

2.2.3.3 Summary

Most of the previous studies investigating outcomes for asthmatic children compared to peers have reported absenteeism [69,78,88], attainment [71,80] or both outcomes [70,73,79,81,82,84,85]. All but one [78] of the previous comparable studies reported children with asthma to miss more days of school compared to peers [69,70,73,79,81-85,88]. Whilst some previous studies have reported poorer attainment for asthmatic children [70,79], others have reported no difference in performance compared to healthy peers [80-82,85]. To my knowledge, no studies have exclusively investigated SEN or school exclusion for asthmatic children compared to peers. However, limited research has reported increased use of special education services [70], increased learning disability [84] and even increased risk of school exclusion in asthmatic pupils [84]. Reports, from a limited number of studies, on whether or not asthmatic pupils have poorer leaver outcomes are mixed [72,89,90]. No studies exploring all-cause hospital admission were identified. However, increased injury admission [99] and mortality due to suicide [100] have been reported. To my knowledge, no studies have previously investigated any of these outcomes for asthmatic children compared to peers on a national scale either worldwide or in the UK. Only one adjusted for maternal confounders [73] and some [70,71] but not all [73,80] have explored absenteeism as a possible mediator. Previous studies have encountered limitations including small sample size [71,78,80,81,85,89] and limitations in case ascertainment [69,78,81,84].

2.3 Epilepsy

Epilepsy is a relatively common neurological condition whereby abnormal brain electrical activity predisposes the person to recurrent seizures [101]. The brain constantly experiences normal electrical activity; however, sudden bursts of intense electrical activity can result in seizure [102]. Isolated seizures can occur for many reasons; therefore epilepsy is normally only diagnosed if at least two unprovoked seizures have occurred in the past and more are likely to occur [101]. Epilepsy is not a singular condition. Over 40 different types of epilepsy have been defined based on their varying symptoms and underlying causes and these have been classed into clinically distinct groups. Furthermore, individuals can suffer from more than one seizure type [103]. Seizures can occur in two broad categories: generalised seizures affect large parts of both sides of the brain often resulting in loss of consciousness and partial (focal) seizures affect localised regions of the brain. Generalised seizures include absence, tonic-clonic, tonic or atonic, and myoclonic seizures [102,104,105]. Status epilepticus, where a seizure lasts for more than thirty minutes, is the most serious type and requires medical attention [104]. Risk factors for epilepsy include brain injury or trauma due to severe head injury, stroke, encephalitis or oxygen deprivation occurring, for example, due to asphyxia at the time of birth. Other possible causes are tumours in the brain, chemical or hormone imbalances, or conditions such as cerebral palsy [102,104]. Epilepsy is referred to as symptomatic epilepsy when the underlying causes are known whereas idiopathic epilepsy describes epilepsy where the cause is not understood.

2.3.1 Prevalence

Epilepsy is most commonly diagnosed in childhood or over the age of 65 [106]. Approximately 1% of children and adolescents are thought to have epilepsy [106,107]. In developed countries worldwide, the prevalence ranges between 3.2 and 5.5 per 1,000 children and incidence is consistently higher in the first year of life [108]. At present approximately 600,000 people in the UK have epilepsy and are prescribed anti-epileptic drugs, equating to 1 person in every 103. Approximately 63,000 children under the age of 18 in the UK have epilepsy; roughly 1 in every 220 children. In Scotland, 54,000 are affected equating to 1 person in every 97 [106].

2.3.2 Treatment

Around two thirds of people with epilepsy can control their seizures via use of anti-epileptic drugs (AEDs). AEDs do not cure epilepsy or its underlying causes; however, they stop seizures from occurring by reducing the excess electrical activity in the brain. Alternative treatments include surgery, vagal nerve stimulation and a ketogenic diet [109].

2.3.3 Contextual review of the literature

A Medline search using the search terms in Table 3 returned 355 articles of which 110 (74 education and 36 health) remained after review of title and 57 (30 education and 27 health) remained after review of abstract. After full text review, 18 education papers and 18 health papers remained relevant to my thesis. These studies were from Nigeria, Holland, Japan, Brazil, USA, Indian and Austria.

Table 3 Search terms for epilepsy literature review

(((Epilep*)
AND
(college or youth or juvenile or young or schol* or adolescent* or child* or student* or pupil* or school)
AND
(school* or education* or attainment or academic or employment* or hospital* or death* or mortality or "learning difficulty" or "learning difficulties" or "learning disability" or "learning disabilities" or exclu* or "special need" or "special needs" or achievement or expulsion or suspension or detention or GPA or grade* or SEN or "assisted support needs" or attendance* or absen* or injur* or leav* or admission*))
NOT
(trial* or intervention* or cross-sectional or "college hospital" or "college teaching hospital" or "absence of")).ti.

**Filters applied: Title, Age 0-18 years, English language

2.3.3.1 School outcomes

Reilly et al [11] conducted a review of the literature investigating low academic achievement or underachievement and identified fifteen studies all of which used small clinic based samples. These compared performance of epileptic children with one or more control group or population mean scores and two also explored underachievement based on IQ and cognitive ability [110,111]. Many studies used academic battery tests despite describing cognitive and behavioural outcomes in their title and the review highlighted a paucity of literature using officially recorded school grades. My review only included articles which

mentioned academic outcomes in their title and which compared epileptic children to one or more comparison groups.

Adewuya et al [112] identified 73 epileptic children and reported poorer mean school recorded grades across all subjects, and specifically in maths and English, over the previous school year compared with 82 controls matched by age, sex, socioeconomic background and comorbid mental disorders. Matching groups together by mental disorders can theoretically control for the possibility of children with epilepsy having worse attainment due to increased prevalence of comorbid difficulties. However, the study did not expand on the nature of identified mental disorders or describe how groups were matched. Further, children with learning impairment who were still able to attend normal school were not identified or excluded and may have contributed to poorer performance in the epilepsy group. Number of days absent was obtained from school records; however, it was not analysed as an outcome or as a potential mediator to explore whether epileptic children had poorer grades due to greater absenteeism. Finally, the sample size was very small and the analyses were not adjusted for additional confounders such as parental or maternity factors or ethnicity.

Nuhu et al [113] and Singh et al [114] both investigated parent and adolescent reported academic performance. Nuhu et al reported poorer performance over the previous three school terms for 77 epileptic children compared with 76 age-sex matched controls. However, the analyses were not adjusted for any other confounders and the sample size was small. Performance was only reported as good or impaired whilst general absence levels were reported as regular or irregular. However, these labels may have been applied inconsistently and it was unclear whether these measures were clearly defined and originally self-reported in this format, or later dichotomised by the researchers. Further, poor performance was adjusted for condition specific variables and absenteeism in subsequent analyses within the epilepsy group; however, absenteeism was not compared or used as a mediator for performance between groups. Finally epilepsy was diagnosed based on eye witness accounts of two or more seizures occurring; therefore, diagnoses may have lacked a sound clinical basis.

Singh et al investigated academic performance, reported by parents and teachers as normal, failed or decreased, and attendance, reported by teachers as decreased or not. This study compared 100 children with epilepsy to 50 healthy siblings and 50 unrelated children with asthma. The rationale behind matching cases to siblings is to control for common shared

factors such as family background, home conditions or parental influence. Similarly, matching to a group with another chronic condition such as asthma can control for potential deficits in academic outcomes occurring due to the generic difficulties of dealing with chronic condition. Epileptic children had more educational problems compared to both control groups: increased absence, more school dropout, poorer performance across all subjects and more failed performance across one or more subjects. However, the groups were compared using univariate analysis with no adjustment for confounders and the authors did not confirm whether the different groups were matched. There appeared to be differences in the demographic profile of each group, for example an excess of boys in the epilepsy group compared to healthy controls, which could have biased the results. Finally, the study sampled from an urban area which limited generalisability to other settings. Austin et al [115] also compared academic achievement between children with epilepsy and a control group of children with asthma and observed lower academic achievement in the epilepsy group. However, this study did not include a healthy comparison group without a chronic condition.

Singh et al [114] and Suurmeijer et al [116] limited their epileptic cohorts to children currently in receipt of anti-epileptic medication. Suurmeijer et al assessed the educational outcomes of 109 epileptic children compared with 109 controls at 11 and 16 years of age via parental questionnaires and observed poorer academic outcomes in the epileptic group. Whilst the analysis was not adjusted for confounders, the groups were matched on a number of variables: gender, age, socioeconomic status, degree of urbanisation of family residence, family size and the ordinal position of the child within the family.

Whilst most studies have reported worse academic attainment for children with epilepsy compared to controls, some have observed no corresponding deficits in IQ. Melbourne et al [117] reported that 33 children with epilepsy had poorer academic and cognitive test scores but similar IQ compared with 33 age-sex matched controls. Epileptic children achieved lower scores on cognitive tests of attention, language and memory and had poorer mathematical ability assessed via the Wide Range Achievement Test (WRAT). Analyses were adjusted for cognitive function, socioeconomic status and home environment and all pupils were selected from the same class which had the further advantage of controlling for any effects of classroom dynamics or class size. Despite this, the study was underpowered due to small sample size and it included children with various types of epilepsy making it hard to generalise to other studies. By comparison, Oliveira et al [118] and Miziara et al [119] observed outcomes for children with specifically rolandic epilepsy and benign childhood

epilepsy with centro-temporal spikes (BECTS) respectively. Oliviera et al observed similar IQ but poorer reading, maths and writing scores, measured on the Educational Achievement Test (EAT), for 19 epileptic children compared to 19 non-epileptic peers matched by age, sex and socioeconomic background. Epilepsy cases were identified via clinical examination and electroencephalogram (EEG); however univariate analyses were employed with no adjustment for additional confounders. This limitation was shared by Miziara et al [119] who reported that 40 epileptic children had lower reading, writing and maths scores, measured using the School Performance Test (SPT), compared with 40 controls matched by age, sex and school level. Additionally, the study was underpowered and utilised a clinic sample likely including more severe epilepsy cases not representative of epileptic children in the community.

Previous studies matched children with epilepsy to sibling controls. Dunn et al [120] assessed reading, maths and writing ability at baseline and 36 months after study entry for 219 children with epilepsy aged 6-14 years and 131 un-matched siblings aged 2-18 years using the Woodcock-Johnson (WJ) test of achievement. Similar reading and maths scores but poorer writing ability was reported for epileptic children at baseline whilst deficits were evident across all domains for children with epilepsy after 36 months. This study defined epileptic children if they had at least one seizure in the previous 3 months whereas the majority of other studies only diagnosed children experiencing at least two seizures. Therefore some of these children may only ever have experienced one seizure and may not have had epilepsy. It is also worth noting that, whilst sibling controls are useful, secondary emotional and physical effects of witnessing and aiding seizures of a loved one may be experienced by close family members including siblings which may adversely impact their outcomes. Berg et al [121] reported increased grade retention, more academic problems and increased placement in special education for 142 epileptic children compared with 142 siblings matched as closely as possible by age. However, the outcomes were not well defined and analyses were not adjusted for additional covariates post matching.

Contrary to the vast majority of literature, Ibekwe et al [122] reported no difference in academic performance between 50 epileptic cases and 50 controls matched by gender, age and socioeconomic status. Children with epilepsy had increased absence compared with their matched counterparts (mean 15.3 days per year versus mean 9.4 days). Academic performance was classed as low, average or high; however, the investigators did not explain how, or by whom, the data were collected or how the categories were defined.

Additionally, similar to other studies, these conclusions were based on univariate analyses with no adjustment for potential confounders.

No studies directly investigating absenteeism, school exclusion or school leaver prospects were identified in this review. However, higher rates of school dropout [114] and more school absenteeism [114,122] for epileptic pupils has been reported previously. Only Berg et al exclusively investigated special educational placement; however, epilepsy has been associated with cognitive impairment, increased learning difficulty [114,117,123], more behavioural problems [114,124,125], lower self-esteem [126,127], reduced psychosocial function [128] and comorbid conditions such as ADHD, conduct disorder, depression and anxiety [114]. Studies have reported either lower IQ in children with epilepsy [122] or no significant difference compared with peers [117,118,129]. Whilst academic underachievement with respect to age and IQ has been observed [110,111,130,131], epileptic children more commonly experience low achievement due to reduced cognitive function [11,111] with deficits in attention [117,123,126,132], memory [117,123,124,133], language [117,123], psychomotor speed [124,132], dexterity [123,132], perception [123], verbal function [134], auditory processing [119], response inhibition [132] and executive function [119,133] all observed. Earlier seizure onset and comorbid ADHD can increase risk for some types of learning difficulty [110] and comorbid depression can lead to low school performance [135]. Children with epilepsy have been reported to have reduced intelligence [114,124], more grade retention [136], increased placement in special education classes [121,136] and increased occurrence of comorbid conditions such as depression, ADHD, anxiety and conduct disorder [114]. Melbourne et al [117] reported more learning difficulty but no increase in absenteeism or grade repetition. In order to exclude the most severely affected epileptic children with complex needs, previous studies have excluded children with cerebral palsy, sensory impairment, speech disorders, epileptic syndromes, low IQ or comorbid psychiatric disorders or neurological deficits [113,114].

2.3.3.2 Health outcomes

Children with epilepsy have poorer health outcomes than their peers and suffer an increased risk of mortality. Many previous studies have identified and analysed deaths attributed to epilepsy within a wider cohort of deceased subjects whilst some have specifically observed sudden unexpected deaths in epileptic patients (SUDEP). Whilst some of the previous studies investigating associations with hospital admission and mortality compared groups of epileptic and non-epileptic children [137-139], most compared epileptic children with the wider population [140-143].

Christensen et al [143] investigated mortality by 30 years of age within a Danish birth cohort of over 1.8 million people, over twenty five thousand of whom were diagnosed with epilepsy in childhood. Approximately eleven thousand people died in the follow up period including eight hundred epileptic subjects and this equated to a fifteen-fold risk of death among those with epilepsy compared to the general population. The mortality rate ratio (MRR) of 14.9 reduced to 4.2 after omitting children with adverse birth outcomes and comorbid neurological disorders. Mortality was higher in younger children, boys and for those diagnosed with epilepsy before the age of five. The main strength of this study was the very large population-wide cohort and the long follow up period. However, epileptic children were identified via hospital admissions; therefore, children treated in private clinics or those managed solely by a general practitioner would not have been identified and this was a limitation.

Ackers et al [140] used a nationwide UK database to identify 6,190 children prescribed anti-epileptic medication and 151 died during follow up which equated to a crude mortality rate (CMR) of 56.2 per 10,000 person years. The standardised mortality ratio (SMR) compared to the general population was 22.4 demonstrating a twenty-fold risk of death for children who had been treated for epilepsy. The prescribing database covered 430 general practices nationwide across England and Wales; however, this only represented 5% of the population. Furthermore the study only explored deaths occurring before the age of 18 years; therefore, deaths in adulthood were not counted.

Berg et al [144] and Harvey et al [145] were limited by inability to calculate SMRs in their respective studies. Berg et al combined cohorts across Canada, USA and Holland to identify 2,229 children experiencing two or more epileptic seizures before the age of seventeen and followed them for an average of 14 years. The combined CMR was 228 per 100,000 people and was greater for those with epilepsy complicated by underlying brain injury or significant intellectual disability compared to uncomplicated epilepsy (743 versus 36 per 100,000). However, comparable age-sex specific death rates in the general population could not be calculated because four different cohorts, across three different countries, comprising different ages and overlapping time periods, were merged together. Therefore, crude death rates were merely compared against those observed within cohorts of children younger than 30 years of age in the USA, Canada and Holland. Harvey et al identified 1,095 deaths before the age of 14 including 93 in children with epilepsy. Identifying epileptic children within a deceased cohort meant that the researchers could not calculate prevalence of epilepsy within the wider population. The MRR was 13.2 for

children with epilepsy compared to children without epilepsy. However, the measure lacked accuracy because the authors could not determine the number of epilepsy cases within the general population having only identified epilepsy deaths within a cohort of deceased children. Therefore, denominator rates of epilepsy in children younger than 14 years of age were merely estimated from five other studies.

Callenbach et al [141] and Camfield et al [142] both identified children who had two or more unprovoked seizures before the age of 16 and excluded those with only neonatal epilepsy or who had acute provoking factors for their seizures. Camfield et al assessed mortality outcomes of 692 children with epilepsy compared to the general population of Nova Scotia. Death rates in children with epilepsy were five times higher compared to an age-sex matched reference population in the 1980's and almost nine times higher compared to a similar reference population in the 1990's. However, on further analysis of epilepsy type, only those with a severe underlying neurological disorder were at an increased risk of death compared to the general population. This particular study undertook survival analysis and demonstrated that 6% of the epileptic children had died after 20 years compared to less than 1% of the reference population. These findings agree with two other studies [141,146]. Callenbach et al [141] reported a seven-fold risk of death in epileptic children compared to the general population. However, after splitting the population by epilepsy type, the study reported an SMR of 20 for children with symptomatic epilepsy and no deaths were identified for children with non-symptomatic epilepsy. Nickels et al [146] also only reported increased mortality in children with epilepsy where they had neurological complications. The authors compared 467 children with epilepsy, identified over a 30 year period and followed up for an average of 7 years, with the general paediatric population and found no increased risk of death for those with uncomplicated epilepsy. Silanpaa et al [147] used a particularly long follow up period of 40 years and reported a three-fold risk of death in children diagnosed with epilepsy before 16 years of age compared to age-sex adjusted mortality in the general population.

The majority of literature investigating hospital admissions focussed on admission due to various types of injury. To my knowledge, no studies have investigated risk of all-cause admission for children with epilepsy compared to peers. Kirsch et al [138] reported no increased risk in various types of injury for 25 children diagnosed with epilepsy between 5 and 16 years of age compared to 25 age-sex matched friends. This study differed from others because all of the children with epilepsy were classed as cognitively normal. However, this limited generalisation of the findings to other studies. The study also lacked

power which may have accounted for the lack of any observed association with specific type of injury admission. Finally, children in this study had epilepsy for at least a year; however all previous injuries were included and some of these may have predated the onset of epilepsy.

Some of these study limitations were also experienced by Baca et al [137]. The authors assessed injuries pertaining to 613 people nine years after diagnosis of epilepsy. Outcomes for those with complicated (133 people) and uncomplicated (368 people) epilepsy were compared; however, only 210 of those with uncomplicated epilepsy were matched and subsequently compared to non-epileptic siblings within 3 years of age. With the exception of head injury, children with uncomplicated epilepsy did not experience excess injury admission compared to siblings; however, the analyses did not appear to be adjusted for any confounders. Children were asked to recall injuries that required medical attention; therefore, recall bias may have occurred. Additionally, the recall did not distinguish between numbers of accidents and numbers of injuries. Finally, lifetime injuries were assessed for all children with epilepsy despite diagnoses occurring between 1 month and 15 years of age. Therefore, some of the injuries may have predated epilepsy diagnosis.

By comparison, Prasad [139] utilised a much larger study cohort and followed children until their first admission following diagnosis of epilepsy. Primary care data pertaining to almost 12 million people across 625 GP practices covering 8% of the UK population were analysed and people receiving more than two prescriptions for antiepileptic medication were defined to have epilepsy. The authors demonstrated greater risk of bone fracture (HR 1.23), thermal injury (HR 1.49) and medicinal poisoning (HR 2.5) for 11,934 people diagnosed with epilepsy between 1987 and 2009 compared to 46,598 peers matched by GP practice and five year age band. Furthermore, a greater risk of medicinal poisoning was observed for males and those in higher age groups. This study had several additional strengths. Participants with prevalent epilepsy or who experienced injury prior to diagnosis were omitted and analyses were adjusted for age, sex, region, deprivation and presence of comorbid ADHD, behavioural disorder, learning disability or cerebral palsy. Finally, participants were aged between 1 year and 24 years of age at diagnosis; however, the results remained after excluding people older than 21 at diagnosis and after using a tighter definition of epilepsy.

2.3.3.3 Summary

To the best of my knowledge, there have not been any previous population based studies investigating the impact of epilepsy on school outcomes. Furthermore no studies have been conducted in the UK. Indeed previous studies have used very small samples [114,117-120,122] and many have used clinic based samples [112,116,117,122] which may have included only the most severe cases. Most comparable studies have reported children with epilepsy to have poorer academic attainment compared to peers [112-120] Compared to the general population, several studies have also reported a strong association between epilepsy and risk of mortality [140-147]. To my knowledge, no studies have previously investigated leaver prospects, absenteeism, exclusion or all-cause hospital admission for children with epilepsy compared to peers. However, within small studies, increased absenteeism of epileptic pupils has been reported [114,122]. Only one study investigated SEN and reported increased risk for children with epilepsy [121]. However, other studies have observed a range of cognitive difficulties in epileptic children [114,117,123-127].

2.4 Attention deficit hyperactivity disorder

Attention deficit hyperactivity disorder (ADHD) is a common neuro-behavioural disorder among children and adolescents and is characterised by persistent and developmentally inappropriate levels of hyperactivity, impulsivity and inattention [148,149]. According to the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM-V), for a diagnosis to be made several inattentive or hyperactive-impulsive symptoms must be present before 12 years of age, have persisted for more than 6 months and occurred in two or more settings such as at home and at school. Symptoms must also interfere with development and negatively impact on social, academic or occupational functioning [150]. Three types of ADHD occur: inattentive type, where inattentive symptoms predominate over all others; hyperactive-impulsive type, where individuals present with predominately hyperactivity and impulsivity; and combined type, where all three symptoms play a role. ADHD can also be classed as mild, moderate or severe depending on the number and severity of symptoms present [148-150].

In the UK a diagnosis of ADHD must be made by a qualified psychiatrist/clinician, however there are no specific laboratory tests for this condition [151] with initial referral and final diagnosis often relying heavily on reports of behaviour patterns from teachers and parents. Initial signs can be more formally evaluated using behaviour rating scales and cognitive tests and physical and neurological examination should be assessed in a developmental context in relation to children of the same age and cognitive level [152]. ADHD can be classified using criteria based on either DSM or the International Classification of Diseases (ICD). Individuals with ADHD have an increased risk of experiencing comorbid disorders [153]. It can therefore present alone or alongside other psychiatric conditions such as depression, bipolar, anxiety or learning disorders, or other disruptive behaviour disorders such as oppositional defiant or conduct disorder [154]. Comorbidities including learning disability, Tourette's syndrome and borderline personality disorder have also been reported [153]. Many of these comorbid conditions result in similar behaviour patterns to ADHD and can therefore complicate diagnoses and even inflate estimated prevalence [151].

2.4.1 Prevalence

In 2013 the World Health Organisation estimated that ADHD affected about 39 million people [155]. However the prevalence of ADHD in published literature varies significantly due to documented methodological differences in case ascertainment [151,156,157]. Differences in diagnostic criteria, impairment criterion and source of information contribute to geographical heterogeneity of prevalence estimates, which have been shown to largely disappear after taking into account these methodological differences [156]. Skounti et al [157]attributed varying worldwide prevalence rates ranging from 2% to 17%, to differences in diagnostic methodology and estimated prevalence between 4% and 10% after taking these differences into account [157]. Meta-analyses following systematic reviews of the literature estimated pooled worldwide prevalence, using only DSM-IV criteria[158], to be 5.29%[159]; ranging from 5% to 7%. An updated review of 135 studies in 2014 [156] reported no increase in numbers of children being diagnosed with ADHD over the last three decades. Childhood prevalence in Scotland is around 5% but treated ADHD ranges between 0.2% and 1.2% [149]. Prevalence of ADHD is four times higher in boys than girls [152,160] and, whilst rends by socioeconomic status have not been widely reported, some studies have observed increased incidence in white children compared to other ethnic groups[161].

2.4.2 Treatment

ADHD can be treated using medication or cognitive behavioural therapy, or both. Though medication does not permanently cure ADHD it can aid concentration in affected individuals, reducing feelings of impulsivity and promoting calmness [162]. Medications licensed in the UK for treatment of ADHD at the time of writing include methylphenidate hydrochloride, dexamfetamine sulfate, lisdexamfetamine dimesylate, atomoxetine and guanfacine hydrochloride. Methylphenidate is the first drug of choice used almost exclusively to treat ADHD symptoms - others drugs may be added or used instead for children who do not respond to first line therapy. Between 1992 and 2008, ADHD medication use amongst children under 16 years of age increased 34 fold but has since stabilised [163]. Whilst the long term impact of ADHD medication is still not well understood, improvements in some health and social outcomes have been documented [162]. Medication has been reported to produce short term relief of symptoms with short term impact on academic outcomes, however whilst medication does improve academic outcomes, it does not completely bridge the attainment gap between children with and without ADHD [164,165]

2.4.3 Contextual review of the literature

A Medline search using the search terms in Table 4 returned 425 articles of which 141 (105 education and 36 health) remained after review of title and 44 (24 education and 20 health) remained after review of abstract. After full text review, 19 education papers and 9 health papers remained relevant to my thesis. The vast majority of these studies were undertaken in the USA with the remainder were from Sweden, Finland, Denmark, Taiwan, Greece, Israel and Holland.

Table 4 Search terms for attention deficit hyperactivity disorder literature review

(((ADHD or (attention and deficit and hyperactivity and disorder) or (attention and deficit and disorder) or "hyperkinetic disorder")
AND
(college or youth or juvenile or young or schol* or adolescent* or child* or student* or pupil* or school)
AND
(school* or education* or attainment or academic or employment* or hospital* or death* or mortality or "learning difficulty" or "learning difficulties" or "learning disability" or "learning disabilities" or exclu* or "special need" or "special needs" or achievement or expulsion or suspension or detention or GPA or grade* or SEN or "assisted support needs" or attendance* or absen* or injur* or leav* or admission*))
NOT
(trial* or intervention* or cross-sectional or "college hospital" or "college teaching hospital" or "absence of")).ti.

**Filters applied: Title, Age 0-18 years, English language

2.4.3.1 School outcomes

To the best of my knowledge no studies have investigated ADHD and educational outcomes on a national population-wide scale. Indeed most studies have compared a small number of children with ADHD against a small number of controls [166-170] and many of these have used clinic based samples [168,171]. A larger study by Barbaresi et al [172] followed 370 children with ADHD and 740 age and sex matched controls without ADHD from 5 through to 18 years of age to compare academic attainment, absenteeism, grade retention and school dropout. Children with ADHD were three times more likely to be retained a grade and 2.7 times more likely to drop out of high school before graduation compared with matched controls. These pupils also experienced increased school absence and lower reading test scores. The academic measure only comprised reading ability, tested using the California Achievement Test (CAT); however the other outcomes were obtained from school records gathered across several years' of follow up. Additional strengths included

the population-based sample and robust five-stage methodology used to identify children with ADHD by combining school and medical records, teacher and parent reports, clinical diagnoses and medication information. Comorbid learning disabilities and psychiatric disorders can often be the driving force behind poor outcomes; however these were not identified or investigated. Further, the predominantly white middle class sample limited generalisability and 43% of the original sample was lost to follow up.

Wu et al [173] compared pupil, parent and teacher questionnaire responses for 333 children with persistent ADHD,166 children with non-persistent ADHD and 266 children without ADHD. Children with ADHD had poorer attitudes towards schoolwork, poorer school interactions, more behavioural problems and lower overall and subject specific (English, maths, Chinese, social science and natural science) school grades which were reported as above average, average, below average or failing. The authors identified a clear gradient effect whereby children with persistent ADHD performed the worst. Questionnaire responses may be subjective, inaccurate or biased and should be treated with caution; however two things were of note in this study. Firstly, children with ADHD had poorer outcomes regardless of which questionnaire responses were used. Secondly, correlations between pupil and parental responses regarding academic attainment were high. Analyses were adjusted for age, sex, comorbidity and pupil IQ; however, the study encountered limitations. The predominately male sample limited generalisability and diagnoses of ADHD based solely on child and parental responses may have been inaccurate. ADHD was classed as non-persistent, for example, if a parental response indicated presence of ADHD whereas a child response did not or vice versa.

Only a few studies have used school grades to measure academic attainment. Ek et al [174] reported that Swedish children with ADHD had lower mean grades, based on sixteen different subjects, at the age of 16 years, compared with controls. They were less likely to qualify for secondary school by passing exams in English, Maths and Swedish and also underachieved in relation to their cognitive capacity – a finding also reported by De Shazo et al [168]. Powers et al [175] analysed school records and reported a lower grade point average and greater class retention for children with ADHD. Similar to Wu et al, these studies both used two comparison groups. Ek et al compared 39 children with ADHD to 80 with behavioural and learning problems and 417 healthy controls because children with ADHD often have comorbid behavioural and learning difficulties which can independently impact educational outcomes. Matching children with ADHD to an ADHD free group with these difficulties can help determine to what extent poorer outcomes are directly attributable to

ADHD. Unfortunately the authors did not report how the comparison groups were matched to children with ADHD. Powers et al compared 48 medicated and 42 non-medicated children with ADHD against 80 healthy controls and additionally measured attainment using the Weschler Individual Achievement Test (WIAT). The medicated group had better outcomes than the non-medicated group; however both ADHD groups had lower grades and greater class retention when compared to healthy controls. The results of this small study suggest that whilst medication may help to improve school outcomes it does not fully bridge the achievement gap. Furthermore, it is possible that the improved outcomes in the medicated group were due to greater parental involvement and proactivity in seeking treatment and care. This study did utilise an ethnically diverse school based sample and the mean length of follow up of 9.13 years was longer than has been reported in other studies.

Bussing et al [167], Faraone et al [176] and McConaughy et al [170] have also utilised more than one control group. Bussing et al [167] reported that 87 children with ADHD had a lower school recorded grade point average, poorer reading, writing and maths scores as measured on the Florida Comprehensive Assessment Test (FCAT), increased placement in special education classes and more class retention and were less likely to graduate compared with 112 healthy controls and 23 children with sub-clinical ADHD who had some ADHD symptoms but not enough for a full diagnosis. Compared to healthy peers, the observed poorer outcomes in the ADHD group remained after adjusting for gender, race and poverty. Children with ADHD had five-fold odds of school dropout and three-fold odds of grade retention. However, these associations disappeared after accounting for children's placements in special education. Faraone et al [176] also reported increased use of special education, more repeated grades and increased learning disability for children with ADHD; however the sample only included boys. The study compared 140 boys with ADHD to 120 controls and 303 siblings.

In the absence of school recorded grades, studies have assessed academic attainment using standardised academic battery tests. Whilst these provide insight into specific reading, writing and mathematical ability, school recorded grades have the advantage of providing a more practical measure of attainment drawing on a wider range of skills. McConaughy et al [170] matched 101 children diagnosed with ADHD to 24 healthy controls and 53 children with behaviour and learning problems who had been referred to an ADHD clinic but did not meet the criteria for a full diagnosis. Compared to both comparison groups, children with ADHD had poorer teacher and parent reported academic outcomes and poorer reading, writing and maths scores measured via the Weschler Individual Achievement Test

(WIAT). Similar to the study by Ek et al, including a comparison group with behavioural and learning difficulties demonstrated that the poorer performance of the ADHD cohort was not wholly attributable to these children having comorbid learning problems.

Barkley et al [166] used the Wide Range Achievement Test (WRAT) to assess school attainment whilst De Shazo et al [168] used the Mini Battery of Achievement (MBA) and both reported that children with ADHD had poorer academic attainment and increased special educational need compared with peers. Barkley et al [166] additionally reported that children with ADHD had increased grade retention and more behavioural problems. They were three times more likely to have comorbid oppositional defiant disorder and four times more likely to have comorbid conduct disorder compared with peers who were matched by age and IQ and by their parents' age, socioeconomic status and education level. Conclusions were however limited because the investigators only compared outcomes for children with and without ADHD regardless of comorbid disorders which may have independently influenced the poor outcomes observed.

De Shazo et al additionally reported that children with ADHD had lower IQ and increased learning disability and underachieved compared to their predicted levels after adjusting for IQ. The authors observed that poorer attainment remained after excluding children with learning difficulties thus demonstrating that deficits could not be wholly attributed to comorbid learning disability. Unfortunately, it was not clear whether the comparison group was matched or how information on learning difficulties and special education placements was obtained. Similar to Babaresi et al [172], these studies sampled within completely [168] or predominately [166] Caucasian cohorts thus limiting generalisability to other ethnic groups. Additionally they were limited by small sample size. Finally it has been hypothesized that, where standardised tests have a long duration, mental fatigue, rather than cognition, may adversely affect results [177] and this could be particularly important for students with ADHD.

Taanila et al [178] investigated mean overall school grades and those obtained in maths and Finnish for children with ADHD, ADHD and specific learning difficulty (SLD), SLD only and healthy controls. The authors reported that children with comorbid ADHD/SLD performed the worst, healthy controls performed the best, and children diagnosed solely with ADHD performed poorer than children diagnosed with SLD. Children were followed from birth and assessed, via teachers' reports, for SLD at age 7 and, via parental report, for ADHD at age 15. However, teenagers may experience behavioural issues during

adolescence which may have been mistaken for ADHD symptoms thus inflating numbers. The analyses, however, were adjusted for gender, socioeconomic status, family type and parental education and this was a strength.

Increased placement in special education and increased grade retention for children with ADHD has been reported across several studies including those by Le Fever et al [179] and Rhode et al [180] who both identified ADHD cases within a sample of children and additionally reported increased exclusion from school. Le Fever et al analysed a sample of 808 pupils within an urban, ethnically and economically diverse district covering three schools and reported that children with ADHD were 3.4 times more likely to repeat a grade, 6.7 times more likely to be excluded and 4.1 times more likely to be placed in special education services. ADHD diagnoses and school outcomes were ascertained from a parental survey; however the response rate was only 63% and academic outcomes were not clearly defined. Whilst analyses were adjusted for gender, race and possession of health insurance, other confounders such as age were not included and neither inclusion criteria nor ages of the children sampled were disclosed. Finally, 17% ADHD prevalence was observed, which exceeded figures reported elsewhere. For example, Rhode et al randomly sampled 1,022 12-14 year old children across 64 state schools and observed 5% ADHD prevalence. This study diagnosed ADHD using a screening instrument and pupils and parents specifically reported on school exclusion and grade repetition. In addition to poorer observed academic outcomes, 50% of children with ADHD also had comorbid oppositional defiant disorder or conduct disorder. These findings, however, were limited due to small sample size, univariate analyses with no adjustment for potentially confounding variables, and lack of matching of groups which resulted in the ADHD cohort having more boys than the non ADHD cohort. Whilst Le fever et al and Rhode et al [180] specifically investigated exclusions, Wu et al [173] reported that children with ADHD had worse behaviour, poorer school interaction and poorer attitudes to school. Similarly, Barkley et al [166] reported more antisocial behaviour and Kuriyan et al [181] reported more disciplinary problems in children with ADHD.

Few studies have investigated school absenteeism related to ADHD [172,182]. Kent et al [182] observed an increased absenteeism and lower grade point average for 326 children with ADHD compared with 213 matched controls. However, whilst this was a larger sample compared with other studies, it only included boys. Grade point average was derived using results in English, maths, social science and science.

Kuriyan et al [181] investigated self-reported educational outcomes at age 19-22 years and employment outcomes between ages 23 and 32 years for 364 subjects diagnosed with ADHD in childhood, compared with 240 matched controls. Subjects with ADHD had lower occupational and educational attainment, were less likely to go to college or obtain a degree, were more likely to be unemployed or be in an unskilled job and were more likely to be unemployed and not in education after leaving school. Matching ensured that the groups were similar in age, race and parental education levels; however, only boys were included in the study which limited generalisability because ADHD may affect boys and girls in different ways. Whilst 30% of the ADHD cohort was lost to follow up, significantly reducing the sample size, the study reported that the demographics of this group were not different compared with subjects that did agree to follow up and therefore this should not have biased the results.

2.4.3.2 Health outcomes

ADHD has been associated with poorer health outcomes. Those affected suffer an increased risk of mortality, in particular suicide[183], and are more likely to inflict self-harm[184]. Dalsgaard et al [185] linked 1.92 million population-wide Danish births between 1981 and 2011 to death records and identified 32,021 children with ADHD based on their admission to a psychiatric hospital. The overall MRR for subjects with ADHD compared to those without was 2.07. Whilst the authors demonstrated a larger effect size for those diagnosed in adulthood (MRR 4.25), increased mortality was also observed for subjects diagnosed before 6 years old (MRR 1.86) or between 6 and 17 years of age (MRR 1.58). Strengths of this study were the large population size and long duration of follow up (32 years); however, diagnosing ADHD based on psychiatric hospital admission is likely to have identified more severe cases and excluded children treated privately, by general practitioners or as outpatients. The analyses were adjusted for sex, age, history of psychiatric illness, maternal and paternal age, and parental employment and education. Whilst observed differences remained after excluding ADHD subjects with oppositional defiant disorder, conduct disorder or substance abuse disorder, these comorbidities may have been under-reported within hospital data.

Barbaresi et al [186] reported an increased risk of suicide (SMR 4.83) in 367 people with ADHD diagnosed in childhood compared with 4,946 age-sex matched controls but did not observe differences in all-cause mortality or accidental death. This study used robust methodology to diagnose ADHD using questionnaires, medical and school records and

clinical diagnoses. However, the sample was 95% Caucasian thus limiting generalisability to other ethnic groups. Schellerman et al [187] compared healthy controls against three age-sex matched cohorts of medicated children receiving prescriptions to treat ADHD: atomoxetine; amphetamine; or methylphenidate. No excess risk of all-cause death was reported for any of the medicated groups compared with healthy peers and no differences in risk of sudden, non-accidental or non-suicidal death or a range of cardiovascular events were identified. Though the study used a large diverse population, a small number of events were observed which meant that many analyses, especially of subgroups, could not be undertaken.

Children with ADHD suffer more injuries[188]. The nature of these injuries are more severe compared to peers without ADHD [184] and injuries are more likely to occur in boys[184]. Affected individuals use significantly more healthcare resources and have higher healthcare costs [189-191]. Few studies have investigated all-cause admission and most have focussed on injury admission instead. Hurtig et al [192] followed approximately 450 children with ADHD and 300 controls from birth until 18 years of age. Children were assessed, by teachers, at 7-8 years of age for signs of hyperactivity, by parents, at 15-16 years of age, for signs of ADHD, and formally for ADHD at 16-18 years of age. The authors reported more admissions due to injury, fractures and intoxication for children presenting with ADHD symptoms at 15 years of age and increased injury admission for those clinically diagnosed between 16-18 years of age. Analyses were adjusted for gender, number of children in the family, family structure, social status and medication use at age 15 years. ADHD was diagnosed between 15 and 18 years of age; however, lifetime hospital admissions were examined. Therefore, many of the hospital admissions may have occurred before children were diagnosed with ADHD and this was a notable limitation of the study. Additionally, children were assessed for ADHD symptoms at 15 years of age when they may have increased behavioural problems due to adolescence. Therefore it is possible that some may have been incorrectly diagnosed as having ADHD. Van Den Ban et al [188] also followed a birth cohort until 18 years of age and identified 1,289 children with ADHD based on receipt of prescriptions for methylphenidate or atomoxetine. These children had a two-fold risk of injury admission compared with 7,332 age-sex matched, non-medicated controls (6 controls per case). The highest risk occurred for children diagnosed between the ages of 12 and 18 years after adjusting for other psychotropic drug use. The authors investigated injury admissions before and after diagnosis of ADHD and utilised a large linked population cohort with up to 10 years follow up; however, some limitations were evident. Diagnoses were only based on prescription dispensing and there was no way of

ascertaining whether medication was taken. Secondly injuries were identified from hospital records; therefore, those treated at accident and emergency or at a general practice would not have been included resulting in potential under-reporting.

Shilon et al [193] and Pastor et al [194] also observed higher risk of admission for injury among children with ADHD but were hindered by limitations in their respective study designs. Shilon et al used a small sample of 29 ADHD cases and 29 controls and only followed subjects for nine months. Furthermore, the outcome was based on parental reporting of injuries requiring medical attention. Whilst the cohorts did not differ demographically, they were not formally matched and the analyses were not adjusted for other confounders. Pastor et al identified 3,741 children with ADHD within a larger survey cohort of 51,984 children aged 6–17 years of age and reported that children with ADHD had two-fold odds of injury compared to non ADHD peers. Analyses were adjusted for age, sex, race and health insurance but not for the presence of comorbid conditions. ADHD was diagnosed via parental report rather than formally by a clinician. Furthermore, injuries were identified by asking parents to recall the number of injuries during the preceding three months and then extrapolating this figure backwards across the previous full year. Therefore, recall bias may have occurred.

2.4.3.3 Summary

A systematic review and meta-analysis by Frazier et al in 2007 [13] investigated academic achievement but included studies on adults and pre-school children and discussed wider cognitive outcomes. Other notable reviews include those by Rowland et al [151], Loe et al [195] and Daley et al [196]. Previous studies have reported poorer academic performance [166-168,170,172-175], attendance [172,182] and leaver outcomes [167,172,181] and associations with SEN [166-168,176,179,180], learning disability [168,176] and school exclusion have also been observed [179,180]. Few studies have reported on all-cause mortality [185] or injury-related hospital admission [188,192-194] but those that have demonstrated increased risk for children with ADHD. Few of the identified studies relied exclusively on medication use to identify children with ADHD [188]. Community and school-based samples have been used [169,172-174,178-180]; however, most studies analysed a clinical sample of ADHD referred children [166,168,171,181], potentially limiting findings to more severely affected children. These may not be representative of children in the community who are likely to be better controlled. Many previous studies have also been limited by small sample size [166-170]. To the best of my knowledge no previous population-wide studies have investigated educational and health outcomes for

children with ADHD compared to peers and no studies of this type have previously been conducted in the UK.

2.5 Depression

Depression is a common mental health disorder which presents with depressed mood, fatigue and a loss of interest or pleasure in daily activities. Associated symptoms include low confidence or self-esteem, inappropriate guilt, disturbed sleep, changes in appetite, poor concentration, feeling sad, angry or withdrawn and suicidal thoughts [197]. However depression in young people may also present with behavioural disorders such as irritability, aggression and poor conduct, poor academic attainment and substance abuse. Young, schoolchildren frequently present with irritability, restlessness and hyperactivity, which may lead healthcare professionals to suspect ADHD rather than depression [198]. The most common signs and symptoms in adolescents and young adults are suicidal thoughts, hopelessness, social isolation, drug or alcohol use, overeating, oversleeping, and rage [199]. Depression can be classed as mild, moderate or severe and different types of depressive disorder exist, for example, recurrent or persistent depressive disorder or bipolar disorder [197,200]. Symptoms may significantly impair many aspects of life including general health, schooling, academic attainment and employment performance and are associated with risk of morbidity and suicide [201]. Main risk factors for depression include family history of the condition and exposure to psychological stress [197].

2.5.1 Prevalence

Depression is common, affecting around 16% of adults at some point in their lifetime [202]. However, the disorder is also relatively common in children and adolescents, with prevalence reported to be 2.6% worldwide [203]. Exact prevalence of depression is difficult to ascertain due to differences in diagnostic methodology and because many of those affected are not receiving active treatment. The condition is often unrecognised in adult [204] and paediatric [199] populations; however, methods of assessment for prevalence and treatment have improved dramatically [205]. Prevalence of depression is lower in childhood but increases sharply in adolescence, particularly in girls [197]. Worldwide prevalence of 2.8% in children under the age of 13 years of age and of 5.6% in adolescents (aged 13 to 18 years of age) has been observed and prevalence in adolescence is higher in girls than boys (5.9% v 4.6%) [206]. Indeed prevalence of depression is thought to be twice as high in

girls compared to boys [207]. Prevalence increased in adolescents and young adults between 2005 and 2014 [208].

2.5.2 Treatment

Depression can be treated with anti-depressant medication or via cognitive behavioural therapy (CBT) or interpersonal psychotherapy (IPT). However, concerns about the use of antidepressants in people younger than 18 years of age mean that treatment choices are not the same for children and adolescents compared to adults [197]. Typical medications include selective serotonin reuptake inhibitors (SSRIs), serotonin norepinephrine reuptake inhibitors (SNRIs) or tricyclic medications [209]. A combination of pharmacological and psychotherapeutic interventions is often used to treat depression [210]. People suffering from severe depression may also require electroconvulsive therapy (ECT) or hospitalisation [209,211].

2.5.3 Contextual review of the literature

The Medline search using the search terms in Table 5 returned 561 articles of which 69 (55 education and 14 health) remained after review of title and 40 (36 education and 14 health) remained after review of abstract. After full text review, 20 education papers and 2 health papers remained relevant to my thesis. The majority of these studies were undertaken in the USA and Canada and the remainder were from China, Holland, Sweden and New Zealand.

Table 5 Search terms for depression literature review

(((depress*)
AND
(college or youth or juvenile or young or schol* or adolescent* or child* or student* or pupil* or school)
AND
(school* or education* or attainment or academic or employment* or hospital* or death* or mortality or "learning difficulty" or "learning difficulties" or "learning disability" or "learning disabilities" or exclu* or "special need" or "special needs" or achievement or expulsion or suspension or detention or GPA or grade* or SEN or "assisted support needs" or attendance* or absen* or injur* or leav* or admission*))
NOT
(trial* or intervention* or cross-sectional or "college hospital" or "college teaching hospital" or "absence of")).ti.

**Filters applied: Title, Age 0-18 years, English language

2.5.3.1 School outcomes

Previous literature suggests that the association between depression and educational outcomes may be bidirectional. Whilst depression can adversely impact school performance through reduced ability to concentrate or lack of motivation to study, academic difficulties may equally lead to feelings of failure and eventual depressed mood [212]. Studies which only investigated relationships in the latter direction were not included in this review.

Previous studies have used longitudinal designs to investigate associations between depression and school leaver outcomes [14,15,213-215]. Three [14,15,214] analysed sub samples from the National Longitudinal Study of Adolescent to Adult Health (Add Health), a school based assessment of adolescents, their families, and their schools which was nationally representative of schools in the USA with respect to type, region, size, urbanisation and ethnicity. Needham et al [15] assessed self-reported depressive symptoms of 14,232 pupils with a mean age of 15 years using the Centre for Epidemiological Studies Depression Scale (CES-D) and evaluated self-reported completion of high school and entry to college approximately 6 years later. Children were classed as having high as opposed to low depressive symptoms, if they scored above a certain threshold on the scale (above 23 for girls and 21 for boys). No differences in school outcomes were reported between children with high and low symptoms after adjusting for gender, race/ethnicity, ability, age, parental education level, household income, family structure and previous pregnancy. However, analysed on a continuous scale, incremental increases in depression score were significantly associated with modest 2% increases in odds of failure to complete high school and failure to enter college. Whilst self-report of high school graduation and college entry should have been relatively robust, self-reported depressive symptoms may have been less accurate, particularly as children were interviewed during adolescence, when mood and behaviour can be particularly variable [216]. Secondly, the study did not investigate alternatives to college enrolment such as employment or commencement of trades or technical schools. Thirdly, presence of commonly occurring comorbid conditions, such as anxiety or obsessive compulsive disorder, which can independently impact academic outcomes, were not considered. Finally, limitations in data collection ensured that children who successfully entered college but dropped out within their first year were recorded as never having entered college which may have inflated the number of children observed with that particular outcome.

Using the same cohort and adjusting for additional community covariates, Fletcher (2008) [14] investigated the impact of dichotomous depression status on high school dropout and post high school enrolment in college including specifically two and four year college courses. Depressed girls were 3.5% less likely to graduate from school compared to girls without depression; however, there was no significant association within boys. The 8% of children classed as depressed were 6% less likely to be enrolled in college, specifically a four year course, after leaving school. A strength of this study was the adjustment for children's grade point average at study entry which was observed to have a negative relationship with depression. Each unit increase in grade point average reduced the odds of depression by nearly 40%. By taking into account innate baseline academic ability, the authors were able to investigate changes in grade point average associated with depression. Another noteworthy attribute was the inclusion of additional community variables in the analyses allowing further adjustment for socioeconomic background by taking account of circumstances specific to different neighbourhoods. This included data on the percentage of people unemployed, in poverty and without a diploma in the area where each pupil resided. Though both of these studies reported no overall association between depression and high school graduation, both did demonstrate significant associations within girls but not boys in subgroup analyses. Other studies have also demonstrated gender specific effects in subgroup analyses [215,217,218]

Fletcher (2010) [214] further assessed school dropout, college entry and years of schooling [214] after restricting the Add Health sample to 2,432 siblings and adjusting for family specific confounders, school specific characteristics, children's health behaviours and comorbid health outcomes. Similar to Needham et al [15], depression was investigated both as a dichotomous outcome and as a continuous scale of increasing symptoms. Depression was significantly associated with fewer years of schooling after adjusting for school specific characteristics; however this was not significant following further adjustment for family specific confounders, children's health behaviours and comorbid health outcomes. When depression was analysed as a continuous scale, increased depressive symptoms were significantly associated with increased school dropout after adjustment for both school specific characteristics and family specific confounders but not after adjustment for additional health behaviours and comorbid conditions. Increased depressive symptoms were not significantly associated with reduced entry. This study had several strengths. Firstly, adjusting for school specific characteristics took into account the fact that some schools have better outcomes than others. Similarly, adjusting for family specific confounders accounted for family and neighbourhood characteristics common to siblings,

including neighbourhood crime, family resources, parental influence and genetic similarities. Finally, the authors adjusted for additional child and health characteristics including marijuana, tobacco and alcohol use, teen pregnancy and sexual behaviour, obesity and other chronic conditions such as asthma, ADHD and diabetes. By using the "Add Health" sample, these studies experienced similar study limitations as those by Needham et al. A large number of different confounders were used across various models in the 2010 study. Whilst this was a strength, it is also possible that some of the analyses were over adjusted by adjusting for mediators on the causal pathway.

Smaller studies have also reported leaver outcomes. Fergusson et al [213] analysed 964 children who had complete follow up data of interest within the Christchurch Health and Development Study, a longitudinal study of a birth cohort followed until 21 years of age. Thirteen percent of children were diagnosed as depressed between the ages of 14-16 years based on separate parent and child interviews using the Diagnostic Interview Schedule for Children. The authors did not report any significant differences between depressed and unaffected subjects as regards their likelihood, by 21 years of age, to have left school with no qualifications or entered a university course, or equivalent training, post high school. Adjusted confounders included maternal education, socioeconomic status, previous sexual abuse, exposure to parental change, intellectual ability, neuroticism, comorbid anxiety or conduct disorder and behaviour patterns such as delinquency, smoking and alcohol and substance abuse. Prevalence of depression between the ages of 16 and 21 years was reported as 33.5% which seems very high compared with other studies. It is therefore possible that prevalence was over-estimated by parents and pupils and this may have biased the findings.

Jonsson et al [215] initially screened a cohort of 2,465 adolescents aged 16-17 years for depressive symptoms and classified them as depressed if they: had ever attempted suicide; scored > 16 on the Beck Depression Inventory (BDI); or scored > 11 on the BDI and at least 30 on the Centre for Epidemiological Studies Depression Scale for Children (CES-DC). On follow up, at 30 years of age, 345 depressed subjects had significantly lower levels of graduation from higher education compared with 243 unaffected sex-matched subjects. The association, however, was only significant for boys after adjusting for socioeconomic status, maternal education and final grade point average on leaving compulsory school. The diagnostic methods appeared to be robust particularly as additional diagnostic interviews were used to confirm the depressed status of subjects using DSM-III criteria. Subjects, however, were sampled from a Swedish university town

with higher levels of attainment compared to the general Swedish population. Furthermore, only higher education courses obtained in Sweden were recorded. Given the observed higher education levels, children in this sample may have been more likely to have attended university abroad and these courses would not have been identified.

Several cross sectional [219-222] and longitudinal studies [217,218,223-225] reported on attained academic grades of children with and without depression. Using a longitudinal design, Shahar et al [224] assessed 499 children in 6th and 7th grade using the Beck Depression Inventory (BDI) and reported that a higher score was significantly associated with lower mean grade point average two years later, based on school recorded grades in English, maths, social studies and science. Depressive symptoms were also negatively correlated with grade point average derived at the start of the study; therefore, adjustment for baseline grade point average, gender and self-criticism was a notable strength. However, the sample only included children in 6th and 7th grade which limited generalisability to other age groups. Lundy et al [225] demonstrated that 335 white and Hispanic children, aged 6 to 11 years of age who scored 60 or above in the anxious/depressed and withdrawn scales of the Childhood Behavioural Checklist (CBCL), had poorer reading, writing and maths scores, measured one month later by the Woodcock-Johnson test of Achievement, compared to children who scored below this clinical threshold. This study, however, did not appear to be adjusted for any confounders. Further, only white and Hispanic children were sampled thus limiting generalisability to other ethnic groups. Depression was diagnosed using only two subscales of the CBCL: anxious/depressed and withdrawn. It is possible that some children were classed as depressed despite only presenting anxious or withdrawn symptoms thus inflating depression prevalence and observed associations.

Whilst Shahar et al and Lundy et al measured depressive symptoms at one point in time and outcomes at a singular follow-up point; other longitudinal studies measured them across multiple time points in order to explore changes in depressive symptoms and associated changes in academic function. Marcotte (2006) [226] measured the same depression symptoms and outcomes at three separate time points and assessed differences in school recorded grades in maths and French between students who were always depressed, never depressed, became depressed and stopped being depressed. Pupils classed as always depressed had significantly worse academic grades; however, there was no association between temporal changes in depressive symptoms and temporal variations in academic outcomes. The study included 644 pupils; however, they were grouped and analysed in sub-categories based on gender and temporal pathways of depressive

symptoms. Therefore, small sample sizes within each category reduced the power of the study which may have hindered the authors' ability to identify small effect sizes. This limitation was shared by other studies stratifying across several trajectories [217,218,223]. Dekker et al [218], for example, sampled 2,076 children between 4 and 16 years of age from the general population of one Dutch province and measured their depressive symptoms at six time points across fourteen years via parental interviews using the Childhood Behavioural Checklist (CBCL). Semi-parametric growth mixture models described trajectories of depressive symptoms across childhood and associations with final educational attainment dichotomised as low (ranging from school dropout to a lower vocational level) or high (ranging from completion of high school to higher education). Children were categorised into six trajectory paths and boys with more severe depression in childhood had a significantly lower educational level by the final stage of follow up compared to boys with depressive symptoms classed as low to moderate. There was little difference observed in girls; however, girls with more severe depression did experience increased mental healthcare usage compared to girls with less severe depression.

Briere et al [223] also used growth mixture models to assess self-reported depressive symptoms, using the Centre for Epidemiological Studies Depression Scale (CES-D), and self-reported mean grades in maths and language, for 6,910 secondary schoolchildren. Self-reports were made annually between the ages of 12 and 16 years and children were grouped based on trajectories of their depressive symptoms: stable-low; increasing; decreasing; transient; or stable-high. Children who had severe but stable depressive symptoms had significantly lower mean grades compared to those with stable but less severe depressive symptoms. Furthermore, academic outcomes at each time point closely mirrored the trajectory of depressive symptoms. However, the results may have been compromised by common or shared method variance because depressive symptoms and academic grades were both self-reported by pupils every year. Pupils were asked to rate their current grades between 0% and 100 % to the best of their knowledge; therefore it is not inconceivable that students may have provided lower ratings when they were more depressed and vice versa. Additionally, most of the sample resided in disadvantaged areas therefore these results may not be generalisable to the wider population.

Rather than investigating relationships in one direction, Verboom et al [217] recorded data on 2,230 children every two years over a six year period and used structural equation modelling to investigate whether teacher and pupil reported depressive symptoms were predictive of teacher reported academic performance or vice versa. Within each set of

observations, more severe depression was correlated with lower academic performance for boys and girls. A temporal bi-directional relationship also existed for girls whereby increased depression was predictive of future reduced academic performance and vice versa. Therefore, in girls, whilst increased depression negatively impacted future attainment, poor academic performance also increased the risk of subsequent depression. There was, however, no temporal relationship between depression and academic ability in either direction for boys. However, several limitations regarding quality, accuracy and consistency of data collected across the three time points were apparent in the study design. Firstly, while academic performance at the first time point was reported by a teacher, performance at later time points was reported by a school mentor who spent less time with pupils. Secondly, reports were initially based on maths and language, however performance in geography, history, chemistry and biology were later incorporated. Finally, the teacher questionnaire completion rate was initially 86%; however, this fell to 70% at the second time point and 42% at the last data collection point.

When measuring depression and educational outcomes at the same point in time it is not possible to decipher whether depression has adversely impacted educational outcomes, poor educational performance have contributed to depression or whether both are true. Therefore, whilst earlier cross-sectional studies adopted simpler analyses, they were limited by potential reverse causation and tended to use smaller samples [219-222]. Hamilton et al [220] compared 15 children hospitalised with depression against 14 children hospitalised with schizophrenia disorders and 20 healthy controls. Compared with controls, children with depression had similar academic attainment; categorised on a scale of one (poor) to five (excellent), based on school report cards, standardised tests and teacher evaluation. However, all of the depressed children had been hospitalised and therefore were likely to have severe depression not representative of children diagnosed with depression in the wider community. The authors did not confirm whether analyses were adjusted for confounders and measured academic attainment in the six-month period before hospitalisation. The non-significant results were likely due to lack of power due to the very small sample size.

Brumback et al [219] demonstrated no difference in academic test scores between depressed and unaffected children based on univariate analysis, using the Wide Range Achievement Test (WRAT). However, the study lacked power due to small sample size and 62 children were diagnosed as depressed within a cohort of 100 children described as having school problems. The study did not define these school problems or explain how depression was

diagnosed; however, these children were most likely not representative of the wider school population. Their problems may have contributed to the high prevalence of depression observed and may have underestimated the impact of depression.

Vincenzi et al [221] and Tesiny et al [222] investigated correlations between depression scores and academic scores. Vincenzi et al reported depression scores to be negatively correlated with reading level, monitored via teacher reports and standardised test scores from the California Achievement Test (CAT), and grade point average, derived using grades in reading, maths, science and social studies. The sample, however, only included 139 black students in 6th grade from an urban area thus limiting generalisability to other settings and ethnic groups. Tesiny et al analysed a larger school sample of 452 boys and 492 girls aged 8-12 years and reported negative correlations between depression and academic outcomes based on standardised test scores and teacher ratings of work, study habits and achievement. The authors used an unconventional method to classify children's depressive symptoms by allowing pupils to rate their classmates on how depressed they perceived them to be using the Peer Nomination Inventory (PNID). This allowed pupils to nominate particular classmates who they thought were most depressed. Whilst this was a novel approach, it lacked clinical input and may have been prone to potential bias or inaccuracy.

My literature search did not identify any studies that exclusively investigated associations between depression and school absenteeism, exclusion or special educational need in a school setting. Several studies, however, reported associations between depression and behavioural, cognitive and social outcomes. Lower levels of social competence [220] and poorer social outcomes [217] have been reported for depressed children and behavioural outcomes have been observed to mirror the course of depressive trajectories [223]. Whilst Brumback et al [219] reported no difference in IQ between depressed and unaffected pupils, Lundy et al [225] reported that children with increased depressive symptoms experienced deficits in general intellectual functioning, language, visual construct, attention, processing speed, executive function, aspects of learning and memory and psychomotor speed. Dekker et al [218] reported that children diagnosed with depression at school had an increased risk of mental health problems in young adulthood and that depressed girls were more likely to use future mental health care services. Fergusson et al [213] observed that depressed children had increased future risk of major depression or anxiety disorder in young adulthood. Behavioural and social problems, neuropsychological impairment and comorbid mental health conditions may all potentially impact absenteeism, risk of exclusion and risk of SEN in addition to academic attainment and leaver prospects.

2.5.3.2 Health outcomes

Depressive symptoms in childhood can also negatively impact subsequent health outcomes in adolescence and beyond [227]. However, very few studies have investigated hospital admissions and mortality following child and adolescent diagnosed depression. Whilst there are associations between child and adolescent depression and increased risk of suicide [201], the majority of studies have reported on suicidal ideation rather than suicide [228]. Previous literature has demonstrated increased risk of child and adolescent depression following injury admission [229]; however, few have reported relationships in the opposite direction. Other studies have assessed screening tools and interventions for future risk of self-harm and suicide in depressed cohorts. The Medline literature search did not identify any previous studies assessing risk of all-cause mortality or all-cause hospital admission for depressed children compared to children without depression or the general population. However, two studies assessed future risk of injury [230,231].

Asbridge et al [230], reported an increased risk of violence related, traffic related and unintentional injuries in depressed schoolchildren aged between 14 and 18 years. Analyses were adjusted for gender, school grade level, weekly spending money, smoking status, past year heavy drinking and marijuana use, sexual behaviour, absenteeism, school grade point average and participation in team sports. However, depression, self-reported via the Centre for Epidemiological Studies Depression Scale (CES-D), was only diagnosed based on depressive symptoms in the past week and injuries were self-reported over the previous six-months. The study was cross-sectional in design and measurements were taken at the same point in time; therefore, whilst crude associations between depression and injury could be investigated, longitudinal relationships in either direction could not. It is also possible that some injuries may have occurred before symptoms of depression became apparent. Children were classed as depressed if they exceeded a cut-off score on the depression scale; however, without clinical diagnoses, some of these cases may have been misclassified. Finally the study did not include other confounders potentially predictive of later injury such as socioeconomic status.

You et al [231] reported depressed secondary schoolchildren aged between 11 and 19 years to have an increased risk of non-suicidal self-injury (NSSI). However, this study was also hindered by study design issues. Depressed and unaffected children were followed up one year later to assess self-injuries; however, these were self-reported if they occurred at any point in the last two years and this window period pre-dated the initial self-reported

measure of depression one year earlier. Therefore, it is possible that several of these NSSI events predated depressive symptoms.

2.5.3.3 Summary

The majority of previous comparable depression studies have reported poorer attainment [224-226] and poorer school leaver outcomes [14,15,213-215] for depressed children compared to peers. To the best of my knowledge, no previous studies have investigated school absenteeism, exclusion or special educational need of depressed children compared to peers. Similarly, no studies have investigated all-cause admission or all-cause mortality; however, increased risk of injury admission for depressed children compared to peers has been reported [230,231]. To the best of my knowledge, no studies comparing educational or health outcomes for depressed children and unaffected peers have analysed population-wide data or adjusted for maternal confounders or increased absenteeism. Furthermore no previous studies have investigated multiple educational and health outcomes and none have previously been conducted in the UK. Additionally several previous studies have experienced limitations such as small sample size [219-221].

3 Data providers, data sources and governance

3.1 Linkage infrastructure in Scotland

Record linkage techniques enable multiple records, pertaining to the same individual but held across multiple datasets, to be brought together [232]. Where common unique identifiers relating to an individual exist across each dataset, records can be easily combined using exact matching. However, in the absence of unique identifiers, probability matching techniques are employed using person identifiers in order to calculate the likelihood that records across different datasets belong to the same person [233-235]. Typical identifiers used are name, date of birth, sex and postcode; therefore, holding identifiable information enables record linkage to produce rich datasets for research purposes and patient based analyses with wide applications in epidemiological research and public health [234-236].

Scotland has some of the best health data compared to other countries and an excellent infrastructure for performing record linkage due to historic and ongoing collection of routinely collected, high quality and well-maintained national administrative datasets and continually evolving linkage methodology [236]. Whilst record linkage has historically brought together health datasets, recent developments have focussed on linking health data to data held across other sectors [237]. The determinants and outcomes of health extend beyond the health sector [2,3] and therefore linking data across different sectors provides an invaluable resource for epidemiological research.

3.1.1 Farr Scotland

This PhD was undertaken under the auspices of the Farr (Scotland) Institute which is part of the Farr Institute of Health Informatics Research [238]. The Farr Institute is a UK research collaboration that is comprised of 21 institutions within academia and various other health partners. It aims to put the UK at the forefront of health informatics by delivering high quality and cutting edge research. Farr (Scotland) is one of the four main Farr centres located across the UK and led by University of Dundee, University College London, University of Manchester and Swansea University. Farr (Scotland) is a joint collaboration between six Scottish Universities (Aberdeen, Dundee, Glasgow, Edinburgh, St Andrews and Strathclyde) and NHS National Services Scotland. It comprises a number of different work streams including one specifically on cross-sectoral record linkage and one on pharmaco-epidemiology. The cross-sectoral work stream aims to establish a number of novel cross-sectoral linkages and use these to undertake exemplar projects to demonstrate

the feasibility and utility of research based on cross-sectoral data. The linkage of education data to prescribing data supported collaborative research between the Farr cross-sectoral linkage and pharmaco-epidemiology work streams.

3.1.2 NHS National Services Scotland

NHS National Services Scotland (NSS) is a non-departmental government body which provides national strategic support services and expert advice to NHS Scotland [239]. NSS works closely with other NHS organisations, particularly the fourteen health boards, in order to deliver effective healthcare to patients and the public and it is accountable to the Scottish Government. Information Services Division (ISD) is part of NHS NSS and collects a wide range of health related administrative data on its behalf. It supports NHS Scotland and the Scottish Government by providing health information and intelligence and statistical services. Nationally collected, high quality and well-maintained national administrative datasets and the ability to link these together using sophisticated record linkage techniques ensure that Scotland has excellent health data resources. ISD has played a pivotal role in record linkage related research over many years with a service dedicated to linking data for research purposes. The electronic Data Research and Innovation Service (eDRIS) is the current name for ISD's record linkage service which assists the research community by providing expert advice, linking data and providing data access in a secure environment.

3.2 Data sources

All of the education data for the project were held by ScotXed (Scottish Exchange of Educational Data) [240] which is part of the Education Analytical Services Division within the Learning and Justice Directorate of the Scottish Government. Health data of interest for the project were held within ISD. The key datasets that were used in the study are described in sections 3.2.1 and 3.2.2.

3.2.1 Education data

ScotXed holds national level education data across various datasets including pupil census data, exam results, attendance (absence and exclusion) data and school leaver data. All datasets contain a pupil specific Scottish candidate number (SCN) and a range of pupil identifiers. Pupil census data are collected annually, in September, to coincide with the start of each new school year, by all local authority maintained primary, secondary and

special schools in Scotland [241]. It includes individual level data for all children who are currently at school from first year of primary school (P1) through to final year of secondary school (S6). The pupil census contains various data including whether each pupil has a special educational need (SEN) during a particular school year. The Scottish Qualifications Authority (SQA) [242] holds examination attainment data for all schoolchildren sitting exams at Scottish schools and these data can be linked to the pupil census using the Scottish Candidate Number for each pupil. ScotXed's pupil leaver destination dataset contains information on pupils' employment or educational status 6 months after leaving school [243,244]. Attendance (absence and exclusion) data are submitted at the end of each school year and contain full attendance histories for each pupil over the previous academic year. Numbers of authorised and unauthorised absences and reasons for absence are recorded. Similarly, numbers of exclusions and length of exclusion are recorded where children have been excluded from school for unacceptable behaviour. Children attending private schools are not included: however, less than 5% of Scottish schoolchildren attend private schools [245]. Therefore the pupil census achieves more than 95% coverage of children attending school in Scotland. These datasets are summarised in Table 6.

Table 6 Summary of education datasets

Dataset name	Purpose/Description	Years collected	Typical data provided
Pupil Census	Collected annually for all local authority maintained primary, secondary and special schools in Scotland	2006 onwards	Individual data for all children who are currently at school from first year of primary (P1) through to final year of secondary (S6) including children in special schools. Includes demographic and personal details, school attended, stage of education, looked after status, free school meals, special educational need, level of English.
Scottish Qualifications Authority (SQA)	Examination attainment data for all school children in Scotland	2006 onwards	Exam results from fourth year (S4) through to sixth year (S6) of secondary school. Course subject and grade awarded.
Pupil leaver destination	Information on pupils' destination 6 months after leaving school	2006 onwards	Employed or unemployed, in higher education or training.
Attendance (absence and exclusion)	Full attendance, absence and exclusion history for each pupil over the previous academic year	2006 - 2010 and 2012	Number of days absent, reason for absence, number of exclusions, reason for exclusion, length of exclusion.

3.2.2 Health data

Information Services Division (ISD) [246] holds information relating to all medicines dispensed in the community in Scotland within its prescribing database known as the Prescribing Information System (PIS). Prescribing information is supplied to ISD by the Practitioner Services Division (PSD) which is responsible for the processing and pricing of all prescriptions in Scotland. Most drugs on PIS are prescribed by general practitioners, with the remainder prescribed by other authorised healthcare professionals such as nurses and dentists. Drug prescriptions issued in hospitals and dispensed in the community are included; however prescriptions dispensed within hospitals are not. These data are augmented with information on prescriptions written in Scotland and dispensed elsewhere in the UK. Prescribing data have been collected since 1993 however PIS data have only become linkable to other datasets from 2009 onwards when CHI recording on the prescribing data exceeded 90%. The Community Health Index (CHI) database is a register of all patients in NHS Scotland; CHI numbers are recorded on health datasets, are unique

to each patient, and enable patient records to be linked together across different health databases [247].

ISD's Scottish Morbidity Database holds permanently linked datasets relating to acute hospital admissions, psychiatric admissions and death registrations in Scotland from 1981 onwards. Records pertaining to the same person are permanently linked together to form groups of records relating to the same patient across all datasets. The Acute and Inpatient Day Case dataset, known as Scottish Morbidity Record 1 (SMR01), collects episode level data on all inpatient and day case discharges from acute specialities in Scottish hospitals. The Mental Health Inpatient and Day Case dataset (SMR04) collects episode level data on patients that are receiving care at psychiatric hospitals. Finally, National Records of Scotland (NRS) death registrations are submitted to ISD and contain an up to date record of all deaths occurring in Scotland. All deaths must be registered within eight days by a Local Registrar.

ISD holds a number of linked maternity and neonatal datasets within its Maternity and Neonatal Linked Database. This contains obstetric histories relating to all mothers giving birth in Scotland from 1975 onwards together with linked records relating to all of her offspring. Records pertaining to the same mother and all of her offspring are permanently linked together in patient groups. The Maternity Inpatient and Day Case dataset (SMR02) collects episode level data every time a mother is admitted for any type of obstetric event including giving birth and admissions for pregnancy related morbidity such as pre-eclampsia. It includes information on both the mother and her offspring. These datasets are summarised in Table 7.

Table 7 Summary of health datasets

Dataset name	Purpose/Description	Years collected	Typical data provided
Maternity Inpatient and Day Case dataset (SMR02)	Episode level data on all maternity events within obstetric specialties in Scotland. Contains data relating to mother and baby.	1975 onwards	Mother's data includes identifiable information (date of birth, name, postcode, CHI number), episode management information (date of admission/discharge, admission reason, admission type, discharge location, specialty), clinical information (diagnosis codes, operative procedure codes and dates), information on previous pregnancies and antecedents (parity, previous abortions, smoking status), pregnancy outcomes (mode of delivery, induced, duration of labour, stillborn, abortion) and demographic data (health board of residence/treatment, socioeconomic status). Baby data includes gender, birth weight, gestational age, APGAR score, presentation.
Acute and Inpatient Day Case dataset (SMR01)	Episode level data on all inpatient and day case discharges from acute specialities in hospitals in Scotland	1981 onwards	Identifiable information (date of birth, name, postcode, CHI number), episode management information (date of admission/discharge, admission reason, admission type, discharge location, specialty), clinical information (diagnosis codes, operative procedure codes and dates) and demographic data (health board of residence/treatment, socioeconomic status).
Mental Health Inpatient and Day Case dataset (SMR04)	Episode level data on patients receiving care at psychiatric hospitals in Scotland	1981 onwards	Identifiable information (date of birth, name, postcode, CHI number), episode management information (date of admission/discharge, admission reason, admission type, discharge location, specialty), clinical information (diagnosis codes, operative procedure codes and dates) and demographic data (health board of residence/treatment, socioeconomic status).
National Records of Scotland (NRS) Death Registrations	Record of all deaths in Scotland	1981 onwards	Identifiable information (date of birth, name, postcode, CHI number), all main and secondary causes of death, date of death, date of registration and demographic data (health board, socioeconomic status).
Prescribing Information System (PIS)	Information relating to all medicines prescribed and dispensed in the community in Scotland.	1993 however PIS data only linkable from 2009 onwards when CHI completeness exceeded 90%.	Patient identifiable information (date of birth, CHI number, sex), prescriber and dispenser details (location, organisation, geography), practice information, costs and drugs information (date prescribed/dispensed/paid, BNF drug codes) and information on prescribed items (drug name, formulation, strength, description, type, dose, quantity).

3.3 Governance and approvals

Education data have previously been linked to SMR02 obstetric data to investigate the association between birth outcomes such as birth weight and gestational age and risk of SEN [16,17]. Previous studies have also sought approval to link PIS prescribing data to SMR01 and SMR02. This study, however, was the first in Scotland to link education records to PIS prescribing data. Furthermore, additional linkage of these datasets to acute and psychiatric hospital admissions, maternity data and deaths ensured that this linkage

was the first of its kind, enabling associations between childhood chronic disease and health and educational outcomes to be studied whilst controlling for a range of potential confounding socioeconomic, demographic, obstetric and maternal factors.

Research projects requiring linkage of data to NSS health datasets or release of patient identifiable information must be approved by an appropriate governance panel. At the time of the study application, the NHS National Services Scotland Privacy Advisory Committee (PAC) was an advisory committee to the Board of NHS National Services Scotland (NSS) and to the Registrar General and provided advice on requests for the release of patient identifiable information held by Information Services Division (ISD) or National Records Scotland (NRS). PAC advised ISD and NRS on the correct balance between protecting personal data and making data available for research, audit and other important uses and ensured that any information releases were carefully controlled. The PAC has since merged with several other governance panels, such as National Caldicott Guardians and the Community Health Index Advisory Group (CHIAG), to form a new overarching Public Benefit and Privacy Panel (PBPP) for Health and Social Care [248]. In addition to PAC, now PBPP, approval, several other governance measures need to be in place particularly for large scale complex research projects which involve multiple linked datasets. The various approvals that were needed for this project are explained next.

3.3.1 Privacy Advisory Committee approval

An application to the NHS National Services Scotland Privacy Advisory Committee (PAC) seeking permission for the linkage of education data held by ScotXed to health data held by NHS National Services Scotland (SMR01/02/04, deaths and prescribing) was approved in September 2014 (Reference: 51/13). PAC requires permission from all data controllers to be granted before it considers linkages to health data and also requires approval from an appropriate ethics body where deemed necessary.

3.3.2 NHS ethics approval

This PhD was an extension to a previous linkage of ScotXed pupil census data, SQA examination data and SMR02 data which examined the association between obstetric practice and pregnancy complications and subsequent educational outcomes [16,17]. The NHS West of Scotland Research Ethics Service confirmed that NHS ethics approval was not required previously, or for this thesis, as no identifiable data were transferred to the

University of Glasgow. Since this study involved linkage of routinely collected data with an acceptably negligible risk of identification, PAC approval was all that was required.

3.3.3 Data sharing agreement

This study linked patient information held across several administrative health datasets within ISD with externally held data controlled by the Scottish Government (ScotXed education). Whilst the application to PAC could approve the overall study, it could only authorise Glasgow University to receive data controlled and held by ISD Scotland within NSS. The PAC approval could not authorise Glasgow University to receive ScotXed data without permission from ScotXed as these data were controlled by the education department of the Scottish Government (SG). Additional data sharing agreements were therefore put in place. A data sharing agreement with ScotXed was developed to allow: education linkage identifiers to be provided to ISD; linkage to the CHI database; and education payload data to be released thereafter into the national safe haven for analyses.

3.3.4 Data processing agreement

ISD linked education data to health data on Glasgow University's behalf and subsequently processed and stored the payload data in the national safe haven for me to analyse. Therefore a second data processing agreement was developed between Glasgow University and ISD.

3.4 Safe haven, data storage and security

Data were not transferred directly to Glasgow University. Instead, the eData Research and Innovation Service (eDRIS) [249] stored the data within the national safe haven for me to access and analyse remotely. Storing in a safe haven environment ensures that research can be done on sensitive data with minimum risk of disclosure by controlling who can have access, what data can be analysed and what outputs can be taken away. This provides a secure environment for the linkage, storage and analysis of personal data. Only approved researchers and the eDRIS research coordinator assigned to a study have access to data within the safe haven. The safe haven was accessed remotely via Cisco Systems VPN (virtual private network) Client; a software application for connecting to a network in a secure way. This was hosted by Citrix [250]. The Citrix platform was configured to remove any functions not needed by end users e.g. internet access or uploading and downloading files. This meant that I essentially worked at a dummy terminal where no data could be

removed, copied or stored. All researchers with access to data are required by eDRIS to be named on the study's PAC application form and to hold a current certificate of training in information governance, to be employed by or sponsored by an approved organisation and to complete and sign an eDRIS user agreement. This agreement specifies eDRIS user responsibilities along with sanctions and penalties for any breaches.

All analyses were completed within the safe haven and outputs were disclosure checked for anonymity prior to release. An eDRIS research coordinator accessed the safe haven to carry out statistical disclosure control on all outputs in order to prevent any accidental disclosure and to ensure that individual level data or results with small frequencies, typically less than five, were not released. Only disclosure controlled outputs were transferred to me. The national safe haven environment was initially provided by NSS via Atos Origin Alliance (AOA) which provided the secure environment for NHS Scotland applications at the time. ISD provided the indexing service and all study data were transferred from data providers to the safe haven using ISD's secure file transfer protocol (SFTP) facility. Approximately half way through the study, the Edinburgh Parallel Computing Centre (EPCC) [251] based at Edinburgh University assumed responsibility for the safe haven environment and provided this service in place of AOA.

4 Methods

4.1 Creating the study cohort using linked data

Whilst education data were available from 2006 onwards, prescribing data were only available for linkage from 2009 onwards. Therefore only education records from 2009 onwards were considered for analyses. The ScotXed school census therefore provided data on all children attending local authority schools (including special schools) in Scotland between 2009 and 2013. These children were predominately between the ages of 5 and 18 years and the majority of pupils had data across more than one school year. In order to study the association between childhood chronic disease and subsequent educational and health outcomes, all children on the school census were linked to the PIS, SMR01, SMR04 and NRS deaths. By linking pupils to their mothers' SMR02 maternity records, the cohort was limited to children born in Scotland and attending school in Scotland.

Chronic conditions of interest were identified by linkage to cashed medications recorded on the PIS. Whilst many medications can have multiple indications, some are used primarily or exclusively to treat one condition. For example, children with type 1 diabetes are treated exclusively with insulin whilst children with ADHD are treated with central nervous system (CNS) stimulants. Other examples are anti-epileptic medication and antidepressants, used to treat epilepsy and depression respectively, and bronchodilator and steroid inhalers, which are used to treat asthma. Whilst various caveats apply, and these will be discussed in due course, in theory children in receipt of these drugs can be reliably assumed to have these conditions. Furthermore, all of the above childhood conditions can plausibly be expected to have a negative impact on one or more health or educational outcomes such as attainment, absence, exclusion, SEN and unemployment. I aimed to identify children with epilepsy, asthma, ADHD, depression and diabetes through PIS records of relevant cashed drugs. PIS provided good CHI coverage from 2009, therefore investigation of medications taken by children as a proxy for disease were limited to prescriptions cashed from 2009 onwards. For children receiving prescriptions, only prescriptions of interest were requested from the PIS database: drugs prescribed for diabetes, asthma, epilepsy, ADHD or depression. Children were linked to hospital admissions (SMR01/SMR04) and deaths to investigate non-fatal and fatal health outcomes. SMR01 and SMR04 provided information on any childhood morbidity resulting in acute or psychiatric hospital admissions. Linkage to SMR02 provided information on obstetric and birth outcomes and maternal antecedents allowing the analyses to be adjusted for potential

or known pregnancy and maternal confounders. The outcomes available as a result of the above linkages included:

- ScotXed/SQA – SENs and reason for this; absence, and exclusion; examination attainment; unemployment
- SMR01/04 – acute hospital admissions / psychiatric hospital admissions
- Death certificates – cause-specific and all-cause mortality

In addition to limitations set by the availability of particular datasets (e.g. PIS data only available from 2009 onwards), some of the outcomes of interest were also limited to particular subgroups. For example, school exams are only undertaken by older pupils in S4, S5 or S6. Similarly, assessment of unemployment was limited to school leavers. The main associations of the chronic conditions of interest were analysed adjusting for potential confounders, such as socioeconomic and demographics factors, outcomes at birth and maternal factors. Relevant interactions were tested and sub-group analyses were performed where appropriate. Sensitivity and mediation analyses were performed whereby attainment and unemployment were re-analysed after excluding children with SEN and adjusting for absenteeism. Final models were run including all diseases together to investigate whether the effects of each disease were independent of the other conditions.

4.2 Primary linkage methodology

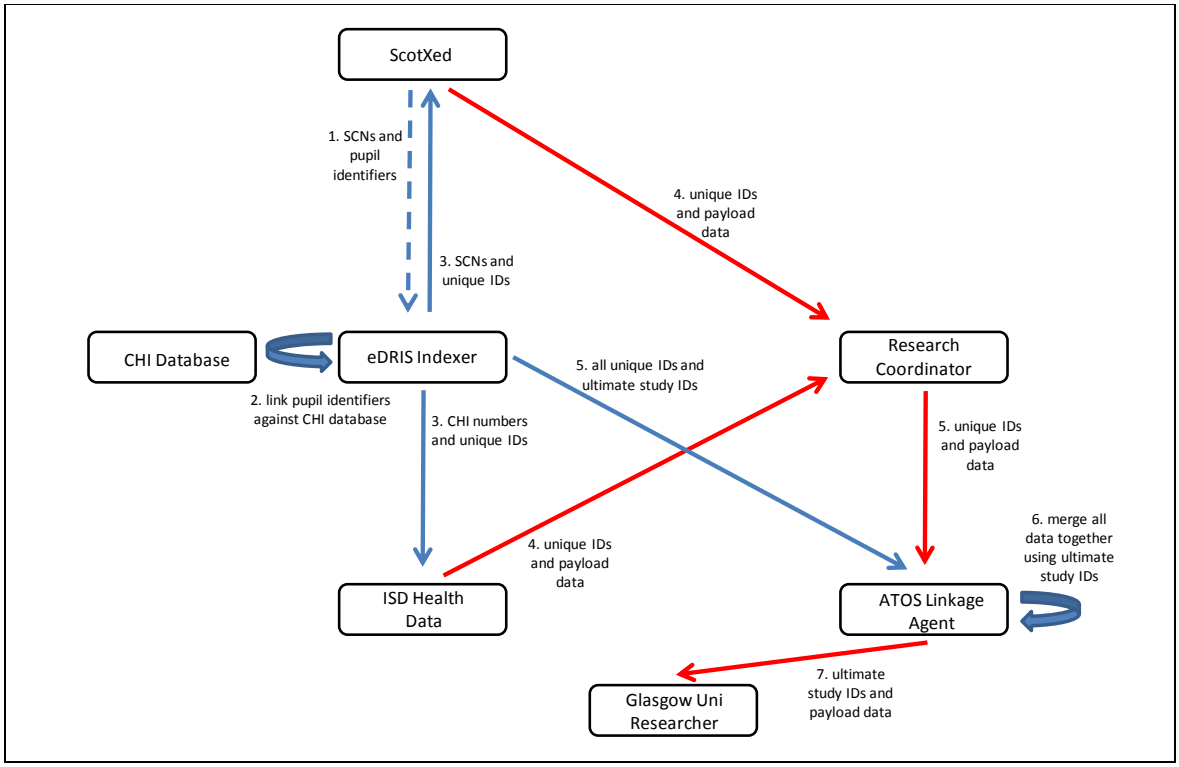
Individual level education records held by ScotXed were linked to individual level health records held by NHS National Services Scotland (SMR01/02/04, deaths and prescribing). All of the health datasets contained patient Community Health Index (CHI) number. The CHI database is a register of all patients in NHS Scotland; the CHI number is a ten digit number unique to each patient and is only recorded on health datasets. No single body has responsibility for CHI and the fourteen health boards are the data controllers. The health sector uses the patient CHI number as a unique identifier which enables different health databases to be linked to each other, at an individual level, using exact matching.

The education sector uses a different unique identifier, the pupil index, known as the Scottish Candidate Number (SCN), by which different education databases can be linked. At present, the two identifiers are not linked to each other therefore probabilistic matching

was used to link the education databases to the CHI register in order to attach a CHI number to each education record. ScotXed held full names for each pupil on their pupil census however they were not permitted to release them to a third party for linkage purposes. Pupil records from ScotXed were therefore probability matched against the CHI database using date of birth, sex and postcode of residence. This methodology has previously been validated and shown to be 99% accurate for singleton births [237].

The linkage of education and health data was a two stage process. The initial probabilistic stage of the linkage was performed by eDRIS and produced a lookup file of each pupil's Scottish Candidate Number (from education) and their corresponding CHI number (from health). This was subsequently used to obtain relevant health and education data via exact (deterministic) matching using CHI number and pupil SCN. The CHI-SCN lookup was held by the eDRIS indexer (indexing is a commonly used term for linking data together either deterministically or probabilistically). The full linkage process is shown in Figure 1 and described in seven stages.

Figure 1 Flow diagram for linkage of education to health data



Based on the diagram above, the specific linkage process for the project was as follows.

1. ScotXed sent pupil specific linking identifiers (date of birth, sex, postcode of residence) and their corresponding Scottish candidate number to eDRIS to probability match against the CHI database.

2. An eDRIS indexer probability-matched each pupil record against the CHI database to obtain a CHI number for each pupil where possible.

3. The eDRIS indexer returned to ScotXed a list of the Scottish candidate numbers which had successfully linked to the CHI database along with a unique ID for each pupil. The Scottish candidate numbers were returned to allow ScotXed to append previously requested educational outcome variables, known as the payload data, from several different data schemes namely the pupil census (including record of SEN), SQA exam results, school leaver destination and absence and exclusion data. Similarly, the eDRIS indexer returned a list of the linked CHI numbers to the various health data providers (various analytical teams within ISD) along with a unique ID for each child. The CHI numbers were sent to each data provider to allow requested health payload data to be appended using CHI number. Health payload data included prescribing data (PIS), acute (SMR01) and psychiatric (SMR04) hospital admissions, deaths (NRS) and maternity records (SMR02). These IDs were unique for each pupil and unique for each data provider, for example different IDs were sent to ScotXed and to each of the data providers within ISD.

4. Having used SCN to match the relevant education data to the unique IDs, ScotXed deleted the SCN's and sent the unique IDs and education variables to the eDRIS Research Coordinator. Similarly, having used CHI number to match the relevant health data to the unique IDs, each health data provider deleted the CHI number and sent the unique IDs and health variables to the eDRIS Research Coordinator.

5. A linkage key, held only by the eDRIS indexer, maintained all data provider specific IDs and an overarching study ID for each unique individual; known as the study index number. This linkage key was sent to the eDRIS linkage agent (a Java script within ATOS origin). Meanwhile, the eDRIS research coordinator forwarded

all education and health payload data and associated unique IDs to the eDRIS linkage agent.

6. Health and education data were merged together by the eDRIS linkage agent using the various unique IDs and the overarching study IDs (study index numbers) within the linkage key forwarded by the eDRIS indexer.

7. The data were then stored in the national safe haven in order for me to access and analyse. All submissions at each step were done via secure file transfer. I had remote access to the final linked health and education datasets held within the national safe haven.

These steps ensured that all of the parties involved in the project could only access certain parts of the data and that no parties involved could ever access all of the data linked together. Table 8 illustrates which parties had access to which parts of the data.

Table 8 Study involvement and access to data

Data	Education data provider	Health data providers	eDRIS indexer	eDRIS linkage agent	Researcher
SCN / unique ID key	Yes	No	Yes	No	No
CHI / unique ID key	No	Yes	Yes	No	No
unique IDs / overall study ID key	No	No	Yes	Yes	No
Health payload data	No	Yes	No	Yes	No
Education payload data	Yes	No	No	Yes	No
Linked health and education dataset with overall study IDs	No	No	No	No	Yes

Linkage was performed using probability matching techniques; therefore each pupil record was able to match to several different CHI records, with different linkage scores assigned to each potential match depending on how well identifiers agreed between the two records in question. In probabilistic matching, records are assigned overall linkage scores based on the level of agreement between individual identifying information on each record. A record linking with the highest linkage score is always accepted as the best match by the linkage algorithm and the next best match is called a 'rival' to that best link. Each pupil census record was assigned an 'optimal linkage' marker (0=unlinked; 1=linked) and a partition code (Table 9) during the linkage stage. The optimal marker described whether the pupil

census record had successfully linked to the CHI database. Records which didn't link to the CHI database had no health information to analyse and were immediately discarded.

The partition code was based on the linkage score and described the strength and quality of each match against the CHI database by providing information on the quality of each individual link and verifying how 'close behind' the next rival link lay. These codes allowed poorer quality rival links to the CHI database to be investigated and, where appropriate, be omitted. Partitions A-E described instances where a pupil record linked to a record on the CHI database with exactly the same date of birth, sex and postcode. Similarly partitions F-H described records that matched exactly by sex and date of birth and by the first six characters of postcode. Partitions I–J described records that matched exactly on sex and postcode but differed slightly on date of birth. Whilst rival matches to the main accepted link could vary in proximity from close to distant, in some cases there was no rival match at all, resulting in for example, assignment of partition matches A or F. Probability matching also allowed the possibility of two different pupil records linking to the same CHI record with exactly the same linkage score. In the case of tied matches, the first occurring CHI match was accepted as the true match and the second CHI match was discarded. An appropriate partition score was then assigned to the first linking record to highlight that a rival CHI at the same score existed. Such cases were highlighted by partitions E, H and K which represented tied matches. The absence of full pupil names or even initials reduced the quality of the linkage and increased the possibility of different pupils erroneously linking to the same CHI number during probability matching if they had identical or very similar dates of birth, postcode and sex. The most obvious case where pupils had identical dates of birth, postcode and sex was twins of the same gender living at the same address. Pupils linking to the same CHI resulted in additional, seemingly duplicate, records where two different pupils were given the same CHI number, and therefore subsequently the same study index number. Unique study index numbers were given to each pupil following the linkage of education data to health data and this study index number was common to each pupil across all of their education and health datasets.

Table 9 Partition code and description

A	Exact match, Unrivalled (next best score > 13 points below)
B	Exact match, Rivalled (Distant: next best score 5.5 to 13 points below)
C	Exact match, Rivalled (Intermediate: next best score 2.5 to 5.5 points below)
D	Exact match, Rivalled (Close: next best score > 0 to 2.5 points below)
E	Exact match, Rivalled (Tied match with another UPI)
F	Close match (6 character postcode match), Unrivalled (next best score > 12 points below)
G	Close match (6 character postcode match), Rivalled (next best score > 0 to 12 points below)
H	Close match (6 character postcode match), Rivalled (Tied match with another UPI)
I	Close match (match in 2 out of 3 DD/MM/YY), Unrivalled (next best score > 5.5 points below)
J	Close match (match in 2 out of 3 DD/MM/YY), Unrivalled (next best score > 0 to 5.5 points below)
K	Close match (match in 2 out of 3 DD/MM/YY), Rivalled (Tied match with another UPI)
L	Match on link weight > 24 points, Unrivalled (next best score > 9 points below)
M	Unlikely matches

4.3 Data and secondary linkage methodology

4.3.1 Education data

Education records were provided by ScotXed as seven separate files which could be matched to each other and to the supplied linked health data using study index number. Pupil census data contained individual pupil records across five school census years; 2009 to 2013. The other education datasets provided data pertaining to those census years. Attendance and exclusions data were not recorded in 2011 or 2013 due to a policy change, whereby from 2010 onwards these data were only collected every second year. Data on special educational needs (SEN) were recorded slightly differently prior to 2010 (section 4.3.1.2.3) and were therefore provided as two separate datasets, namely (i) Reason for Support and (ii) Student Need. Attainment data from SQA and leaver destination data made up the collection of education datasets provided.

The pupil census data formed the spine of the study cohort; therefore, records within all of the other datasets (both health and education) were appended onto the pupil census data. All of the datasets, including the pupil census data, were initially quality checked and, where necessary, de-duplicated to allow data to be re-matched together in an appropriate way. Once this process was complete, the various education and health datasets were manipulated to include either one record per pupil overall or alternatively one record per pupil within each school census year. These were then appended onto the pupil census data in a series of one-to-many matches using common study index number, or one-to-one matches using both the study index number and school census year. This process is explained in more detail in sections 4.3.1.1 and 4.3.1.2. Much of the initial data cleaning,

data manipulation and re-linkage of data files were undertaken using SPSS version 21. Later data manipulation, recoding and analyses were undertaken using Stata MP version 14.

4.3.1.1 Cleaning the pupil census data and matching of linked education data

Pupil records which failed to link to the CHI database (i.e. optimal=0) were supplied as part of the pupil census dataset but were immediately excluded from the analyses because they had no corresponding linked health information to analyse. Records which did link to the CHI database but which did not have an associated SMR02 birth record were also excluded. These steps limited the analyses to only children born in Scotland. Records without an SMR02 were excluded because final models were adjusted for a number of obstetric factors and this was not possible in the absence of those variables. Each pupil census record also had a corresponding 'partition score' produced during the linkage detailing the strength of the match against the CHI database. This was required to enable investigation of duplicate records within the pupil census data and potential bad links to the CHI database.

The pupil census is recorded annually, in September, at the start of every new school year and produces one record for each pupil currently attending a local–authority primary, secondary or special school in Scotland. Whilst each census year should contain one record per pupil, several duplicate records were present within census years in the supplied pupil census data. Some of these records were identical duplicates where all data items were the same on each record. Other duplicate pairs had the same study index number, census year and date of birth but had different values across other variables. These duplicates occurred for two reasons:

1. Records with duplicated study index number, census year and date of birth most likely occurred due to different pupils erroneously linking to the same CHI number if they had same sex and identical or very similar dates of birth and postcode. Identical study index numbers, for example, were present for same sex twins living at the same address.

2. Records duplicated on all data items, could also occur for the reason above but were more likely records that were erroneously duplicated on the original recording of the pupil census data.

The pupil census dataset was de-duplicated before proceeding with any analyses. The majority of pairs of duplicate records had partition scores E or H indicating that these records most likely belonged to sets of twins. In the absence of full names, each of these records would have simultaneously linked to two different CHI numbers at the same linkage score in the CHI database. However, in both cases, the first CHI number, and therefore the same CHI number, on the CHI file, would have been accepted by the linkage algorithm as being the correct match. The second of the two CHI numbers in each case, the tied rival CHI match, would have been discarded therefore leaving each twin linking to the same CHI but with only one link being correct. In the absence of full names, and in light of the resultant inability to accurately match same sex twins or higher order multiples living at the same address, I restricted the study cohort to singleton births. This was done by selecting all school records where the corresponding linked SMR02 maternity record indicated a singleton birth child. Residual records with partition score E or H were then excluded, regardless of whether they were part of a duplicate pair, as there was no way of deciphering which of these remaining twins had been assigned the correct CHI number and which one had been wrongly assigned.

All other instances where duplicated records contained the same partition were also deleted because it was impossible to choose the correct link. Occurrence of these duplicate partition scores similarly indicated residual twins, siblings or alternatively different children with very similar identifiers matching to the same CHI number with varying quality of rival matches being flagged. For duplicates that had different partitions, the record with the best partition was retained and the remaining duplicates were deleted. This scenario occurred where two different pupil records linked to the same CHI record but one linked at a higher score than the other. Finally, some residual exact duplicates were deleted. The pupil census data provided information on pupils' school stage and this ranged from P1 (first year of primary school) through to S6 (final year of secondary school) with stage coded as SP for children attending special schools. In order to strictly limit the cohort to pupils between first year of primary school and final year of secondary school and those attending special schools, a small number of pupils with the following additional codes were excluded from the cohort: AD (adult learner), S7 and S9. Finally some of the pupils on the school census were older than 18 years or younger than 5 years. Whilst most children start school at five years it is not uncommon for a child to start school aged four. Similarly, whilst most pupils have left school by 17 or 18 years of age, it is possible that some, through perhaps grade retention, may leave at 19. Records with ages less than 4 years or greater than 19 years were omitted.

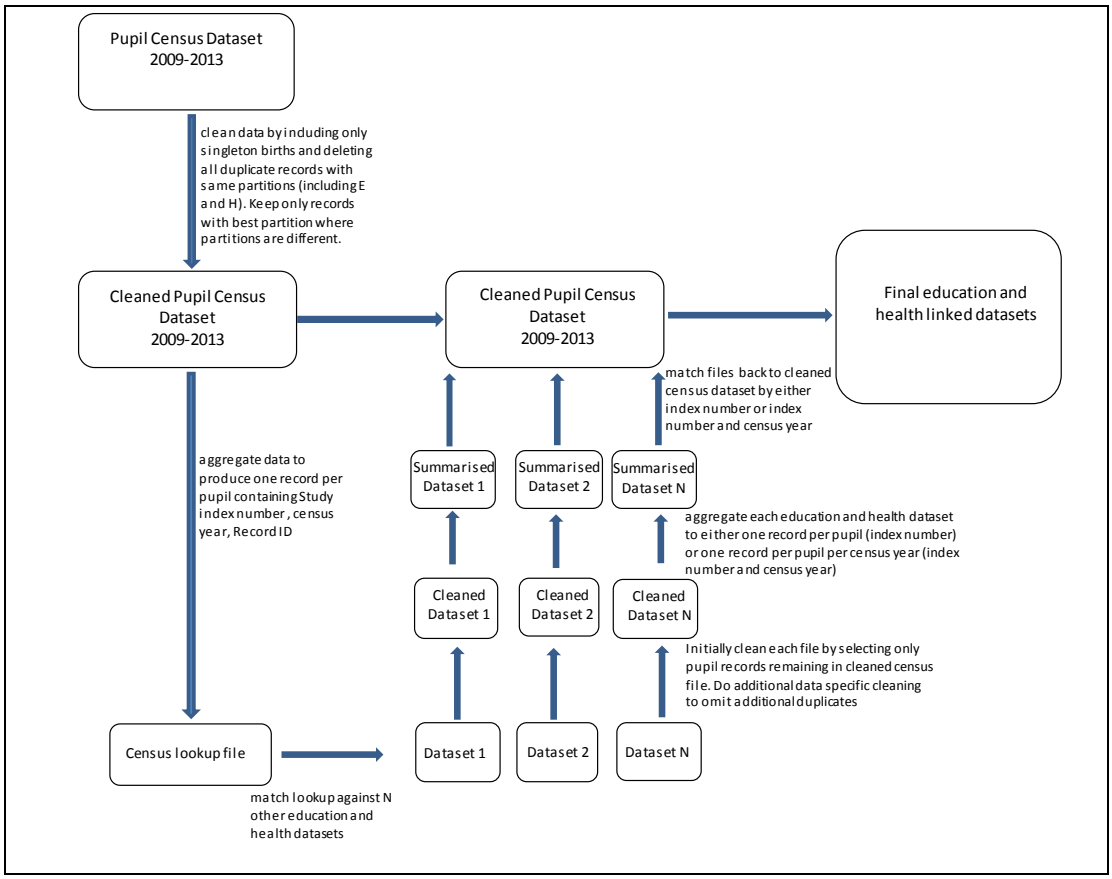
All records relating to the same pupil were given the same study index number during the linkage process. However, several records with the same index number had more than one date of birth recorded across different school census years. These discrepancies either occurred due to: residual bad links, where two different pupils on the school census had linked to the same CHI number and been given the same study index number; or due to a legitimate link where the pupil had one or more erroneous dates of birth recorded within their set of school census records. In cases where different dates of birth existed for the same pupil across different school census years I cross referenced the date of birth on each pupil census record against the date of birth recorded on the corresponding linked SMR02 record. The SMR02 record goes through regular quality checks and is deemed to be very accurate [252]; therefore, in cases where the dates of birth differed, but were within one year, the birth date on the pupil census was replaced with the SMR02 date of birth. Records where the dates of birth were more than a year apart were deemed to be residual poor links and were excluded.

All of the measures described above aimed to make the final datasets as clean and free of linkage errors as possible though some imperfections remained. For example, a few pupils had gaps in their recorded school census years between 2009 and 2013. Whilst some of these gaps may have been genuine, others could be attributed to inconsistencies within large scale administrative data. Pupils who had gaps in their recorded school census years were not treated differently and their records were analysed similarly to other pupils in the cohort.

4.3.1.2 Cleaning and matching the linked education data files

After removing duplicates and erroneous links to reduce the pupil census cohort to one record per pupil per census year, the next step involved creating a file of study index number, census year and unique record ID to be matched against all of the other education datasets. Whereas unique pupil IDs were created at the linkage stage and were used to create a study index number for each pupil, unique record IDs, supplied by ScotXed, were unique for every single record in the pupil census dataset regardless of whether or not they pertained to the same pupil. This step enabled pupil records, deleted from the pupil census file, to be deleted from the other education files too, thus reducing duplicate records and bad links in these datasets. Each of the other education files were then further cleaned and manipulated before being matched back to the full pupil census file. This process is illustrated in Figure 2.

Figure 2 Cleaning the pupil census data and associated health and education data file



Attendance data are collected at the time and then collated by the school at the end of every school year. Therefore, this dataset summarised absences for each pupil within each census year and was expected to contain one record per pupil per census year. Similarly, the leaver dataset was expected to contain one record per pupil because each pupil only leaves school once and has one destination recorded 6 months after leaving school. The other datasets varied in structure however some obvious conditions remained. For example, the attainment data were expected to contain no more than one record per census year per exam for each pupil. Exclusions and records of SEN are provided on a record by record basis and, within each school year, pupils can be excluded more than once and can be diagnosed with more than one type of SEN. Therefore, whilst it was considered acceptable for these data to contain more than one record per pupil per census year, they should not have contained any identical duplicates. The majority of duplicates and bad links in these datasets were removed by matching the lookup file of study index numbers, census years and unique record IDs from the cleaned pupil census and only retaining pupils who were present within that file. The study index numbers were derived directly from the original pupil census linkage to the CHI database and were common across all of the education datasets. Therefore, by removing duplicates and bad links in the census data it

was expected that this would significantly reduce the number of duplicate records in these datasets. Whilst this was indeed the case, a small number of residual duplicate records had to be deleted.

After removing duplicate records and bad links, each education dataset went through further data manipulation to make them suitable for appending to the pupil census. The absence, exclusion and SEN data were reduced to summarised datasets containing one record per pupil per census year and then appended onto the pupil census using study index number and census year. Similarly, the attainment and leaver data were reduced to summarised datasets containing one record per pupil and appended onto the pupil census using study index number (Figure 2). This matching process is described in sections 4.3.1.2.1 through 4.3.1.2.5.

4.3.1.2.1 Absence data

Attendance data were recorded at the time and then collated at the end of every school year by the school. Absence data used to be recorded annually; however following a policy change by the Scottish Government in 2010, these data were recorded every two years thereafter but only for the second year [253]. Absence data were therefore only recorded in census years 2009, 2010 and 2012. Despite this limitation, three years of absence data were still available for this project. The cleaned absence dataset contained one record per pupil per census year and each record consisted of several attendance and absence categories with a corresponding number of half days recorded against them. The absence data contained: authorised absences such as late starts, sickness, work experience and authorised family holidays; and unauthorised absences such as truancy, temporary exclusion and unauthorised family holidays. Guidance says that absence should be recorded as unauthorised until an explanation, for example an appropriate sick note, is given. All types of authorised and unauthorised absence were included in this thesis. The data also contained a summary of the total number of half days attended and the total number of half days which could possibly have been attended by the pupil throughout the whole census year.

On discussion with ScotXed it was agreed that late starts and half days spent on work experience should be added to recorded attendances and that external leave which occurred with parental and school consent should be deducted from the number of possible attendances. The total number of full day absences per census year was the outcome of

interest and was calculated for each pupil as: Number of days absent = (Number of possible half day attendances – Number of actual half day attendances) /2

Whilst the absence data contained several duplicates, none of these remained after matching the cleaned pupil census lookup file to the absence data to limit it to pupils on the cleaned version of the pupil census. Once the absence data were appended onto the full pupil census, an absence cohort was produced by firstly selecting only census years 2009, 2010 and 2012 and then further selecting census records with a linked absence record (those with a non-blank attendance year). Attendance records are submitted for all children; therefore, only a small number of census records did not have a corresponding absence record and this was attributed to linkage limitations. Finally a small number of records which had the number of possible attendances recorded as zero were excluded from the analyses.

4.3.1.2.2 Exclusion data

Exclusion data were recorded at the time and then collated along with attendance data by the school at the end of every school year. Similar to attendance data, changes in policy meant that from 2010 exclusion data were also recorded every two years but only for the second year. Exclusion data were therefore only available for census years 2009, 2010 and 2012. The dataset contained records pertaining to all children who had received one or more exclusions from school during census years 2009, 2010 or 2012 and provided start and end dates for each exclusion. The total number of exclusions within each census year was calculated for each pupil by aggregating the file by study index number and census year. Similarly the total length of exclusion throughout each census year was calculated by summing all individual exclusion lengths together for each pupil within each census year. After matching the pupil census look-up file to the exclusions data, the exclusions data were aggregated to contain one record per pupil per census year. Once the exclusions data were appended onto the full pupil census by study index number and census year, an exclusions cohort was produced by firstly reselecting only census years 2009, 2010 and 2012 and then further selecting records with a linked exclusion record (a non-blank attendance year). Exclusion data and absence data were both recorded as part of the attendance dataset therefore including only those records with a non-blank attendance year ensured that pupils with no exclusions were genuinely non-excluded pupils as opposed to pupils on the census who just had not linked to their corresponding attendance record. Records with numbers of possible attendances recorded as zero were excluded.

4.3.1.2.3 Special educational need data

Special educational need (SEN) data are recorded within the school census and children can be recorded as having more than one SEN in any given census year. SENs were recorded under a variable called 'Reason for Support' until 2010 and thereafter they were recorded under a variable called 'Student Need'. The codes used were broadly consistent over the study period; however, some additional SEN codes were introduced post 2010 which were not previously available. The SEN data were provided by ScotXed as two separate datasets in line with this change.

Children have SEN if they are unable to benefit fully from school education without help beyond that normally given to schoolchildren of the same age [254]. Both schools and local authorities have a statutory duty to identify children with SEN and provide support. Several recorded SEN codes were excluded from my definition because they were not relevant for the purposes of this particular study. Reasons for additional support excluded from the definition included interrupted learning, English as an additional language, looked after or more able pupils, young carer, bereavement, substance misuse, family issues, risk of exclusion, and reasons classed as other or not-disclosed. I included SEN attributed to learning disability, learning difficulty, sensory impairment, physical motor impairment, communication problems, autism, physical and mental health problems and social, emotional or behavioural difficulties. The full range of codes is shown in Table 10 including those that classified the child as having a SEN and those that did not meet the definition.

Table 10 List of special educational need (SEN) categories, codes and descriptions

SEN code	SEN description	Derived SEN group
10	Learning disability	Learning disability
11	Dyslexia	Learning difficulty
12	Other specific learning difficulty *(eg numeric)*	
13	Other moderate learning difficulty	
20	Visual impairment	Sensory impairment
21	Hearing impairment	
22	Deafblind	
23	Physical or motor impairment	Physical motor impairment
24	Language or speech disorder	Communication problems
44	Communication support needs*	
25	Autistic spectrum disorder	ASD
26	Social, emotional and behavioural difficulty	Social, emotional and behavioural difficulty
27	Physical health problem	Physical health problem
28	Mental health problem	Mental health problem
40	Interrupted learning	Not classed as SEN
41	English as an additional language	
42	Looked after	
43	More able pupil	
45	Young carer*	
46	Bereavement*	
47	Substance misuse*	
48	Family issues*	
49	Risk of exclusion*	
98	Not disclosed/declared	
99	Other	

* code only included in student need dataset from 2010 onwards

The reason for support and student need datasets were merged to create one dataset containing all SEN codes between 2009 and 2013 and the pupil census lookup file was matched to this to include only those pupils on the cleaned pupil census file. The majority of the codes which were only recorded after 2010 were not relevant and were excluded from the dataset. However, children classed as having communication support needs were included in the analyses despite this reason only being incorporated into the data collection post 2010. Several duplicate records were initially present; however, these disappeared after matching the pupil census lookup file to the SEN data. Before the SEN data were appended onto the full pupil census data, the SEN data were reduced and manipulated to contain one record per pupil per census year. Having re-classified SEN codes as described above, the SEN data were aggregated by study index number and census year. The data were further manipulated to contain one row of data per pupil per census year with all SEN codes recorded in that year listed horizontally. The SEN data were appended onto the full pupil census by study index number and census year.

Children were simply classed as having a SEN for the main analyses if they had a record of any of the SENs of interest in a given census year. Pupils with no SEN data or whose SEN code had been excluded were classed as not having a SEN. Specific SEN groups were also derived for subsequent analyses as highlighted in Table 10. Each SEN description was grouped into a broader SEN category to allow specific types of SEN to be investigated against each chronic condition. Children were classed as having each specific type of SEN if they had a record of any of the SENs of interest within that category in a given census year. Children with each of those types of SEN were compared against children who did not have any eligible SEN.

4.3.1.2.4 Leaver destination data

The leaver destination dataset contained data between 2009 and 2013. The pupil census lookup file was matched to the leaver destination dataset in order to omit pupils that had been deleted from the pupil census dataset; however, some duplicated records remained after this step. These were not exact duplicates as they differed either by leaving date, leaver status or both. Both records were deleted within all of these duplicated pairs as it was impossible to determine which record contained the correct information. The data were then aggregated by study index number to ensure that the dataset included one record per pupil. The data contained various codes classifying each pupil's destination six months after leaving school. These were derived in order to form more manageable grouped codes as shown in Table 11. On discussion with ScotXed it was agreed to treat leaver status as a dichotomous variable detailing whether pupils had successfully moved into either education, employment or training. Failure to move into education, employment or training is described by the education sector as having NEET status [255]. Leaver status was therefore categorised into two categories: 0= in education, employment or training (EET); 1= not in education, employment or training (NEET).

Table 11 Leaver destination category and associated NEET status

sldr_status	group	NEET status
Activity Agreement	Activity Agreement	No
Employed	Employed	No
Employed - with Training	Employed - with Training	No
Employed - without Training	Employed - without Training	No
FE Further Education	Further Education	No
Get Ready For Work	Training	No
HE Higher Education	Higher Education	No
Part-time Employment	Part-time Employment	No
Training	Training	No
Voluntary Work Voluntary work	Voluntary Work	No
Unknown	Unknown	Yes
Unemployed	Unemployed	Yes
Unemplyed Not Seeking Unemployed - not seeking unemployed - not seeking	Unemployed Not Seeking	Yes
Unemployed Seeking Unemployed - seeking	Unemployed Seeking	Yes
Economically Inactive	Economically Inactive	Yes
Excluded	Unknown - Exclude	omitted from analyses
Moved Outwith Scotland	Unknown - Exclude	omitted from analyses
Unavailable - Ill Health	Unavailable - Ill Health	omitted from analyses

4.3.1.2.5 Attainment data

The attainment data provided by ScotXed contained exam information for a range of full exams, modules, sub-modules, single unit awards and group awards for all pupils attending school between 2009 and 2013. The data contained course codes, course names and overall course results including results for every course module and sub module. Group awards and single unit awards were also included. Only some of these records were needed; therefore, the first task involved deciding which course codes were of significance and which could be omitted from the data. The second task involved deciding how best to summarise each pupil's attainment. For example, attainment could either have been analysed on a yearly basis or at the point of leaving school. The third task involved deciding how best to prepare a summarised attainment file from the original attainment dataset which could then be appended onto the full pupil census data. These tasks were all

inter-related as decisions on how best to structure and analyse the attainment data impacted how it would be cleaned and vice versa.

Pupils normally sit exams in fourth, fifth and sixth year of secondary school, abbreviated as school stages S4, S5 and S6. Therefore, only pupils who reached these stages were included in the analyses. A small number of records in the attainment dataset had student stage coded as S3 (third year of secondary school); however, these records were excluded from the attainment dataset. Pupils in special schools were also omitted from the analyses because their actual stage at school could not be identified. These children had their school stage recorded as SP and therefore were easily identified within the census data. I could have analysed the attainment data on a yearly basis, assessing each pupil's attainment at the end of each school term. Analysing the data in this way would have resulted in a longitudinal design and repeated measures on each pupil, whereby they would have contributed a minimum of one year and a maximum of three years of exam results to the analyses. An advantage of analysing the data in this way was that every year of exam data between 2009 and 2013 would have been eligible for analysis, regardless of school stage; therefore no additional data would have been omitted. The attainment measures that I planned to use, however, were designed to measure attainment on leaving school. These measures were provided by ScotXed and, on discussion, it was deemed that analysing attainment as a final measure on leaving school was a more relevant method.

A second option involved analysing attainment as a summation of all attainment on leaving school. This option reduced the dataset to a one record per pupil cohort. Data were available until 2013; therefore, attainment for all pupils who had left school by this point would have been summarised. Analysing attainment on leaving school had a limitation whereby valid attainment records belonging to pupils who had not yet left school by 2013 would not have been included in the analyses. A further limitation was apparent for earlier school leavers in the cohort. Pupils in S5 and S6 who left school in 2009 had attainment data pre-dating 2009 and this was not included in the dataset; therefore, these pupils were unable to contribute their full attainment data to the analyses. Similarly, pupils in S6 who left school in 2010 had attainment data pre-dating 2009. These particular subsets of pupils therefore would have to have been omitted from the analyses. Finally, a small number of pupils who had reached S6 did not have an accompanying leaver record. Solely basing the analyses on school leavers would have excluded these pupils from the analyses. The biggest limitation was that S4 and S5 leavers were not available for analyses in 2013 due to

the introduction of Curriculum for Excellence (a new examination system introduced across Scotland) [256].

The favoured option involved analysing each pupil's final summarised attainment at 2013. This also reduced the cohort to a one record per pupil model but had the advantage of including records for all pupils between S4 and S6 regardless of whether they had an official school leaver record. Records for S5/S6 pupils in 2009 and S6 pupils in 2010 still had to be omitted for the reasons given above. It also became apparent that I would have to omit some additional pupils to ensure that every child had the same opportunity to progress to S6 by 2013. This enabled them to contribute the same number of attainment years so as not to bias overall attainment. For example, pupils progressing through to S6 by 2013 could contribute 3 years of attainment data whilst pupils in S4 or S5 in 2012 could only contribute one or two years respectively with naturally lower attainment as a result. The exposure for each pupil was therefore adjusted by excluding particular stage and year combinations. As a result, S4 pupils in 2012 and 2013 and S5 pupils in 2013 were omitted from the analyses. Many of the excluded records pertaining to S4 pupils in 2012 included valid school leavers and this was a limitation; however, I decided that this was the most consistent and unbiased analysis plan out of the three discussed. The full range of census years and school stages that contributed to the final attainment measures are described in Table 12.

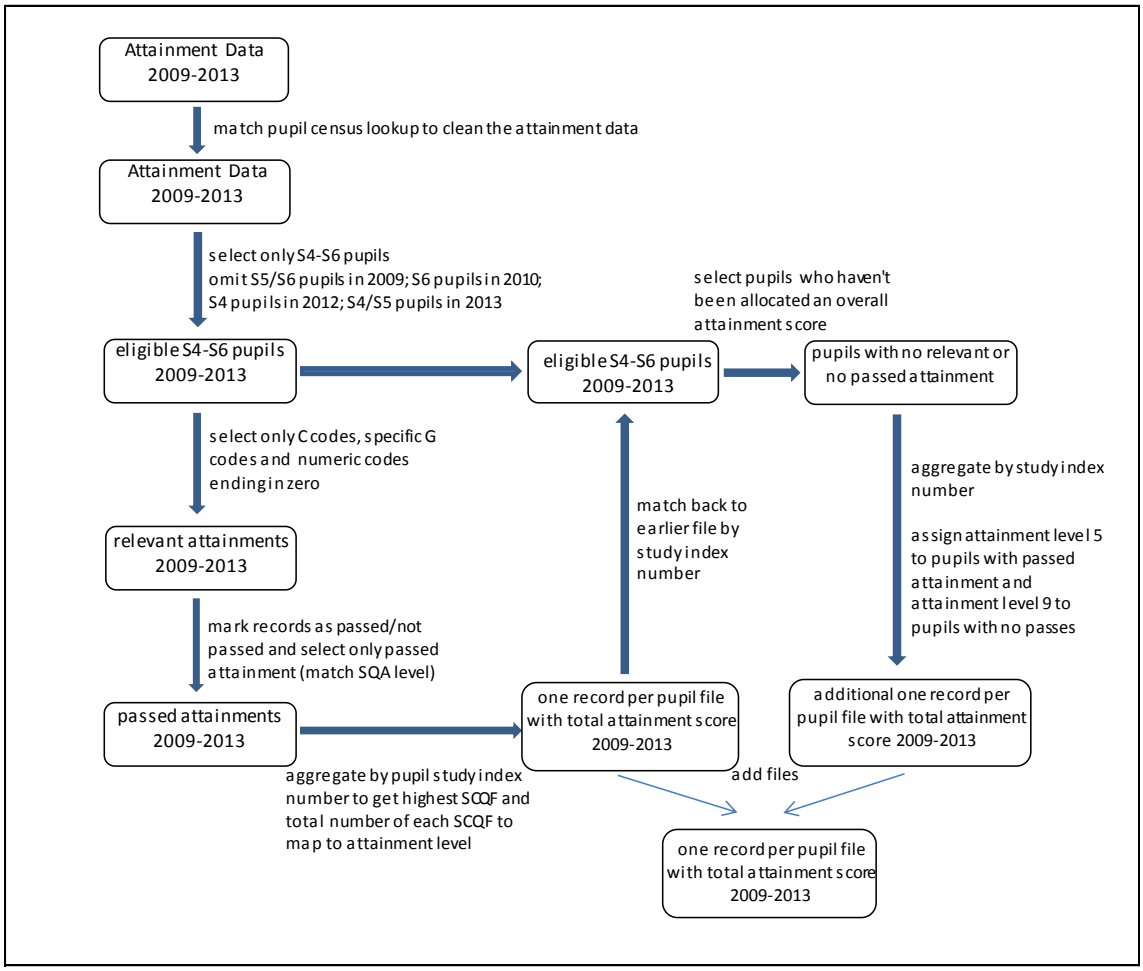
Table 12 Inclusion of pupils in attainment cohort by school stage and census year

Census Year	Stage		
	S4	**S5**	**S6**
2009	yes	no	no
2010	yes	yes	no
2011	yes	yes	yes
2012	no	yes	yes
2013	no	no	yes

The pupil census look-up file was matched to the attainment data in order to omit pupils that had been deleted from the census dataset. Subject names (English, maths etc.) and Scottish Qualifications Authority (SQA) course level descriptions (standard grade, higher etc.) were then appended onto the attainment data using a look-up file as these were not provided by ScotXed. The ultimate aim was to produce a one record per pupil dataset containing an overall summarised attainment score for each pupil and this process is illustrated in Figure 3.

In order to create this file, the data were initially cleaned by selecting relevant attainment codes. The first stage involved limiting the data to include only pupils in S4, S5 and S6. Pupils in particular stages not permitted to contribute to the attainment calculation were then excluded from the attainment dataset. The following pupils were omitted: S5/S6 pupils in 2009; S6 pupils in 2010; S4 pupils in 2012; S4/S5 pupils in 2013. The data from ScotXed contained all examination information for all pupils sitting exams including main exam modules, sub modules, administrative examination codes, individual awards and group awards. Only some of the records provided were needed to enable the relevant attainment information to be derived. Several other records were excluded because they only provided attainment relating to, for example, sub-modules or components of main exam modules or administrative information on exams taken. Retaining these records would have resulted in erroneous counting of additional non-relevant attainment. Standard grade exams, for example, were recorded using four-character numeric codes but resulted in five separate pupil records, only one of which was relevant to this thesis. Records with codes ending in zero contained attainment information relating to the main exam of interest (e.g. 0300 biology) and were retained whilst records containing other numeric codes (e.g. 0301-0304) related to attainment within exam components such as problem solving, knowledge and understanding and practical skills and were excluded from the attainment dataset. The other types of exam followed a similar pattern. Relevant course codes for all other types of exam (Intermediate 2 through to Advanced Higher) began with the letter C and these records were retained since they provided information on attainment from the main exam of interest. Records containing attainment within additional sub modules for these exam types began with the letters D, F and H and were deleted from the attainment data. Additionally, all records containing exam codes beginning with the letter X were deleted as these were external assessment codes for administration purposes and were not of interest for this project. A few records with course codes beginning with the letter ‘G’ remained and these related to group awards called Baccalaureate awards. The introduction of Curriculum for Excellence meant that a number of new Baccalaureate awards were introduced in 2013. To maintain consistency, I only retained Baccalaureate awards which were available across all of the census years during the study period. I therefore retained course codes relating to Baccalaureate awards in languages (G9CW), sciences (G9CX), social sciences (GE9W) and expressive arts (GE9X).

Figure 3 Process to clean the attainment data and create a summarised attainment dataset



In order to identify successfully attained grades rather than failed exam results, only records with results which related to passes were retained. Pupils who passed their attainment were therefore flagged in the attainment dataset. Results which constituted a pass were represented by grades 1, 2, 3, 4, 5, 6, 77 (grade D at higher), P (pass) and D (distinction). Some records had exam result codes relating to failed attainment: 7 (grade 7 for standard grade), 8 and 9 (no award), 999 (no attempt) and E (no result). A very small percentage of records (0.1%) had no corresponding result recorded for the exam and these records were omitted from the attainment data. All pupils who appeared on the attainment dataset and who had passed an exam in one or more relevant full courses (standard grade, higher etc.) were coded into four categories of attainment based on their number of awards attained and the quality of those awards. The quality of academic award was derived based on the result of each exam and the corresponding SQA level of the particular course. SQA levels included Standard Grade, Access 1-3, Intermediate 1-3, Higher and Advanced Higher. Each pupil's exam results and the corresponding SQA levels for these results were combined to obtain a corresponding Scottish Credit Qualifications Framework (SCQF) [257]

level for each award. SQA levels and results mapped to corresponding SCQF levels as shown in Table 13. The SCQF level of each award was used in a final algorithm to obtain an overall attainment score for each pupil. The file was then aggregated by study index number to produce one summarised attainment record for each pupil. Several variables were retained and new ones derived for use in this one record per pupil summarised dataset. The aggregated dataset contained the total number of attainments at each SCQF level for each pupil. This was used to derive an ordinal four-category attainment variable: 1 (high); 2 (general/broad); 3 (basic); 4 (low). Children with more than one attainment at level SCQF 7 or three or more attainments at SCQF level 6 were coded as having high attainment. Children with more than seven attainments at SCQF level 4 or more than five attainments at SCQF level 5 or less than three attainments at SCQF level 6 were coded as having general/broad attainment. Children with five or more attainments at SCQF level 3 or two or more attainments at SCQF level 4 or four or less attainments at SCQF level 5 were coded as having basic attainment. Finally, children with one or more attainments at SCQF level 2 or less than five attainments at SCQF level 3 or less than two attainments at SCQF level 4 were coded as having low attainment. This mapping is summarised in Table 13.

Table 13 Algorithm to map SQA level and result to SCQF level and final four category attainment level

			Attainment Level			
SQA level code	SQA level (result)	SCQF level	low	basic	broad	high
8	Access 2	2	>=1			
9	Access 3	3	<5	>=5		
1	Standard Grade (foundation 5 or 6)	3				
10	Intermediate 1	4	<2	>=2		
1	Standard Grade (general 4 or 5)	4			>7	
11	Intermediate 2	5		<=4	>=5	
1	Standard Grade (credit 1 or 2)	5				
12	Higher	6			<3	>=3
13	Advanced Higher	7				>=1
47	Baccalaureate	7				

The four-category measure of attainment was calculated based on pupils' number of passed attainments within relevant courses as described above. This attainment measure therefore did not include pupils who had either failed all of the exams they had sat or had only passed very low attainment courses such as a single module. In order to identify pupils who had failed all of their exams, the final one record per pupil attainment dataset was appended onto an earlier version of the attainment data which contained all relevant

attainments regardless of result. Pupils with relevant attainment but without any passes were identified and given an attainment value equal to 9 and a summarised one record per pupil dataset was created. Both of these files were then appended onto the full original attainment dataset to identify remaining pupils who only had attainment records of a standard below those deemed to be relevant. These pupils were assigned an attainment category of 5 to highlight that they failed to achieve even the lowest level attainment. These three datasets each contained one record per pupil and were added together and appended onto the full pupil census dataset using study index number. Pupils with passed relevant attainment, failed relevant attainment and no relevant attainment were initially coded separately in order to allow me to discuss the number of pupils in each category with ScotXed and therefore sense check the data and decide how best to deal with pupils falling into the latter two categories.

Once the summarised attainment data had been appended onto the pupil census, the specific census year and student stage combinations described in Table 12 were re-selected. The census data contained one record per pupil per census year whereas the attainment data contained one summarised record per pupil. This introduced the problem of which values to use for variables such as age and deprivation category which could change across different census years as pupils grew older or moved house. The last census record for each pupil was selected and therefore the last recorded age and deprivation category was used in the analyses. Selecting the cohort in this way meant that pupils in special schools were omitted from the analyses as their specific stage at school could not be identified. Children who had relevant attainment but had failed all exams were excluded from the analyses. Children who had passed low levels of attainment were recoded into the lowest attainment category. Remaining children without any form of attainment record were excluded from the attainment analyses. Residual pupils who had no attainment records despite being of the correct age to sit exams possibly occurred for four reasons.

1. Pupil was present on the initial September school census but left throughout the year and before sitting exams in the following May.

2. Pupil had attainment records which were not recorded on the SQA dataset. This could occur where a pupil was present on the September school census but moved to a private non-local authority school during the school year before sitting exams within that setting.

3. Pupil had attainment which did not link to the census due to linkage limitations.

4. Pupil truly had no attainment because they did not sit any exams.

4.3.2 Prescribing data

4.3.2.1 Cleaning the prescribing data

Prescribing data were provided by eDRIS and contained information on all medications prescribed from 2009 onwards for the chronic conditions of interest for this project namely diabetes, asthma, epilepsy, ADHD and depression. Only prescriptions pertaining to children present on the school census between 2009 and 2013 who had linked to the CHI database were provided. Furthermore prescibing data were only provided for medications likely to be prescribed for the chronic conditions specified above. Drugs likely to be prescribed for each of the chronic illnesses appeared in different sections of the British National Formulary (BNF) and children issued prescriptions within those BNF drug groups were defined as having those chronic conditions. After discussion with a pharmacist and clinical pharmacologist, all prescriptions within each of the following groups were provided by eDRIS.

1. BNF Chapter 6 (Diabetes)

2. BNF Chapter 3 (Asthma)

3. BNF section 4.8 (Epilepsy)

4. BNF section 4.4 (ADHD)

5. BNF section 4.3 (Depression)

The pupil census formed the spine of the cohort and, similar to the education datasets, the prescribing data were appended onto the cleaned pupil census data after duplicates had been removed and the data had been summarised in an appropriate way. The census look-up file was linked to the PIS data to select only prescriptions relating to those children on the cleaned version of the school census. For this project, children were defined as having a chronic disease of interest in a particular census year if they received medication within the relevant BNF chapter/section during that census year and satisfied additional condition

specific criteria (section 4.3.2.2). Each prescription had a corresponding prescribing date and this was used to derive the calendar year in which the medication was prescribed. The school census is taken in September every year, therefore, for each census year, medications prescribed between the preceding January and the following December of the same year were used to identify children with a chronic illness of interest. The first and last census years that each pupil was present on the school census between 2009 and 2013 were calculated. Children's first and last census years were compared against the calendar year of all of their respective prescription dates. Only prescriptions taken by pupils whilst at school were of relevance; therefore, medications prescribed before a pupil had started school or after a pupil had left school between the years 2009 and 2013 were excluded from the analyses.

The prescribing data contained several prescribing records per pupil within each census year. Therefore the data were reduced to a one record per pupil per census year summary dataset before being appended onto the pupil census data by study index number and census year. This summarised dataset contained derived counts of numbers of relevant disease-specific prescriptions received by each pupil within each census year. Children who met the criteria outlined in section 4.3.2.2 were defined as having each chronic condition. This dataset also included yearly counts of other less-relevant medications prescribed to pupils within the same disease specific BNF chapters. Pupils who only received these types of medications were excluded from the study because it was hard to decipher whether or not they had the chronic condition of interest. This was done to ensure that children with each chronic condition were compared against a cohort of children not receiving any drugs prescribed for that condition. The process is explained in more detail in the relevant chronic disease sections (4.3.2.2.1 to 4.3.2.2.5).

Prescribing data were appended onto the pupil census data using study index number and census year. Some pupils appeared to have a chronic illness marker in one year but not the next and vice versa. This occurred either because they truly had not received medications in certain years or due to linkage limitations. The chronic disease specific flags were therefore prospectively or prospectively and retrospectively assigned to pupil records for other census years as explained in section 4.3.2.2.

4.3.2.2 Identifying chronic illness using prescribing data

4.3.2.2.1 Diabetes

Children were defined as having diabetes in a given school census year if they cashed at least one prescription for insulin during the calendar year containing their school census date (yellow highlight in Table 14). Diabetes diagnoses were prospectively applied to all census years following a pupil's first insulin prescription in the period 2009-2013.
Children who were not issued with a prescription for insulin across any of the census years but who were prescribed metformin or 'other' drugs from BNF section 6.1.2 in at least one census year were excluded from the analyses comparing children receiving diabetes medication against children not receiving diabetes medication.

Table 14 Prescriptions used to define diabetes

BNF section	Category	Drug
6.1.1	Insulin	Insulin Aspart Insulin Detemir Insulin Glargine Biphasic Isophane Insulin Biphasic Insulin Aspart Insulin Lispro Soluble Insulin (Neutral Insulin) Isophane Insulin Biphasic Insulin Lispro Insulin Glulisine
6.1.2	Metformin	Metformin Hydrochloride
6.1.2	Other not commonly used	Pioglitazone And Metformin Hydrochloride Vildagliptin And Metformin Hydrochloride Gliclazide Liraglutide Glibenclamide Exenatide Sitagliptin Pioglitazone Glimepiride Dapagliflozin Repaglinide Lixisenatide Glipizide Saxagliptin Linagliptin Tolbutamide

[1] received medications included in the definition of diabetes are highlighted in yellow

4.3.2.2.2 Asthma

Children were defined as having asthma in a given census year if they cashed either

1. More than one prescription during the calendar year containing their school census date for an inhaled steroid prescribed with either an inhaled long acting or short acting beta agonist (yellow highlight in Table 15)

2. More than one prescription during the calendar year containing their school census date for an inhaled steroid (green highlight in Table 15) and more than one prescription during the same calendar year for either an inhaled long acting or short acting beta agonist (blue highlight in Table 15)

Asthma diagnoses were prospectively applied to all census years following the first year in the period 2009-2013 where conditions 1 or 2 above were met. Children who were issued with one or fewer prescriptions in 1 or 2 above in a given school census year or who received only leukotriene receptor antagonists (LTRA) or drugs within the 'other not commonly used' category in Table 15 were not defined as having asthma in that year. Children who did not satisfy conditions 1 or 2 in any of the census years but who were prescribed LTRAs or 'other not commonly used' drugs in at least one census year were excluded from the analyses comparing children receiving asthma medication against children not receiving asthma medication.

Table 15 Prescriptions used to define asthma

BNF section	Category	Drug
3.1 Bronchodilators	Short acting beta agonist	Salbutamol Terbutaline Sulfate Salbutamol With Ipratropium Fenoterol Hydrobromide
	Long acting beta agonist	Salmeterol Formoterol Fumarate Bambuterol Hydrochloride Indacaterol
3.2 Corticosteroids	Steroid	Beclometasone Dipropionate Fluticasone Propionate Ciclesonide Mometasone Furoate Budesonide
	Steroid with long acting beta agonist	Budesonide With Formoterol Fumarate Beclometasone Dipropionate And Formoterol Fluticasone Propionate And Formoterol Fluticasone Furoate And Vilanterol Salmeterol With Fluticasone Propionate
3.3	Leukotriene receptor antagonist (LTRA)	Montelukast Zafirlukast
3.1 - 3.3	Other not commonly used	Theophylline[1] Ipratropium Bromide[1] Tiotropium[2] Orciprenaline Sulfate[2] Ephedrine Hydrochloride[2] Aminophylline[2] Aclidinium Bromide[3] Glycopyrronium Bromide[3] Sodium Cromoglicate[4] Nedocromil Sodium[4]

[1] received medications included in the definition of asthma are highlighted in blue, green and yellow

[2] from section 3.1 but not commonly used so classed as other

[3] from section 3.2 although not a corticosteroid

[4] from section 3.3 but not commonly used so classed as other

4.3.2.2.3 Epilepsy

Children were defined as having epilepsy in a given census year if they cashed at least one prescription from BNF section 4.8 during the calendar year containing their school census date (yellow highlight in Table 16). Epilepsy diagnoses were prospectively applied to all census years following a pupil's first prescription from BNF section 4.8 in the period 2009-2013. No pupils were excluded from the analyses comparing children receiving epilepsy medication against children not receiving epilepsy medication.

Table 16 Prescriptions used to define epilepsy

BNF section	Category	Drug
4.8	All	Sodium Valproate Lamotrigine Carbamazepine Levetiracetam Topiramate Gabapentin Clobazam Pregabalin Clonazepam Ethosuximide Phenytoin Oxcarbazepine Lacosamide Midazolam Hydrochloride Zonisamide Phenobarbital Rufinamide Vigabatrin Perampanel Primidone Stiripentol Eslicarbazepine Retigabine Tiagabine

[1] received medications included in the definition of epilepsy are highlighted in yellow

4.3.2.2.4 Attention deficit hyperactivity disorder

Children were defined as having ADHD in a given census year if they cashed at least one prescription from BNF section 4.4 during the calendar year containing their school census date (yellow highlight in Table 17). ADHD diagnoses were prospectively and retrospectively applied to all census years in the period 2009-2013 if a pupil received at least one prescription from BNF section 4.4 in any census year. No pupils were excluded from the analyses comparing children receiving ADHD medication against children not receiving ADHD medication.

Table 17 Prescriptions used to define attention deficit hyperactivity disorder

BNF section	Category	Drug
4.4	Atomoxetine	Atomoxetine
4.4	Stimulant	Methylphenidate Hydrochloride Dexamfetamine Sulfate Lisdexamfetamine Dimesylate

[1] received medications included in the definition of ADHD are highlighted in yellow

4.3.2.2.5 Depression

Children were defined as having depression in a given census year if they cashed at least one relevant prescription during the calendar year containing their school census date: a tricyclic antidepressant from BNF section 4.3.1; a selective serotonin reuptake inhibitor (SSRI) from BNF section 4.3.3; one of the serotonin norepinephrine reuptake inhibitors (SNRI) commonly classed alongside SSRIs namely mirtazapine or venlafaxine from BNF section 4.3.4 (yellow highlight in Table 18). Depression diagnoses were applied prospectively and retrospectively to all census years in the period 2009-2013 if a pupil satisfied the definition for depression in any census year. Children who did not satisfy the definition above across any of the census years but who were prescribed alternative medications from BNF sections 4.3.2 or 4.3.4 in at least one census year were excluded from the analyses comparing children receiving medication for depression against children not receiving medication for depression.

Table 18 Prescriptions used to define depression

BNF section	Category	Drug
4.3.1	Tricylic drugs	Amitriptyline Trazodone Hydrochloride Nortriptyline Imipramine Hydrochloride Clomipramine Hydrochloride Lofepramine Dosulepin Hydrochloride Doxepin Trimipramine Amitriptyline Hydrochloride With Perph Mianserin Hydrochloride
4.3.3	Selective Serotonin Reuptake Inhibitors (SSRI)	Citalopram Fluoxetine Sertraline Escitalopram Paroxetine Fluvoxamine Maleate Mirtazapine[2] Venlafaxine[2]
4.3.2 and 4.3.4	Monoamine oxidase inhibitors and other	Duloxetine Flupentixol Moclobemide Reboxetine Agomelatine Phenelzine Tranylcypromine Tryptophan

[1] received medications included in the definition of depression are highlighted in yellow

[2] Serotonin Norepinephrine Reuptake Inhibitor (SNRI) from section 4.3.4 but commonly classed as an SSRI

4.3.3 Maternity, hospital admission and death data

4.3.3.1 Cleaning and matching maternity data to the pupil census

Maternity records (SMR02) were provided by eDRIS for all pupils who had successfully linked to the CHI database and these were appended onto the cleaned pupil census data using study index number. The SMR02 data did not contain any duplicate records therefore no additional data cleaning was required to enable the data to be correctly appended onto the pupil census data. SMR02 records are generated every time a mother is admitted to an obstetric ward either to deliver a baby or for events pre and post-delivery. The SMR02 records provided by eDRIS included maternity records for pupils on the pupil census. One SMR02 record is generated for each delivery regardless of the number of babies born; therefore twins and multiple births are presented on the same SMR02 record. Whilst

linking twins and multiple births to other datasets can sometimes be problematic, the pupil census data had been limited to singleton births during the original de-duplication process; therefore, the SMR02 records could easily be appended using study index number.

Information from the SMR02 maternity records allowed the various multivariate analyses to be adjusted for a number of obstetric and birth factors and maternal antecedents. Maternity variables used in the multivariate analyses were maternal age, estimated gestational age, birthweight, maternal smoking during pregnancy, Apgar score, parity and mode of delivery. Several of these covariates were recoded and re-categorised and missing values were managed where appropriate. Maternal age at delivery was categorised into four groups namely: aged less than 25; 25-29; 30-34; and over 34 years. The definition of maternal parity is the number of times that a woman has previously delivered a live or stillborn foetus with an estimated gestational age greater than 24 weeks. Women who have never previously given birth are described as nulliparous. Parity was categorised as: 0=nulliparous; 1=parous; 2=multiparous. Apgar score measured at 5 minutes evaluates the physical condition of a new born infant and is an acronym for: appearance, pulse, grimace, activity and respiration. Scores range between 1 and 10 and healthier babies have higher scores. This variable was categorised as 1-3, 4-6 or 7-10. Birthweight, gestational age and baby gender were used to derive age and sex specific birthweight centiles for the study population. Gestational ages less than 24 weeks were grouped together as were gestational ages greater than 43 weeks. Maternal smoking described whether the mother smoked at any point during her pregnancy and this was coded as yes, no or not known. Mode of delivery was categorised on SMR02 as: spontaneous vaginal delivery; cephalic delivery; assisted vaginal delivery; breech delivery; elective caesarean section; emergency caesarean section; other. Sociodemographic covariates used in the various multivariate models were recorded on the pupil census data. Demographic confounders included pupils' age in each census year, pupil ethnicity and area deprivation derived from postcode of residence using the Scottish Index of Multiple Deprivation (SIMD) [258] for 2012 whereby pupils were allocated into general population quintiles. The SIMD is derived from 38 indicators across 7 domains (income, employment, health, housing, geographic access, crime and education, and skills and training) using information collected for data zones of residence (median population, 769). Pupil age was categorised into three groups namely <11, 11-14 and >14 years. Ethnic group was categorised into white, Asian, black/African/Caribbean, mixed, other and not known.

4.3.3.2 Cleaning and matching acute and psychiatric hospital admissions and deaths to the pupil census

Acute hospital admissions (SMR01), psychiatric hospital admissions (SMR04) and mortality (NRS) data were provided by eDRIS for all children who had successfully linked to the CHI database. These datasets were used to provide health outcomes over follow up for pupils on the school census. School census data were available between 2009 and 2013; therefore mortality and admission data were also limited to include only data post 2009. The dataset supplied by eDRIS, referred to as the morbidity and mortality file from this point onwards, contained all acute and psychiatric admissions and deaths from 2009 onwards.

The school census look-up was matched to the morbidity and mortality file to identify subsequent admissions and deaths for pupils within the final de-duplicated and cleaned school census dataset. A series of files containing study index numbers and corresponding dates of first prescriptions within each chronic disease group were produced using linked pupil census and prescribing data and these were appended onto the morbidity and mortality file. This was done to enable identification of admissions and mortality for each child following their first relevant prescription in the period 2009-2013. First prescription dates were identified where pupils had received medication meeting the criteria for a diagnosis of diabetes, asthma, epilepsy, ADHD or depression based on the definitions contained in section 4.3.2.2. Children with no such diagnoses had no prescription dates and so admissions and mortality following their first school census date between the years 2009-2013 were identified instead. The earliest school census date was 21st September 2009 therefore admissions and deaths were censored on 21st September 2014 to produce a maximum follow up period of 5 years.

Flags were produced to identify whether a pupil had been admitted to hospital or died over the follow up period and, where applicable, the date of their first admission and date of death were identified. The total number of all-cause acute (SMR01) and psychiatric (SMR04) hospital admissions were summed over the duration of the follow up period and added together for each child to produce a count of the total number of admissions to hospital. Similarly the total length of stay in hospital over the follow up period pertaining to all-cause acute and psychiatric admissions was calculated for each child. Length of stay on SMR is coded as zero days if a patient is admitted and discharged on the same day; therefore, one day was added to each recorded length of stay in order to count day case admissions and hospital stays of less than one day.

This process was repeated to investigate risk of selected cause-specific admissions relating to each condition. This was done by identifying the most common cause of admission for each condition and following the same process for all-cause admissions described above. SMR01 and SMR04 record primary and secondary causes of admission using the International Statistical Classification of Diseases, 10th Revision (ICD10) [259] and include up to six diagnostic codes per episode. The primary diagnostic code is used as the indicator for cause of admission. The main causes of admission were investigated for pupils who had been defined as having each of the five chronic conditions. Each condition was then investigated to assess its impact on risk of cause-specific admission most commonly associated with that condition. The most common ICD10 recorded cause of admission for pupils defined as having each of the five chronic conditions were as follows:

Diabetes: E00-E99 - Endocrine, nutritional and metabolic diseases; R00-R99 - Symptoms, signs and abnormal clinical and lab findings not elsewhere classified

Asthma: J00-J99 - Diseases of the respiratory system

Epilepsy: G00-G99 - Diseases of the nervous system

ADHD: S00-T98 - Injury, poisoning and certain other consequences of external causes

Depression: S00-T98 - Injury, poisoning and certain other consequences of external causes; F00-F99 - Mental and behavioural disorders

Admissions related to abnormal clinical and lab findings not elsewhere classified (R00-R99) were combined with admissions related to endocrine, nutritional and metabolic diseases (E00-E99). This was done because these types of admission often occur for diabetic children when they experience complications related to diabetic control.

4.4 Statistical methods

4.4.1 Yearly educational outcomes (absences, exclusions and special educational need)

After appending all of the education datasets to the main pupil census dataset by study index number or a combination of study index number and census year, five initial analyses files were produced, one for each educational outcome of interest from absences through to attainment. These contained all of the variables necessary and were in the appropriate format to conduct the various multivariate analyses. Pupil outcomes that were analysed on a yearly basis included number of absences, number of exclusions and record of SEN. These longitudinal datasets contained repeated measures on the same pupil across several census years dependant on how many years they attended school between 2009 and 2013. Therefore appropriate analytical methods were chosen to account for the correlations between observations relating to the same pupil repeated across different years. Generalised estimating equations (GEE) were used because they accurately analyse longitudinal data by suitably accounting for within-subject correlation using a user-defined correlation structure and enable population averaged effects to be investigated [260-262]. Residual clustering can be adjusted for using robust standard errors and I incorporated these into each of my analyses.

4.4.1.1 Correlation structure

Before performing the various generalised estimating equation models, the user written quasi-likelihood under the independence model criterion (QIC) [263] statistic was used to determine the most appropriate correlation structure between repeated measures for the same pupil. Several correlation structures were investigated namely stationary, non-stationary, fixed, autoregressive, exchangeable and independent. Only independent, exchangeable and unstructured correlations were suitable descriptors of the datasets in this project because some pupils had less than two observations and observations were unequally spaced throughout time [260]. The most appropriate correlation structure for each of the datasets was an independent structure. This was used when running the longitudinal generalised estimating equations for each of the yearly outcomes.

4.4.1.2 Generalised Estimating Equations

Associations between sociodemographic and maternal characteristics of pupils and their chronic disease status (yes/no) were investigated via cross tabulations and p values. One of the assumptions of chi-square tests is that all observations or measurements are independent of each other. Therefore, chi-squared tests could not be used to investigate relationships within any of the longitudinal cohorts used in this thesis because they included repeated measurements on the same pupil across multiple census years. A series of univariate longitudinal generalised estimating equations were therefore employed to investigate associations between sociodemographic and maternal characteristics and chronic disease (yes/no). The chronic disease outcomes were dichotomous (yes/no) therefore I used generalised estimating equations with a binomial distribution and logit link function. Univariate generalised estimating equations with an appropriate correlation structure and an appropriate distribution were then performed to investigate the association between sociodemographic and maternity characteristics and each of the yearly outcomes (absences, exclusions and SEN). The distribution chosen was dependant on the nature of the outcome as described in sections 4.4.1.3 and 4.4.1.4. Univariate associations between sociodemographic and maternal characteristics and number of days absent from school were investigated using generalised estimating equations with a negative binomial distribution and log link function. For the purposes of creating initial demographic tables, I summarised exclusions by whether children had ever been excluded (yes/no) as opposed to summarising the number of exclusions per year. This was done because only 2% of the cohort had ever been excluded (section 4.4.1.3). Therefore, univariate associations between sociodemographic and maternal characteristics and whether pupils' had any record of SEN or were ever excluded at school were investigated using generalised estimating equations with a binomial distribution and logit link function.

Univariate longitudinal generalised estimating equations with an appropriate correlation structure and an appropriate distribution were next performed to investigate the impact of each chronic condition (yes/no) on each of the yearly outcome measures (absences, exclusions and SEN) as explained in sections 4.4.1.3 and 4.4.1.4. These models were next adjusted for sociodemographic confounders: age category, sex, SIMD 2012 quintile and ethnicity. I further adjusted the models for maternal confounders: maternal age, sex and gestation specific birth weight centile, estimated gestational age, mode of delivery, Apgar score, maternal smoking during pregnancy and parity. Univariate interactions with age category, sex and SIMD 2012 deprivation quintile were investigated for each chronic

disease status against each outcome. Where interactions existed, univariate and multivariate subgroup analyses by age category, sex and SIMD 2012 deprivation quintile were performed. Categories of variables believed to be at less risk of adverse outcomes were chosen as the reference category in all analyses. Children may experience more than one condition; therefore, I reran the final multivariate models including all conditions together. Each chronic condition specific analysis was therefore finally adjusted for presence of any of the other identified chronic conditions in addition to sociodemographic and maternal confounders. Where appropriate, output was tabulated and presented graphically in the form of forest plots.

4.4.1.3 Absence and Exclusion

The final absence and exclusion datasets each contained one record per pupil per census year. The predictors were presence of any of the five conditions of interest in turn. The outcomes were the number of days absent and the number of exclusions during a given census year. Both of these outcomes were count variables and therefore could be modelled using count data methods such as Poisson or negative binomial regression. Histograms of the number of days absent and the number of exclusions per census year were both observed and summary statistics were produced. It was apparent from these that the outcomes were more appropriately modelled by a negative binomial distribution because the mean and variance were not equal. Different pupils contributed different numbers of census years of data to the models depending on how many years they had attended school between 2009 and 2013. Children who contributed more census years of data had more opportunity to be absent or be excluded. Similarly, within each census year, pupils who had more possible attendances recorded had potential to have higher numbers of absences and more exclusion. An offset variable was therefore used in order to adjust for the exposure time for each pupil. The number of possible attendances per school term was used as an appropriate offset variable since absences and exclusions were analysed on a yearly basis and this adjusted for differing exposure times for each pupil within each census year. Generalised estimating equations with an independent correlation structure, a negative binomial distribution and a log link were employed using the *xtgee* command in Stata to investigate each chronic condition against the respective outcomes of number of absences per census year and number of exclusions per census year.

Around 98% of the cohort had never been and excluded and therefore, given the very high percentage of children with no exclusions, alternative methods of analysing the exclusions

data were considered. These included using a zero-inflated Poisson (ZIP) or zero-inflated negative binomial (ZINB) model. Both of these models can be used to describe count data with a large number of zero counts but cannot be used when employing GEE analyses within a repeated measures framework. I also considered whether to re-categorise exclusions into a dichotomous variable detailing whether a pupil had ever been excluded (yes/no) during a particular census year. I finally considered analysing the number of exclusions as a categorised ordinal variable with an upper cut off: 0, 1, 2, 3, 4, or 5 or more exclusions. However, I concluded that re-categorising exclusions in both of those ways would lose valuable information on the distribution of exclusions. Secondly, whilst only a small percentage of pupils had ever been excluded, the number of excluded pupils was still large due to the very large population cohort. Finally, the bigger issue, above zero inflation, was to adjust for within-subject correlation. The approach has been recommended by experts in the count data field [264]. Therefore, despite the zero inflation, the exclusions data could be validly analysed using a negative binomial model.

4.4.1.4 Special Educational Need

Pupils were classed as having a SEN in a given school year if any SEN had been recorded on their school census record for that year. SEN was therefore analysed as a yearly dichotomous binomial outcome (yes/no). Generalised estimating equations with an independent correlation structure, binomial distribution and logit link function were employed using the *xtgee* command in Stata to investigate each chronic condition against the outcome of SEN. The same analyses were repeated for different types of SEN: learning disability, learning difficulty, sensory impairment, physical motor impairment, communication problems, autism, physical and mental health problems and social, emotional or behavioural difficulties. SEN categories were not mutually exclusive therefore children could exist across multiple categories and be diagnosed with more than one SEN in the same census year. Pupils with each respective type of SEN were, therefore, identified and compared against pupils without any type of SEN. This enabled the impact of each type of SEN to be investigated against a common comparison group.

4.4.2 Overall educational outcomes (attainment and unemployment)

Pupil's academic attainment and unemployment were analysed as one-off overall final outcomes across the full time period 2009-2013. The final leaver and attainment datasets contained one summary record per pupil; therefore repeated measures were not an issue

and standard data analyses were employed rather than longitudinal generalised estimating equations. Similarly, whereas the longitudinal datasets described in sections 4.4.1.3 and 4.4.1.4, defined pupils' chronic disease status on a yearly basis, pupils within the leaver and attainment datasets were defined as having a chronic disease if they were prescribed relevant drugs (section 4.3.2.2) at any point between 2009 and 2013.

Chi squared tests for trends were performed to investigate the association between sociodemographic and maternity characteristics and unemployment. Spearman correlations and chi square tests for trends were performed to investigate the association between sociodemographic and maternity characteristics and attainment. Univariate logistic regression models were used to investigate the impact of each chronic condition (yes/no) on each of the one off outcome measures (attainment and unemployment). Unemployment was recorded as a dichotomous outcome: 0= in education, employment or training (EET); 1= not in education, employment or training (NEET). Therefore binomial logistic regression was used to assess the impact of each chronic condition on unemployment. Attainment was analysed as an ordinal four-category measure: 1=high attainment; 2=general/broad attainment; 3=basic attainment; 4=low attainment. Therefore, generalised ordinal logistic regression was used to assess the impact of each chronic condition on the level of final grade attainment [265]. Ordinal logistic regression performs a series of binary logistic regressions at each ordered category, each time calculating the odds of higher categories occurring compared to that category combined with all lower categories for a given predictor variable or set of predictor variables. In the case of a four-category ordered outcome variable, the model returns the respective odds of: categories 2 or above occurring compared to category 1; categories 3 or above occurring compared to categories 2 or below; category 4 occurring compared to categories 3 or below. When the proportional odds assumption is met, these odds are all the same; however, when they differ the assumption is violated. The generalised ordinal logistic regression model therefore allows partial proportional odds to be fitted whereby proportional odds are met for some variables but not for others. I ran partial proportional odds models for each chronic condition using the *gologit2* command in Stata and utilised the *autofit* option in order to formally test the parallel odds assumption for all variables in the models. I defined higher numbered categories to represent poorer attainment; therefore, a higher odds ratio within each attainment category represented worse attainment outcomes.

The respective models were adjusted for sociodemographic confounders: age, sex, SIMD 2012 quintile and ethnicity. Age categories (<11, 11-14 and >14 years) were included in

longitudinal models investigating absence, exclusion and risk of SEN because pupil ages ranged between 4 and 18 years. However, school leavers and children of attainment age only ranged between 14 and 18 years. Therefore these outcomes were adjusted using alternative measures of pupil age. Models were investigated including age in years as a continuous or discrete measure and the potential non-linearity of age was investigated using a quadratic age squared term. All of these models were assessed using the Akaike information criterion (AIC) and Bayesian information criterion (BIC) statistics in order to ascertain which model provided the best fit [266]. Unemployment was most appropriately modelled using age as a discrete variable. Attainment was most adequately modelled using age as a continuous variable. I further adjusted the models for maternal confounders: maternal age, sex and gestation specific birth weight centile, estimated gestational age, mode of delivery, Apgar score, maternal smoking during pregnancy and parity.

Univariate interactions with sex and SIMD 2012 deprivation quintile were investigated for each chronic disease status against each outcome. Where interactions existed, univariate and multivariate subgroup analyses by sex and SIMD 2012 deprivation quintile were performed. I did not investigate interactions with age because, whilst the full population of schoolchildren had ages ranging between 4 and 19 years, school leavers and children of attainment age only ranged between 14 and 18 years.

Sensitivity analyses were performed whereby pupils with a recorded SEN were omitted from both the leaver destination and attainment cohorts and the various analyses were repeated. This enabled me to investigate whether any associations with poorer attainment and unemployment were due to the chronically ill children who also had SEN or whether associations also existed in chronically ill children even if they did not have a SEN. Furthermore, attainment and unemployment were investigated again after adjusting for number of days absent. This investigated whether absenteeism was a possible mediator in the association between each chronic condition and attainment and unemployment.

Children may experience more than one condition; therefore, I reran the final multivariate models including all conditions together. Each chronic condition specific analysis was therefore finally adjusted for presence of any of the other identified chronic conditions in addition to sociodemographic and maternal confounders. Where appropriate, output was tabulated and/or presented graphically in the form of forest plots.

4.4.3 Health outcomes

Children were followed up for a maximum of 5 years from 21st Sept 2009 until 21st Sept 2014 for the following outcomes:

1. First all-cause acute or psychiatric hospital admission
2. First cause-specific acute or psychiatric hospital admission
3. Total number of all-cause acute and psychiatric hospital admissions
4. Total length of stay in hospital
5. All-cause mortality

The earliest possible entry date to the study for any child was the date of the 2009 school census (21st Sep 2009). However, the specific entry date and the age at entry were different for each pupil. Children without one of the five defined chronic conditions and children with ADHD and depression entered the study on the pupil census date of their first known school census year. Children with diabetes, asthma or epilepsy entered on the date of their first known prescription if this was after their first known census date. The maximum follow up period was five years until 21st Sept 2014 regardless of whether a pupil had left school within this period. Some children may have left school between 2009 and 2013 with no prescriptions for any of the chronic conditions but subsequently received medications in the post school period. This study only investigated outcomes for pupils who received medication whilst at school therefore children who left school with no record of treatment for one of the chronic conditions were defined as not suffering from the condition. The follow up period varied for each pupil and ranged from one year to five years.

4.4.3.1 Risk of admission and death over follow up

Univariate and multivariate Cox proportional hazards models were initially used to investigate the impact of each chronic illness on risk of all-cause hospital admission, cause-specific hospital admission and all-cause mortality. Survival models typically use time from study entry as the time scale and adjust for age at study entry. However, age can be used as an alternative time scale [267]. The latter is less commonly used; however, it is

increasing in popularity and the pros and cons of both methods have been discussed elsewhere with respect to investigating subjects entering a follow up study at different ages and different time points [268,269].

Cox proportional hazards models require proportional hazards to hold if the point estimates and standard errors are to be robust. I ran univariate and multivariate Cox proportional hazards models for each chronic condition using time from study entry adjusted for age and then again using age and tested the proportional hazards assumption each time using the *estat phtest* command within Stata. For models investigating risk of admission, I initially defined time from study entry as the period elapsed between each pupils first school census date (for pupils who didn't receive any prescriptions of interest) or first prescription date (for pupils who did receive a prescription of interest) during the study period, and the date of their first subsequent hospital admission within 5 years of follow up or the censor date of the study. Models investigating risk of death used the same methods; however, date of death was used instead of first subsequent date of admission if a death occurred over the follow up period. I adjusted for pupil age at entry each time to take into account pupils entering the study at different ages. I then reran the models using pupil age at admission, death or censor date as the time scale. Models which satisfied the proportional hazards assumption were accurately described in this way. They described a constant hazard ratio over the follow up period and gave similar results regardless of whether time was modelled as time on study or age at event. Therefore, I presented output from proportional hazards models using time on study adjusted for age as this is the more common method.

Alternative models are required to describe hazard ratios which do not satisfy the proportional hazards assumption because the hazard ratio varies over different periods of time or at different ages. Alternative models include Poisson piecewise regression models with time dependent effects [267] and extended Cox regression models with heaviside step functions for time dependent variables [270]. Both of these models manually split the follow up time into various small subsections and assume the equivalent of proportional hazards within each section rather than over the full follow up period. Whilst the Cox models define excess risk in terms of hazard ratios, the Poisson models define it as incidence rate ratios. However these can be interpreted in the same way [267]. Both models yield similar results; however, I chose to use Poisson piecewise regression models in my analyses because these models also provide baseline incidence rate ratios within each category. Cox regression models with time varying covariates (tvc) analysed continuously over follow up

rather than within defined subsections were also investigated using stcox with the tvc option. However, they were not feasible to use in this study due to the large population size which resulted in an extremely long computation time for each model.

Chronic conditions which violated the proportional hazards assumption on Cox regression modelling were therefore re-run using Poisson piecewise regression models with time dependent effects. Similar to the Cox regression models, these were run using firstly time from study entry adjusted for age and then using actual age at event. These measures were defined in exactly the same way as for the Cox regression models described above. However, this time I chose to split the time from study entry (first school census date or first relevant prescription date in the study period) until admission or death into categories: >= 0 and < 1; >= 1 and < 2; >= 2 and < 3; >= 3 and < 4; >= 4 and <= 5 years. For models using pupil age (at admission, death or censor date) as the time scale, I split age at event into categories: 4-6, 7-8, 9-10, 11-12, 13-14, 15-16 and 17-23 years. These models allowed the hazard ratios to be described within different portions of the follow up period and again for different age categories at the event of interest.

All of the Cox proportional hazards and Poisson piecewise models were adjusted for sociodemographic and maternity confounders: sex, SIMD 2012 quintile, ethnicity, maternal age, sex and gestation specific birth weight centile, estimated gestational age, mode of delivery, Apgar score, maternal smoking during pregnancy and parity. The survival models using time from study entry as the time scale were additionally adjusted for age category (<11, 11-14 and >14 years) at time of study entry.

Univariate interactions with sex were investigated for each chronic disease status against each outcome. Where interactions existed, subgroup analyses by gender were performed to investigate time to first all-cause and cause-specific admission. These were not performed for all-cause mortality due to an insufficient number of deaths. Several survival plots were produced for each model. Where appropriate, output was tabulated and presented graphically in the form of forest plots to show hazard ratios and incidence rate ratios across follow up time and for different ages at event. Children may experience more than one disease; therefore, I reran the final multivariate models including all diseases together. Each chronic condition specific analysis was therefore finally adjusted for presence of any of the other identified chronic conditions in addition to sociodemographic and maternal confounders.

4.4.3.2 Total admissions and total length of hospital stay

The impact of each chronic condition on the outcomes of total number of admissions and total length of hospital stay over the follow up period were analysed via recommended count models [264]. The specific number of days followed up was used as the offset exposure variable for each child since children with longer available follow up had increased chance of more admissions. Around 80% of pupils had no admissions and zero length of stay in hospital over the follow up period and both distributions were skewed with large amounts of dispersion. Whilst Poisson regression count models assume that the mean and variance are equal, negative binomial regression models allow for greater dispersion. Two variants of the negative binomial model are used depending on how the dispersion varies according to the mean. A quadratic negative binomial model (NB2) is more commonly used however there are times when a linear negative binomial model (NB1) fits the data better. In both cases, the variance is constant for all observations. However a third variant, the NB-P model, provides even further flexibility by allowing the dispersion to vary for different observations. I additionally considered models which would specifically take into account the zero inflation aspect of the data. All three variants of the negative binomial model (NB1, NB2 and NB-P), a zero inflated negative binomial model (ZINB) and a zero inflated Poisson model (ZIP) were therefore investigated and compared using AIC and BIC [266] statistics. The most suitable model was a negative binomial with varying dispersion (NB-P) [264].

5 Results

5.1 Linkage, data cleaning and producing analyses files

5.1.1 Education data

Education records were provided by ScotXed as seven separate datasets. The number of records and the number of pupils within each of the originally supplied datasets is shown in Table 19. The SEN data were supplied as two separate files with one containing data for 2009 and the other containing data from 2010 onwards. These were collated after cleaning each file; therefore the SEN data have been summarised as one dataset. Table 19 also shows the number of records and the number of pupils after de-duplicating and cleaning each file. The cleaning process is explained in more detail in the following sections.

Table 19 Education records pre and post data cleaning

		supplied data		cleaned data	
Dataset	**Census Years**	**Records**	**Pupils**	**Records**	**Pupils**
Pupil Census	2009 - 2013	3,368,837	961,382	2,793,185	766,244
Attendance	2009/10/12	1,919,318	856,565	1,597,433	702,225
Exclusions	2009/10/12	71,949	32,222	60,422	26,917
SEN	2009 - 2013	839,806	189,378	659,291	141,640
Attainment	2009 - 2013	11,226,244	532,698	9,318,995	297,471
Leaver	2009 - 2013	263,736	259,597	219,958	219,366

5.1.1.1 Cleaning the pupil census data

Using date of birth, sex and postcode as linkage identifiers in the probability match, eDRIS managed to link 98.4% of all schoolchildren on the pupil census between 2009 and 2013 to the CHI database. These links were deemed to have a precision of 98% and sensitivity of 99.5%. There were 3,368,837 records on the full census file pertaining to 961,382 pupils across five census years and 3,315,699 records (98.4%) successfully linked to the CHI database and were retained (908,244 pupils). The cohort was further restricted to include only 2,909,149 records which had an associated linked SMR02 maternity record. I deleted 406,539 records which did not have a corresponding SMR02 record and then deleted 80,680 records pertaining to non-singleton births. After limiting the cohort to exclusively pupil records which had linked to the CHI database and records which had an associated SMR02, I found that 15,918 records within the reduced cohort of 2,828,480 records had a mismatch between the pupil census date of birth and the date of birth recorded on SMR02.

Furthermore 1,236 of these pupils had more than one date of birth recorded across their groups of school records. The school census date of birth was replaced with the corresponding SMR02 date of birth for 15,907 records where the two dates did not match but where the difference was less than one year. This was done because the SMR02 date of birth was deemed to be more accurate than the pupil census date of birth. The remaining eleven records where the dates were more than a year apart were deemed to be poor links and these were omitted from the cohort. No discrepancies in recorded date of birth remained after this step and each pupil had only one recorded date of birth across all of their records.

Following these steps, 15,392 records were deleted because they had a partition code equal to E or H and a further 2,047 were deleted because they were duplicated by partition where the partition was not equal to E or H. Several records also had a duplicated record where the partitions were different. In those cases, the record with the worst partition was excluded from the analysis and this resulted in 17,213 records being deleted (section 4.2). No further duplicates remained after these exclusions. The dataset went through two final cleaning stages. I deleted 64 records where the pupil age was recorded as greater than 19 years or less than 4 years. Finally, 568 records were deleted where stage of school was recorded as AD (adult learner), S7 or S9 in order to restrict the cohort to children aged between 4 and 19 years old and in P1 through to S6 or in special schools (SP). The final cleaned pupil census data contained 2,793,185 records pertaining to 766,244 pupils across five census years 2009-2013.

The pupil census file was aggregated by study index number and census year to contain one record per pupil per census year. There were 2,793,185 records in this lookup file which was matched against all of the other education datasets to ensure that pupils excluded from the census file were also excluded from these datasets. Figure 4 is a flow diagram illustrating each step of data cleaning performed on the school census data and shows how many records remained at each cleaning stage. Table 20 shows the number of pupil records per census year in the census file after deleting erroneous links and duplicate records.

Figure 4 Data cleaning performed on ScotXed school census data

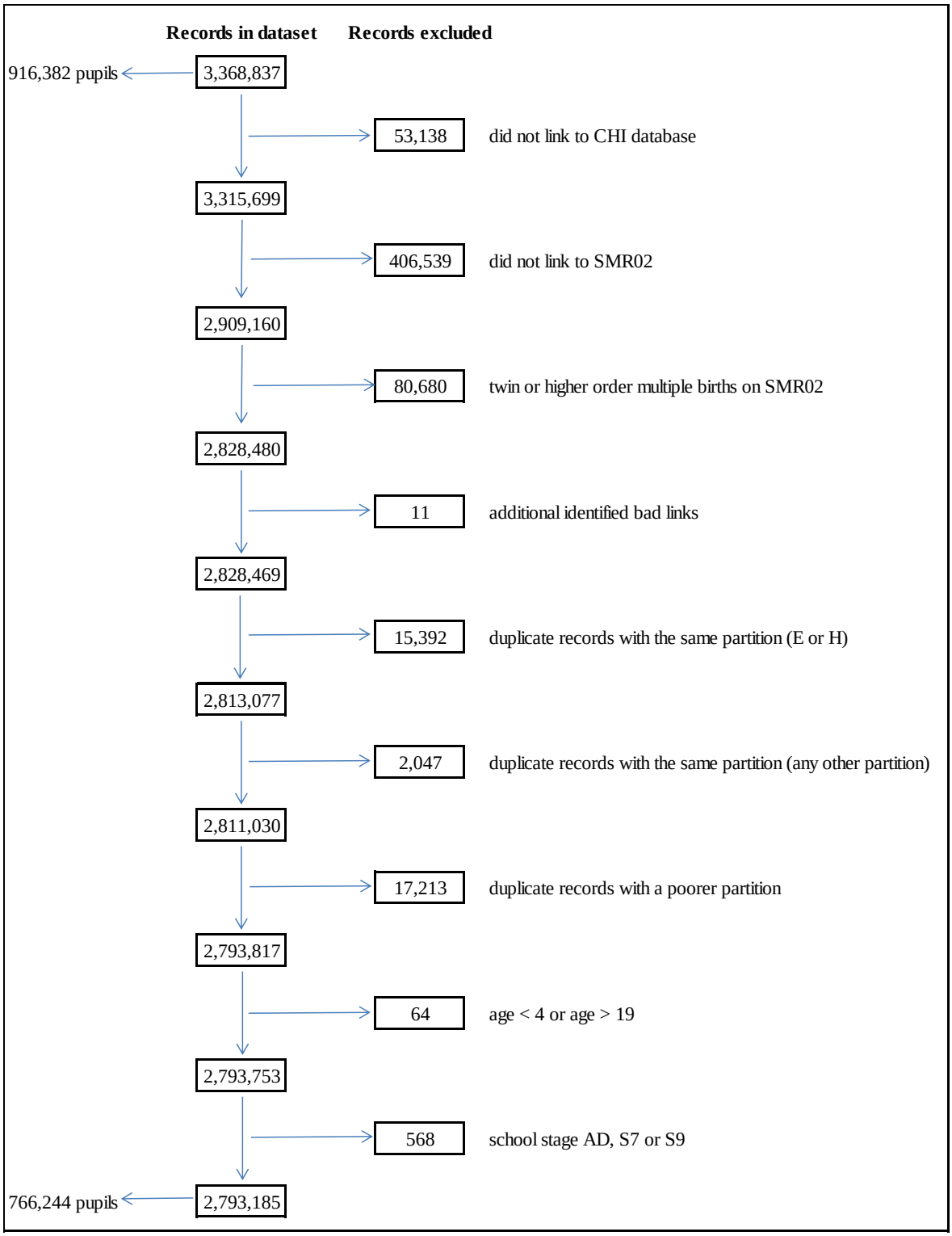


Table 20 Cleaned pupil census data – number of pupils per school year

Year	Number of records
2009	565,269
2010	560,488
2011	556,315
2012	555,599
2013	555,514
Total	2,793,185

5.1.1.2 Cleaning the linked education datasets

The remaining education datasets also contained duplicate records; however, the majority of these disappeared after matching the pupil census lookup file to each dataset by index number, census year and record ID and then only selecting records pertaining to those pupils on the final cleaned version of the pupil census. Table 19 in section 5.1.1 presents the number of records and the number of pupils in each of the education datasets after initially cleaning them by matching them against the pupil census lookup file. A residual number of duplicates remained and these were removed. This process is explained in the following sections.

5.1.1.3 Absence data

The absence dataset contained 1,919,318 records which reduced to 1,597,433 records after matching to the pupil census lookup. This matching process removed all of the duplicate records within the absence dataset. The data contained one record per pupil per census year and so this was already in a suitable format to be matched back onto the pupil census data without further manipulation.

The absence data were matched by index number and census year to the 2,793,185 pupil records within the full pupil census 2009-2013. Absence data were only available for 2009, 2010 and 2012; therefore 1,681,356 pupil records remained after selecting census records in those years. 83,923 records were deleted because they had no absence data and a further 36 records were deleted because they had a value of zero days possibly attended. The final absence dataset contained 1,597,397 records pertaining to 702,210 pupils across census years 2009, 2010 and 2012.

5.1.1.4 Exclusions data

The exclusion dataset contained 71,949 records which reduced to 60,422 records after matching to the pupil census lookup. The dataset contained a record for every exclusion episode; therefore the data had to be manipulated before being matched on to the full pupil census file. A summarised exclusions dataset containing 33,254 pupil records was produced. This contained a maximum of one record per pupil per census year. The exclusion data were matched by index number and census year to the 2,793,185 pupil records within the full pupil census 2009-2013. Exclusion data were only available for 2009, 2010 and 2012 therefore 1,681,356 pupil records remained after selecting records in those years. 83,923 records were deleted because they had no attendance data and a further 36 records were deleted because they had a value of zero days possibly attended. The final exclusions dataset contained 1,597,397 records pertaining to 702,210 pupils across census years 2009, 2010 and 2012.

5.1.1.5 Special educational need data

The Reason for Support dataset contained 68,588 SEN records in 2009 relating to 51,799 pupils. This reduced to 53,765 records after matching against the pupil census lookup file. The dataset contained no further duplicate records. Several recorded SEN codes did not fulfil my definition of SEN and were therefore excluded. 47,003 records of interest remained. The file was aggregated by index number and census year to provide 36,020 records each containing a yearly summary of SEN records for each pupil. The maximum number of different SEN recorded in a year for a single pupil was nine.

Similarly the Student Need dataset contained 771,218 SEN records between 2010 and 2013 relating to 181,013 pupils. This reduced to 605,526 records after matching to the pupil census lookup. Several remaining duplicate records were deleted to leave 601,131 records. 501,955 records remained after retaining only eligible SEN codes. The file was aggregated by index number and census year to provide 267,219 records each containing a yearly summary of SEN recorded for each pupil. The maximum number of different SEN recorded in a year for a single pupil was eleven. The two datasets were then combined to provide 303,239 records. The SEN data were matched by index number and census year to the 2,793,185 pupil records within the full pupil census 2009-2013.

5.1.1.6 Leaver status data

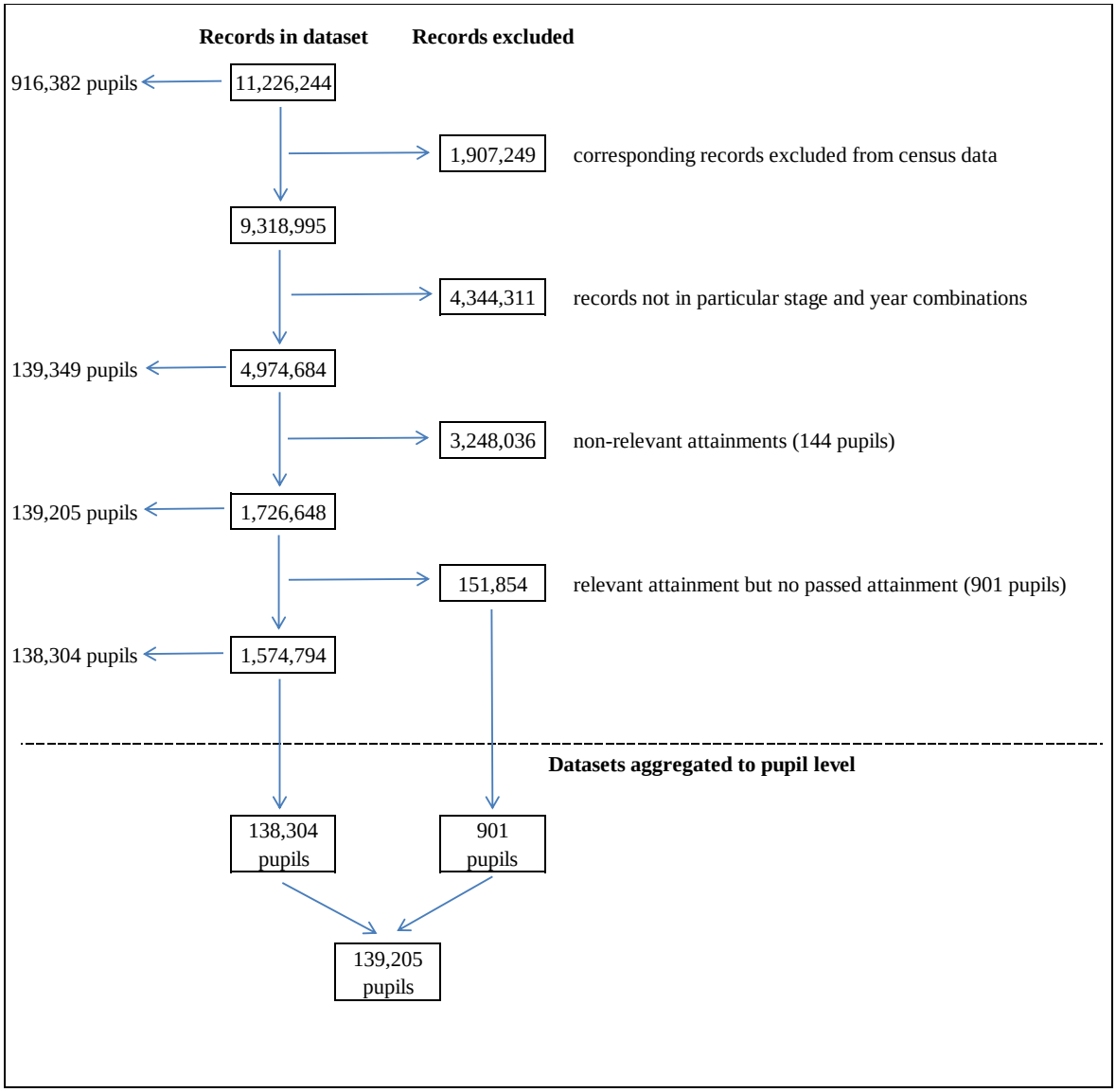
The leaver dataset contained 263,736 records between 2009 and 2013. 219,958 records remained after matching the leaver data to the pupil census lookup file in order to omit pupils that had been deleted from the census dataset. 1,184 duplicates (592 pairs) remained in the leaver dataset. With the exception of one pair, these were not exact duplicates but instead differed either on leaving date, leaver status or both. As there was no way of telling which record was the correct one both records were deleted in these 591 duplicated pairs. One of the records was retained in the remaining exact duplicate pair. This resulted in a further 1,183 records being deleted. Finally 851 records were deleted where leaver status was recorded as either excluded, moved outwith Scotland or unavailable due to ill health. The remaining 217,924 records were matched by index number to the 2,793,185 pupil records within the full pupil census. The final census record, with appended leaver data, was then selected for each pupil to enable demographic variables at time of leaving to be retained.

5.1.1.7 Attainment data

Figure 5 shows the number of records within the attainment file at each stage of data cleaning as previously explained in section 4.3.1.2.5. There were 11,226,244 records within the original attainment dataset and this reduced to 9,318,995 records after matching against the pupil census lookup file in order to omit pupils that had been deleted from the census dataset. The attainment data were further reduced to include only records pertaining to pupils in stages S4, S5 and S6 who attended school in particular years (Table 12 in section 4.3.1.2.5). The dataset was reduced to 4,974,684 records (139,349 pupils) after omitting pupils in S5 and S6 in 2009, S6 in 2010, S4 in 2012, and S4 and S5 in 2013. Only attainments of interest were retained (namely numeric, C and specific G codes) which further reduced the dataset to 1,726,648 records (139,205 pupils). 1,574,794 records pertaining to 138,304 pupils related to attainment at pass level or above and these were summed in order to derive an attainment level (1=high; 2=general/broad; 3=basic; 4=low) for each pupil. The remaining 151,854 attainment records (901 pupils) related to failed attainment. These pupils were classed as having low attainment. Summarised attainment relating to 139,205 pupils was therefore included in the final attainment cohort. These records were matched by index number to the 2,793,185 pupil records within the full pupil census. The last available census record, with appended attainment data, was then selected for each pupil to enable relevant demographic variables to be retained. A very small percentage of pupils (2.3%) of attainment age had no linked attainment data and this was

most likely due to linkage limitations or those pupils leaving school midway through term whereby they appeared on the September school census but left school before sitting any end of year exams (section 4.3.1.2.5)

Figure 5 Data cleaning performed on ScotXed attainment data



5.1.2 Prescribing data

The prescribing dataset provided by eDRIS contained 4,511,428 prescription records pertaining to 292,490 pupils. This reduced to 3,094,548 records after selecting only prescriptions pertaining to those pupils on the cleaned school census data. The dataset contained all prescriptions between 2009 and 2013 for each child who had attended school at any point between 2009 and 2013. Therefore several prescriptions related to drugs taken either before starting or after leaving school. Since only drugs taken whilst at school were of interest, prescription records were only retained in the dataset if they were prescribed during any calendar year in which the pupil actually attended school. This reduced the dataset to contain 1,936,853 records. Table 21 shows the total number of prescriptions specific to each chronic condition present on the PIS dataset. The first column shows the total number of prescriptions between 2009 and 2013 within BNF chapter 6 (diabetes), BNF chapter 3 (asthma), BNF section 4.8 (epilepsy), BNF section 4.4 (ADHD) and BNF section 4.3 (depression). The second column shows the number of prescriptions cashed whilst at school i.e. medications prescribed before children had started or after they had left school during the period are not counted. Finally the third column shows the number of prescriptions whilst at school for each chronic condition after including only specific BNF codes of interest previously outlined in methods section 4.3.2.2.

Table 21 Total number of prescriptions specific to each chronic condition

Chrionic condition	Total prescriptions*	Total prescriptions at school**	Total eligible prescriptions at school***
Diabetes	234,434	142,782	140,973
Asthma	2,111,481	1,391,092	1,275,121
Epilepsy	252,205	144,609	144,609
ADHD	319,607	229,342	229,342
Depression	176,821	29,028	28,986

*Prescription within original BNF section (see methods section 4.3.2.1) prescribed between 2009 and 2013

**Prescription within original BNF section (see methods section 4.3.2.1) prescribed whilst at school between 2009 and 2013

***Perscription within tighter definition (see methods section 4.3.2.2) prescribed whilst at school between 2009 and 2013

The prescriptions data were aggregated by index number and census year and summarised as a one record per pupil per census year dataset before then being matched by index number and census year to the 2,793,185 pupil records within the full pupil census across years 2009 to 2013. Diabetes medication was prescribed during the year for 10,604 of the pupil census records. 716 pupil census records linked to ineligible diabetes prescriptions and were excluded from the diabetes analyses (section 4.3.2.2.1). Similarly, asthma medication was prescribed for 148,727 census records and 320,387 pupil census records linked to ineligible asthma prescriptions and were excluded from asthma specific analyses

(section 4.3.2.2.2). Epilepsy medication and ADHD medication was prescribed during the year for 15,295 and 30,168 pupil census records respectively (sections 4.3.2.2.3 and 4.3.2.2.4). Finally, anti-depressant medication was prescribed during the year for 20,328 of the pupil census records and 28 census records linked to ineligible anti-depressant prescriptions and were excluded from depression specific analyses (section 4.3.2.2.5). The full school census dataset held records across 2009-2013 pertaining to 766,244 pupils. 3,330 pupils received at least one relevant diabetes prescription during at least one of their school years and were therefore classed as diabetic whilst 197 pupils were excluded from the diabetes analyses because they were ineligible. Therefore diabetes prevalence was 0.43%. The same criteria were used to diagnose children as having epilepsy, ADHD and depression. 5,314 children were classed as having epilepsy (0.69%) and 7,413 were classed as having ADHD (0.97%). 5,342 were classed as depressed whilst 7 pupils were excluded from the depression analyses. Therefore depression prevalence was 0.70%. 45,900 pupils received more than one relevant asthma prescription during at least one of their school years and were therefore classed as asthmatic (section 4.3.2.2.2) whilst 82,528 pupils were excluded from asthma specific analyses. Therefore asthma prevalence was 5.99%.

5.1.3 Maternity, hospital admission and death data

953,410 linked SMR02 maternity records were originally provided by eDRIS. However, after matching this to the pupil census lookup file, and omitting several records as described in section 5.1.1.1, 766,244 records remained. Only children who had a linked SMR02 maternity recorded were included in the final analysis cohort. Similarly 655,786 hospital admission and 1,213 death records were provided by eDRIS, of which 464,600 and 561 remained respectively after matching to the pupil census lookup file.

5.2 Characteristics of children by chronic disease status

Children with diabetes were not significantly different to those without in terms of deprivation, gender, mode of delivery, 5 minute Apgar score, gestational age, maternal age and maternal parity. In contrast, children with diabetes were significantly older, had significantly higher sex-gestation-specific birthweight centiles and their mothers were significantly more likely to have smoked during pregnancy. They were also significantly more likely to be of white descent and less likely to be of black or Asian descent (Table 22).

Significantly more boys than girls had asthma. Children with asthma were significantly older, more deprived and had significantly lower 5 minute Apgar scores, sex-gestation-specific birthweight centiles and gestational ages. Their mothers were significantly younger and significantly more likely to have smoked during pregnancy and been nulliparous. Children with asthma were significantly more likely to be of Asian descent and significantly less likely to be white. Finally, they were significantly more likely to have been delivered by caesarean section or assisted vaginal delivery and significantly less likely to have been delivered by spontaneous vaginal delivery (Table 23).

Children with epilepsy were not significantly different to those without in terms of gender, ethnicity or maternal parity. However, children with epilepsy were significantly older, more deprived, and had significantly lower Apgar scores, gestational ages and sex-gestation-specific birthweight centiles. Their mothers were significantly younger and significantly more likely to have smoked during pregnancy. Finally, children with epilepsy were significantly more likely to have been delivered by caesarean section and significantly less likely to have been delivered by spontaneous vaginal delivery (Table 24).

Children with ADHD were not significantly different to those without in terms of mode of delivery. Significantly more boys than girls had ADHD. Children with ADHD were significantly more deprived and had significantly lower 5-minute Apgar scores, gestational ages and sex-gestation-specific birthweight centiles. They were more likely to have been born to mothers who were younger than 25 years of age, multiparous, and who smoked during pregnancy. Children with ADHD were significantly more likely to be aged between 11 and 14 years. This was explained by the fact that children with ADHD left school earlier. 64.3% of children with ADHD left school before 16 years of age compared with 28.4% of children without ADHD (Table 25).

Children with depression were not significantly different to those without in terms of 5-minute Apgar score or maternal parity. Significantly more girls than boys had depression and depressed children were significantly more likely to be white and significantly less likely to be of Asian or black descent. Children with depression were significantly older, had significantly lower sex-gestation specific birthweight centiles and their mothers were significantly more likely to have smoked during pregnancy and given birth between 25 and 29 years of age. Finally depressed children were significantly less likely to reside in the most deprived areas, had significantly lower gestational age and were significantly less likely to have been born by emergency caesarean section (Table 26).

Table 22 Characteristics of children with and without diabetes

		No Diabetes N=2,781,865		Diabetes N=10,604		Total N=2,792,469		
		n	%	n	%	n	%	P value
Sociodemographic factors								
Sex								
	male	1,416,247	50.9	5,343	50.4	1,421,590	50.9	0.586
	female	1,365,618	49.1	5,261	49.6	1,370,879	49.1	
	Missing	0		0		0		
Age category								
	<11	1,395,183	50.2	3,099	29.2	1,398,282	50.1	<0.001
	11-14	671,028	24.1	3,258	30.7	674,286	24.1	
	>14	715,654	25.7	4,247	40.1	719,901	25.8	
	Missing	0		0		0		
Deprivation quintile								
	1 (most deprived)	623,841	22.5	2,131	20.1	625,972	22.4	0.131
	2	556,100	20.0	2,270	21.4	558,370	20.0	
	3	538,476	19.4	2,098	19.8	540,574	19.4	
	4	545,375	19.6	2,117	20.0	547,492	19.6	
	5 (least deprived)	514,232	18.5	1,980	18.7	516,212	18.5	
	Missing	3,841		8		3,849		
Ethnic group								
	White	2,633,184	96.4	10,322	98.2	2,643,506	96.4	0.002
	Asian	61,453	2.2	85	0.8	61,538	2.2	
	Black	5,372	0.2	1	0.0	5,373	0.2	
	Mixed	23,861	0.9	82	0.8	23,943	0.9	
	Other	7,760	0.3	19	0.2	7,779	0.3	
	Missing	50,235		95		50,330		
Maternity factors								
Maternal age (years)								
	<=24	748,816	26.9	2,612	24.6	751,428	26.9	0.525
	25-29	800,888	28.8	3,334	31.4	804,222	28.8	
	30-34	809,310	29.1	3,120	29.4	812,430	29.1	
	>=35	422,798	15.2	1,538	14.5	424,336	15.2	
	Missing	53		0		53		
Maternal smoking								
	No	1,809,176	72.2	7,305	77.4	1,816,481	72.2	<0.001
	Yes	697,096	27.8	2,133	22.6	699,229	27.8	
	Missing	275,593		1,166		276,759		
Parity								
	0	1,250,176	45.2	4,897	46.4	1,255,073	45.2	0.346
	1	961,456	34.8	3,567	33.8	965,023	34.8	
	>1	554,099	20.0	2,092	19.8	556,191	20.0	
	Missing	16,134		48		16,182		

		No Diabetes N=2,781,865		Diabetes N=10,604		Total N=2,792,469		
		n	%	n	%	n	%	P value
Maternity factors								
Mode of delivery								
	SVD	1,802,207	64.8	6,865	64.7	1,809,072	64.8	0.643
	Cephalic	69,105	2.5	267	2.5	69,372	2.5	
	Assisted VD	332,531	12.0	1,333	12.6	333,864	12.0	
	Breech Delivery	7,941	0.3	26	0.2	7,967	0.3	
	Elective Caesarean	210,248	7.6	802	7.6	211,050	7.6	
	Emergency Caesarean	359,314	12.9	1,306	12.3	360,620	12.9	
	Other Delivery	514	0.0	5	0.0	519	0.0	
	Missing	5		0		5		
Gestation (weeks)								
	<24	100	0.0	0	0.0	100	0.0	0.076
	24-27	4,120	0.1	10	0.1	4,130	0.1	
	28-32	25,909	0.9	99	0.9	26,008	0.9	
	33-36	129,644	4.7	507	4.8	130,151	4.7	
	37	136,915	4.9	607	5.7	137,522	4.9	
	38	351,340	12.6	1,482	14.0	352,822	12.6	
	39	574,531	20.7	2,061	19.4	576,592	20.7	
	40	832,831	30.0	3,352	31.6	836,183	30.0	
	41	627,548	22.6	2,166	20.4	629,714	22.6	
	42	95,090	3.4	298	2.8	95,388	3.4	
	43	1,996	0.1	9	0.1	2,005	0.1	
	>43	469	0.0	9	0.1	478	0.0	
	Missing	1,372		4		1,376		
Sex-gestation-specific birthweight centile								
	1-3	114,889	4.1	342	3.2	115,231	4.1	<0.001
	4-10	249,978	9.0	820	7.7	250,798	9.0	
	11-20	331,924	11.9	1,095	10.3	333,019	11.9	
	21-80	1,633,154	58.8	6,296	59.4	1,639,450	58.8	
	81-90	237,005	8.5	1,018	9.6	238,023	8.5	
	91-97	150,138	5.4	715	6.8	150,853	5.4	
	98-100	62,104	2.2	305	2.9	62,409	2.2	
	Missing	2,673		13		2,686		
5 min Apgar								
	1-3	12,557	0.5	62	0.6	12,619	0.5	0.776
	4-6	25,997	0.9	81	0.8	26,078	0.9	
	7-10	2,713,873	98.6	10,386	98.6	2,724,259	98.6	
	Missing	29,438		75		29,513		

Table 23 Characteristics of children with and without asthma

		No Asthma N=2,324,071		Asthma N=148,727		Total N=2,472,798		
		n	%	n	%	n	%	P value
Sociodemographic factors								
Sex								
	male	1,158,074	49.8	88,141	59.3	1,246,215	50.4	<0.001
	female	1,165,997	50.2	60,586	40.7	1,226,583	49.6	
	Missing	0		0		0		
Age category								
	<11	1,165,279	50.1	65,867	44.3	1,231,146	49.8	<0.001
	11-14	553,813	23.8	41,629	28.0	595,442	24.1	
	>14	604,979	26.0	41,231	27.7	646,210	26.1	
	Missing	0		0		0		
Deprivation quintile								
	1 (most deprived)	510,710	22.0	39,964	26.9	550,674	22.3	<0.001
	2	458,412	19.8	32,407	21.8	490,819	19.9	
	3	449,201	19.4	28,737	19.4	477,938	19.4	
	4	462,014	19.9	25,823	17.4	487,837	19.8	
	5 (least deprived)	440,511	19.0	21,579	14.5	462,090	18.7	
	Missing	3,223		217		3,440		
Ethnic group								
	White	2,202,662	96.5	139,780	95.4	2,342,442	96.5	<0.001
	Asian	49,178	2.2	4,445	3.0	53,623	2.2	
	Black	4,393	0.2	344	0.2	4,737	0.2	
	Mixed	19,337	0.8	1,438	1.0	20,775	0.9	
	Other	6,423	0.3	453	0.3	6,876	0.3	
	Missing	42,078		2,267		44,345		
Maternity factors								
Maternal age (years)								
	<=24	613,910	26.4	44,390	29.8	658,300	26.6	<0.001
	25-29	669,640	28.8	43,586	29.3	713,226	28.8	
	30-34	683,812	29.4	40,252	27.1	724,064	29.3	
	>=35	356,661	15.3	20,499	13.8	377,160	15.3	
	Missing	48		0		48		
Maternal smoking								
	No	1,521,803	72.7	92,704	69.2	1,614,507	72.5	<0.001
	Yes	571,578	27.3	41,177	30.8	612,755	27.5	
	Missing	230,690		14,846		245,536		
Parity								
	0	1,032,161	44.7	70,468	47.7	1,102,629	44.8	<0.001
	1	808,818	35.0	49,203	33.3	858,021	34.9	
	>1	470,004	20.3	28,182	19.1	498,186	20.3	
	Missing	13,088		874		13,962		

	No Asthma N=2,324,071		Asthma N=148,727		Total N=2,472,798		
	n	%	n	%	n	%	P value
Maternity factors							
Mode of delivery							
SVD	1,516,045	65.2	92,085	61.9	1,608,130	65.0	<0.001
Cephalic	57,367	2.5	4,050	2.7	61,417	2.5	
Assisted VD	275,435	11.9	18,246	12.3	293,681	11.9	
Breech Delivery	6,535	0.3	489	0.3	7,024	0.3	
Elective Caesarean	174,639	7.5	11,985	8.1	186,624	7.5	
Emergency Caesarean	293,604	12.6	21,840	14.7	315,444	12.8	
Other Delivery	444	0.0	32	0.0	476	0.0	
Missing	2		0		2		
Gestation (weeks)							
<24	51	0.0	10	0.0	61	0.0	<0.001
24-27	2,823	0.1	611	0.4	3,434	0.1	
28-32	19,731	0.8	2,420	1.6	22,151	0.9	
33-36	103,945	4.5	9,165	6.2	113,110	4.6	
37	111,828	4.8	8,525	5.7	120,353	4.9	
38	291,512	12.5	19,698	13.3	311,210	12.6	
39	480,906	20.7	30,269	20.4	511,175	20.7	
40	700,966	30.2	42,305	28.5	743,271	30.1	
41	528,531	22.8	30,976	20.8	559,507	22.6	
42	80,530	3.5	4,520	3.0	85,050	3.4	
43	1,699	0.1	111	0.1	1,810	0.1	
>43	407	0.0	37	0.0	444	0.0	
Missing	1,142		80		1,222		
Sex-gestation-specific birthweight centile							
1-3	94,515	4.1	6,977	4.7	101,492	4.1	<0.001
4-10	207,484	8.9	13,823	9.3	221,307	9.0	
11-20	275,916	11.9	18,033	12.1	293,949	11.9	
21-80	1,367,277	58.9	86,066	57.9	1,453,343	58.8	
81-90	198,902	8.6	12,184	8.2	211,086	8.5	
91-97	125,805	5.4	8,161	5.5	133,966	5.4	
98-100	52,000	2.2	3,286	2.2	55,286	2.2	
Missing	2,172		197		2,369		
5 min Apgar							
1-3	10,380	0.5	783	0.5	11,163	0.5	<0.001
4-6	20,951	0.9	1,800	1.2	22,751	0.9	
7-10	2,268,367	98.6	144,534	98.2	2,412,901	98.6	
Missing	24,373		1,610		25,983		

Table 24 Characteristics of children with and without epilepsy

		No Epilepsy N=2,777,890		Epilepsy N=15,295		Total N=2,793,185		
		n	%	n	%	n	%	P value
Sociodemographic factors								
Sex								
	male	1,413,986	50.9	7,811	51.1	1,421,797	50.9	0.830
	female	1,363,904	49.1	7,484	48.9	1,371,388	49.1	
	Missing	0		0		0		
Age category								
	<11	1,393,194	50.2	5,202	34.0	1,398,396	50.1	<0.001
	11-14	670,400	24.1	4,092	26.8	674,492	24.1	
	>14	714,296	25.7	6,001	39.2	720,297	25.8	
	Missing	0		0		0		
Deprivation quintile								
	1 (most deprived)	622,145	22.4	4,012	26.3	626,157	22.4	<0.001
	2	555,284	20.0	3,253	21.3	558,537	20.0	
	3	537,737	19.4	2,963	19.4	540,700	19.4	
	4	544,828	19.6	2,776	18.2	547,604	19.6	
	5 (least deprived)	514,080	18.5	2,253	14.8	516,333	18.5	
	Missing	3,816		38		3,854		
Ethnic group								
	White	2,629,665	96.4	14,475	96.1	2,644,140	96.4	0.948
	Asian	61,218	2.2	396	2.6	61,614	2.2	
	Black	5,350	0.2	23	0.2	5,373	0.2	
	Mixed	23,813	0.9	134	0.9	23,947	0.9	
	Other	7,748	0.3	31	0.2	7,779	0.3	
	Missing	50,096		236		50,332		
Maternity factors								
Maternal age (years)								
	<=24	746,994	26.9	4,610	30.1	751,604	26.9	<0.001
	25-29	799,744	28.8	4,677	30.6	804,421	28.8	
	30-34	808,718	29.1	3,930	25.7	812,648	29.1	
	>=35	422,381	15.2	2,078	13.6	424,459	15.2	
	Missing	53		0		53		
Maternal smoking								
	No	1,807,483	72.2	9,455	69.6	1,816,938	72.2	<0.001
	Yes	695,283	27.8	4,134	30.4	699,417	27.8	
	Missing	275,124		1,706		276,830		
Parity								
	0	1,248,370	45.2	7,032	46.2	1,255,402	45.2	0.253
	1	960,061	34.8	5,213	34.2	965,274	34.8	
	>1	553,338	20.0	2,989	19.6	556,327	20.0	
	Missing	16,121		61		16,182		

	No Epilepsy N=2,777,890		Epilepsy N=15,295		Total N=2,793,185		
	n	%	n	%	n	%	P value
Maternity factors							
Mode of delivery							
SVD	1,800,091	64.8	9,396	61.4	1,809,487	64.8	<0.001
Cephalic	68,996	2.5	406	2.7	69,402	2.5	
Assisted VD	332,154	12.0	1,767	11.6	333,921	12.0	
Breech Delivery	7,911	0.3	66	0.4	7,977	0.3	
Elective Caesarean	209,983	7.6	1,170	7.6	211,153	7.6	
Emergency Caesarean	358,237	12.9	2,484	16.2	360,721	12.9	
Other Delivery	513	0.0	6	0.0	519	0.0	
Missing	5		0		5		
Gestation (weeks)							
<24	96	0.0	4	0.0	100	0.0	<0.001
24-27	3,999	0.1	135	0.9	4,134	0.1	
28-32	25,659	0.9	349	2.3	26,008	0.9	
33-36	129,203	4.7	992	6.5	130,195	4.7	
37	136,648	4.9	949	6.2	137,597	4.9	
38	350,713	12.6	2,193	14.4	352,906	12.6	
39	573,513	20.7	3,222	21.1	576,735	20.7	
40	832,211	30.0	4,178	27.3	836,389	30.0	
41	627,148	22.6	2,716	17.8	629,864	22.6	
42	94,872	3.4	526	3.4	95,398	3.4	
43	1,987	0.1	18	0.1	2,005	0.1	
>43	478	0.0	0	0.0	478	0.0	
Missing	1,363		13		1,376		
Sex-gestation-specific birthweight centile							
1-3	114,353	4.1	947	6.2	115,300	4.1	<0.001
4-10	249,341	9.0	1,515	9.9	250,856	9.0	
11-20	331,223	11.9	1,884	12.3	333,107	11.9	
21-80	1,631,173	58.8	8,686	56.9	1,639,859	58.8	
81-90	236,924	8.5	1,152	7.5	238,076	8.5	
91-97	150,128	5.4	755	4.9	150,883	5.4	
98-100	62,087	2.2	331	2.2	62,418	2.2	
Missing	2,661		25		2,686		
5 min Apgar							
1-3	12,367	0.4	256	1.7	12,623	0.5	<0.001
4-6	25,735	0.9	343	2.3	26,078	0.9	
7-10	2,710,430	98.6	14,541	96.0	2,724,971	98.6	
Missing	29,358		155		29,513		

Table 25 Characteristics of children with and without attention deficit hyperactivity disorder (ADHD)

		No ADHD N=2,763,017		ADHD N=30,168		Total N=2,793,185		
		n	%	n	%	n	%	P value
Sociodemographic factors								
Sex								
	male	1,396,225	50.5	25,572	84.8	1,421,797	50.9	<0.001
	female	1,366,792	49.5	4,596	15.2	1,371,388	49.1	
	Missing	0		0		0		
Age category								
	<11	1,384,628	50.1	13,768	45.6	1,398,396	50.1	0.001
	11-14	665,313	24.1	9,179	30.4	674,492	24.1	
	>14	713,076	25.8	7,221	23.9	720,297	25.8	
	Missing	0		0		0		
Deprivation quintile								
	1 (most deprived)	616,528	22.3	9,629	32.0	626,157	22.4	<0.001
	2	551,140	20.0	7,397	24.6	558,537	20.0	
	3	535,277	19.4	5,423	18.0	540,700	19.4	
	4	543,024	19.7	4,580	15.2	547,604	19.6	
	5 (least deprived)	513,258	18.6	3,075	10.2	516,333	18.5	
	Missing	3,790		64		3,854		
Ethnic group								
	White	2,615,068	96.4	29,072	98.8	2,644,140	96.4	<0.001
	Asian	61,526	2.3	88	0.3	61,614	2.2	
	Black	5,354	0.2	19	0.1	5,373	0.2	
	Mixed	23,789	0.9	158	0.5	23,947	0.9	
	Other	7,693	0.3	86	0.3	7,779	0.3	
	Missing	49,587		745		50,332		
Maternity factors								
Maternal age (years)								
	<=24	737,997	26.7	13,607	45.1	751,604	26.9	<0.001
	25-29	796,503	28.8	7,918	26.2	804,421	28.8	
	30-34	806,826	29.2	5,822	19.3	812,648	29.1	
	>=35	421,638	15.3	2,821	9.4	424,459	15.2	
	Missing	53		0		53		
Maternal smoking								
	No	1,802,785	72.4	14,153	52.0	1,816,938	72.2	<0.001
	Yes	686,336	27.6	13,081	48.0	699,417	27.8	
	Missing	273,896		2,934		276,830		
Parity								
	0	1,241,856	45.2	13,546	45.4	1,255,402	45.2	<0.001
	1	956,004	34.8	9,270	31.0	965,274	34.8	
	>1	549,281	20.0	7,046	23.6	556,327	20.0	
	Missing	15,876		306		16,182		
Mode of delivery								
	SVD	1,790,256	64.8	19,231	63.7	1,809,487	64.8	0.705
	Cephalic	67,989	2.5	1,413	4.7	69,402	2.5	
	Assisted VD	330,685	12.0	3,236	10.7	333,921	12.0	
	Breech Delivery	7,895	0.3	82	0.3	7,977	0.3	
	Elective Caesarean	209,105	7.6	2,048	6.8	211,153	7.6	
	Emergency Caesarean	356,567	12.9	4,154	13.8	360,721	12.9	
	Other Delivery	515	0.0	4	0.0	519	0.0	
	Missing	5		0		5		

	No ADHD N=2,763,017		ADHD N=30,168		Total N=2,793,185		
	n	%	n	%	n	%	P value
Maternity factors							
Gestation (weeks)							
<24	100	0.0	0	0.0	100	0.0	<0.001
24-27	3,987	0.1	147	0.5	4,134	0.1	
28-32	25,448	0.9	560	1.9	26,008	0.9	
33-36	127,977	4.6	2,218	7.4	130,195	4.7	
37	135,647	4.9	1,950	6.5	137,597	4.9	
38	348,644	12.6	4,262	14.1	352,906	12.6	
39	570,638	20.7	6,097	20.2	576,735	20.7	
40	828,395	30.0	7,994	26.5	836,389	30.0	
41	623,869	22.6	5,995	19.9	629,864	22.6	
42	94,508	3.4	890	3.0	95,398	3.4	
43	1,959	0.1	46	0.2	2,005	0.1	
>43	474	0.0	4	0.0	478	0.0	
Missing	1,371		5		1,376		
Sex-gestation-specific birthweight centile							
1-3	113,621	4.1	1,679	5.6	115,300	4.1	<0.001
4-10	247,694	9.0	3,162	10.5	250,856	9.0	
11-20	329,204	11.9	3,903	12.9	333,107	11.9	
21-80	1,622,841	58.8	17,018	56.4	1,639,859	58.8	
81-90	235,829	8.5	2,247	7.5	238,076	8.5	
91-97	149,371	5.4	1,512	5.0	150,883	5.4	
98-100	61,785	2.2	633	2.1	62,418	2.2	
Missing	2,672		14		2,686		
5 min Apgar							
1-3	12,383	0.5	240	0.8	12,623	0.5	<0.001
4-6	25,653	0.9	425	1.4	26,078	0.9	
7-10	2,695,989	98.6	28,982	97.8	2,724,971	98.6	
Missing	28,992		521		29,513		

Table 26 Characteristics of children with and without depression

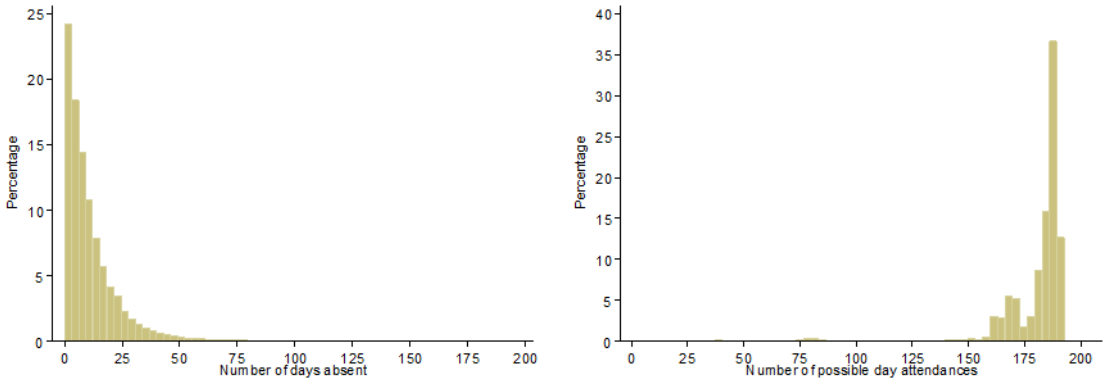
		No Depression N=2,772,829		Depression N=20,328		Total N=2,793,157		
		n	%	n	%	n	%	P value
Sociodemographic factors								
Sex								
	male	1,415,013	51.0	6,781	33.4	1,421,794	50.9	<0.001
	female	1,357,816	49.0	13,547	66.6	1,371,363	49.1	
	Missing	0		0		0		
Age category								
	<11	1,396,081	50.3	2,315	11.4	1,398,396	50.1	<0.001
	11-14	668,063	24.1	6,425	31.6	674,488	24.1	
	>14	708,685	25.6	11,588	57.0	720,273	25.8	
	Missing	0		0		0		
Deprivation quintile								
	1 (most deprived)	622,240	22.5	3,912	19.3	626,152	22.4	0.001
	2	554,415	20.0	4,108	20.2	558,523	20.0	
	3	536,288	19.4	4,407	21.7	540,695	19.4	
	4	543,503	19.6	4,097	20.2	547,600	19.6	
	5 (least deprived)	512,569	18.5	3,764	18.6	516,333	18.5	
	Missing	3,814		40		3,854		
Ethnic group								
	White	2,624,459	96.4	19,653	97.6	2,644,112	96.4	0.004
	Asian	61,342	2.3	272	1.4	61,614	2.2	
	Black	5,368	0.2	5	0.0	5,373	0.2	
	Mixed	23,784	0.9	163	0.8	23,947	0.9	
	Other	7,736	0.3	43	0.2	7,779	0.3	
	Missing	50,140		192		50,332		
Maternity factors								
Maternal age (years)								
	<=24	746,269	26.9	5,335	26.2	751,604	26.9	0.027
	25-29	798,007	28.8	6,396	31.5	804,403	28.8	
	30-34	806,749	29.1	5,889	29.0	812,638	29.1	
	>=35	421,751	15.2	2,708	13.3	424,459	15.2	
	Missing	53		0		53		
Maternal smoking								
	No	1,804,453	72.2	12,468	69.4	1,816,921	72.2	<0.001
	Yes	693,912	27.8	5,494	30.6	699,406	27.8	
	Missing	274,464		2,366		276,830		
Parity								
	0	1,246,276	45.2	9,117	44.9	1,255,393	45.2	0.747
	1	958,016	34.8	7,246	35.7	965,262	34.8	
	>1	552,385	20.0	3,935	19.4	556,320	20.0	
	Missing	16,152		30		16,182		
Mode of delivery								
	SVD	1,795,891	64.8	13,578	66.8	1,809,469	64.8	0.006
	Cephalic	69,037	2.5	365	1.8	69,402	2.5	
	Assisted VD	331,511	12.0	2,410	11.9	333,921	12.0	
	Breech Delivery	7,909	0.3	68	0.3	7,977	0.3	
	Elective Caesarean	209,571	7.6	1,582	7.8	211,153	7.6	
	Emergency Caesarean	358,386	12.9	2,325	11.4	360,711	12.9	
	Other Delivery	519	0.0	0	0.0	519	0.0	
	Missing	5		0		5		

	No Depression N=2,772,829		Depression N=20,328		Total N=2,793,157		
	n	%	n	%	n	%	P value
Maternity factors							
Gestation (weeks)							
<24	100	0.0	0	0.0	100	0.0	0.040
24-27	4,093	0.1	41	0.2	4,134	0.1	
28-32	25,767	0.9	241	1.2	26,008	0.9	
33-36	129,212	4.7	978	4.8	130,190	4.7	
37	136,608	4.9	989	4.9	137,597	4.9	
38	350,173	12.6	2,727	13.4	352,900	12.6	
39	572,596	20.7	4,130	20.3	576,726	20.7	
40	830,170	30.0	6,214	30.6	836,384	30.0	
41	625,629	22.6	4,232	20.8	629,861	22.6	
42	94,667	3.4	731	3.6	95,398	3.4	
43	1,976	0.1	29	0.1	2,005	0.1	
>43	474	0.0	4	0.0	478	0.0	
Missing	1,364		12		1,376		
Sex-gestation-specific birthweight centile							
1-3	114,416	4.1	882	4.3	115,298	4.1	0.015
4-10	248,909	9.0	1,947	9.6	250,856	9.0	
11-20	330,401	11.9	2,706	13.3	333,107	11.9	
21-80	1,628,258	58.8	11,587	57.1	1,639,845	58.8	
81-90	236,365	8.5	1,707	8.4	238,072	8.5	
91-97	149,849	5.4	1,026	5.1	150,875	5.4	
98-100	61,966	2.2	452	2.2	62,418	2.2	
Missing	2,665		21		2,686		
5 min Apgar							
1-3	12,490	0.5	133	0.7	12,623	0.5	0.182
4-6	25,902	0.9	176	0.9	26,078	0.9	
7-10	2,704,981	98.6	19,962	98.5	2,724,943	98.6	
Missing	29,456		57		29,513		

5.3 Number of days absent from school

The final absence dataset contained 1,597,397 pupil records after retaining only census years 2009, 2010 and 2012. The histograms in Figure 6 below show the distribution of the number of days absent and the possible number of days attended per year for all 702,210 pupils included in the final analyses. The number of days absent during a census year ranged from 0 to 190 days. The mean was 11.71 days (SD 13.90) and the median was 8 days (IQR 11.5). The number of possible attendances ranged from 0.5 to 206 days. The mean was 178.84 days (SD 21.20) and the median was 186 days (IQR 9.5).

Figure 6 Number of days absent and number of possible day attendances per census year



The absence and possible attendance distributions were very over dispersed because the variances were much higher than would be expected for a Poisson distribution where the mean and variance should be roughly equal. The mean number of absences was 11.71 and the variance was 193.14 whilst the mean number of possible days attended was 178.84 and the variance was 449.52. Therefore these were best described via a negative binomial distribution.

5.3.1 Characteristics of children by number of days absent

Diabetes, asthma, epilepsy, ADHD and depression were all significantly associated with an increase in median annual number of days absent from school (Table 27). The median number of days absent was significantly greater amongst girls, older pupils, more deprived pupils and those with lower 5-minute Apgar scores, lower sex-gestation-specific birthweight centiles and gestational ages less than 27 weeks or greater than 43 weeks. Children whose mothers were younger, smoked during pregnancy and had increased maternal parity also had a significantly greater median number of days absent. Children of black descent had a lower median number of days absent whilst children of Asian descent had increased absenteeism. Finally children born by caesarean section or assisted vaginal delivery had a significantly lower median number of days absent (Table 28).

Table 27 Number of days absent from school by presence or absence of chronic conditions

		N	mean	SD	median	IQR	P value
Diabetes							
	No	1,591,411	11.7	13.9	8.0	11.5	<0.001
	Yes	5,986	16.4	17.8	11.0	15.5	
	Missing	0					
Asthma							
	No	1,517,852	11.5	13.8	7.5	11.5	<0.001
	Yes	79,545	14.8	15.9	10.0	14.5	
	Missing	0					
Epilepsy							
	No	1,589,164	11.7	13.8	7.5	11.5	<0.001
	Yes	8,233	18.8	23.3	11.0	18.5	
	Missing	0					
ADHD							
	No	1,580,137	11.7	13.8	7.5	11.5	<0.001
	Yes	17,260	15.6	18.4	10.0	16.0	
	Missing	0					
Depression							
	No	1,584,916	11.6	13.7	7.5	11.5	<0.001
	Yes	12,481	25.0	28.5	15.5	24.5	
	Missing	0					

Table 28 Characteristics of children by number of days absent

		N	mean	SD	median	IQR	P value
Sociodemographic factors							
Sex							
	male	812,566	11.5	13.8	7.5	11.5	<0.001
	female	784,831	11.9	14.0	8.0	11.5	
	Missing	0					
Age category							
	<11	799,224	9.0	9.4	6.5	9.0	<0.001
	11-14	390,450	13.1	15.5	8.5	13.0	
	>14	407,723	15.6	18.0	10.5	15.5	
	Missing	0					
Deprivation quintile							
	1 (most deprived)	365,474	16.2	17.5	11.0	16.0	<0.001
	2	317,814	13.1	14.8	9.0	13.0	
	3	307,128	11.0	12.6	7.5	10.5	
	4	312,826	9.4	11.0	6.5	9.0	
	5 (least deprived)	291,879	7.8	9.4	5.5	8.0	
	Missing	2,276					
Ethnic group							
	White	1,509,679	11.7	13.9	7.5	11.5	<0.001
	Asian	34,732	12.5	12.1	9.5	12.5	
	Black	2,535	9.0	10.1	6.0	9.5	
	Mixed	13,341	11.3	13.4	7.5	11.5	
	Other	5,162	15.1	18.9	9.5	15.0	
	Missing	31,948					
Maternity factors							
Maternal age (years)							
	<=24	430,791	14.8	15.9	10.5	14.0	<0.001
	25-29	467,530	11.6	13.7	8.0	11.5	
	30-34	464,329	9.9	12.3	6.5	9.5	
	>=35	234,715	9.9	12.5	6.5	9.5	
	Missing	32					
Maternal smoking							
	No	1,030,586	9.8	11.7	6.5	9.5	<0.001
	Yes	404,236	15.7	17.0	11.0	15.5	
	Missing	162,575					
Parity							
	0	716,017	10.9	12.6	7.5	11.0	<0.001
	1	553,754	11.3	13.5	7.5	11.5	
	>1	318,965	14.3	16.8	9.5	14.5	
	Missing	8,661					
Mode of delivery							
	SVD	1,044,081	12.2	14.5	8.0	12.0	<0.001
	Cephalic	38,457	12.4	14.7	8.0	12.0	
	Assisted VD	189,934	10.3	12.2	7.0	10.0	
	Breech Delivery	4,623	12.6	14.9	8.0	12.0	
	Elective Caesarean	118,293	11.0	13.0	7.5	11.0	
	Emergency Caesarean	201,728	10.7	12.6	7.0	11.0	
	Other Delivery	279	11.2	12.3	8.0	11.0	
	Missing	2					

	N	mean	SD	median	IQR	P value
Maternity factors						
Gestation (weeks)						
<24	51	10.7	13.5	6.0	11.0	<0.001
24-27	2,337	12.5	14.9	9.0	12.0	
28-32	14,791	12.5	14.5	8.0	12.5	
33-36	74,142	12.7	14.9	8.5	13.0	
37	78,813	12.7	15.0	8.5	12.0	
38	202,450	12.2	14.4	8.0	12.0	
39	327,339	11.6	13.9	7.5	11.5	
40	480,730	11.6	13.8	7.5	11.5	
41	358,511	11.2	13.2	7.5	11.0	
42	55,900	11.5	14.0	7.5	11.0	
43	1,172	11.9	14.0	8.0	12.0	
>43	285	12.3	13.3	9.0	11.5	
Missing	876					
Sex-gestation-specific birthweight centile						
1-3	66,428	14.2	16.2	9.5	14.5	<0.001
4-10	144,204	13.1	15.1	9.0	13.0	
11-20	191,088	12.5	14.7	8.0	12.0	
21-80	937,290	11.5	13.6	7.5	11.0	
81-90	135,550	10.6	12.7	7.0	10.5	
91-97	85,974	10.6	12.7	7.0	10.5	
98-100	35,278	10.5	12.4	7.0	10.5	
Missing	1,585					
5 min Apgar						
1-3	7,375	13.6	17.0	8.5	13.0	<0.001
4-6	14,922	12.3	14.6	8.0	12.0	
7-10	1,559,816	11.7	13.9	8.0	11.5	
Missing	15,284					

5.3.2 Association between diabetes and number of days absent

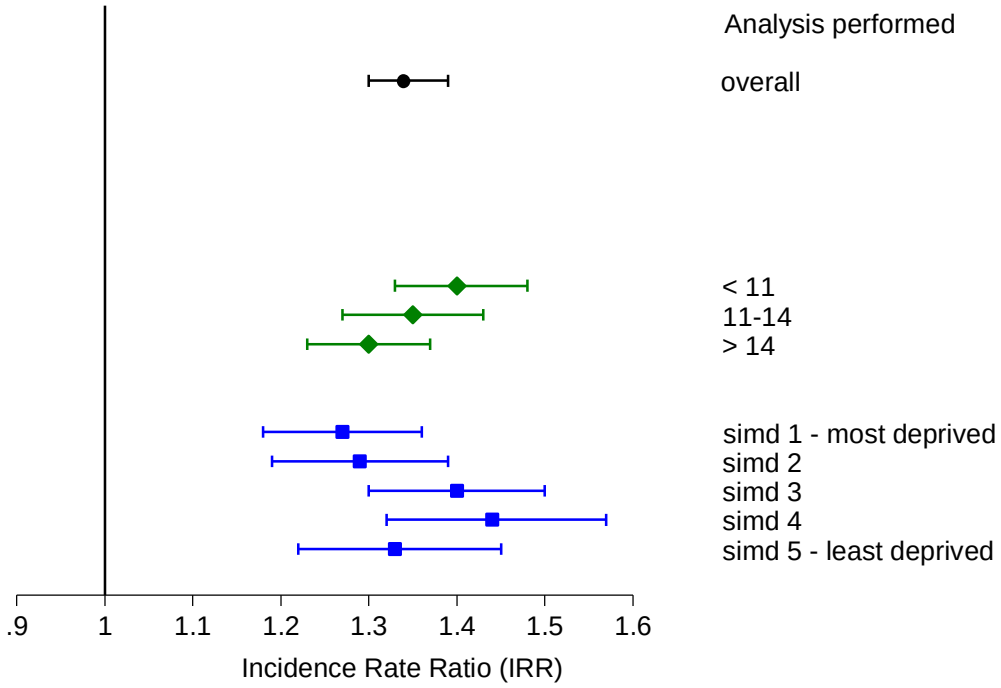
The diabetes analysis dataset contained 1,596,954 pupil census records after excluding 443 where ineligible diabetes drugs had been prescribed (section 4.3.2.2.1). Children with diabetes had significantly higher numbers of days absent from school on univariate analysis (IRR 1.42, 95% CI 1.36-1.47) and after adjusting for sociodemographic (IRR 1.31, 95% CI 1.27-1.36) and maternity (IRR 1.34, 95% CI 1.30-1.39) confounders. In the fully adjusted model, diabetic pupils had a 1.34 fold risk of absence (34% increased risk) compared to pupils without diabetes. There were significant univariate interactions between diabetes and both age (p=0.036) and deprivation (p=0.043) but not with gender (p=0.187). On subgroup analyses the association between diabetes and absence was stronger for younger and less deprived pupils; however, the interactions of diabetes with age (p=0.126) and deprivation (p=0.189) were no longer significant after adjusting for sociodemographic and maternity confounders. Table 29 presents incidence rate ratios and confidence intervals from univariate and multivariate analyses; overall and within age and deprivation subgroups. The fully adjusted multivariate incidence rate ratios and confidence intervals are summarised visually in Figure 7.

Table 29 Association between diabetes and absence from school by age and area deprivation

	Univariate model		Multivariate[1]		Multivariate[2]	
	IRR	95% CI	IRR	95% CI	IRR	95% CI
Overall	1.42***	1.36-1.47	1.31***	1.27-1.36	1.34***	1.30-1.39
< 11 years	1.36***	1.28-1.44	1.39***	1.32-1.46	1.40***	1.33-1.48
11 - 14 years	1.28***	1.21-1.36	1.31***	1.23-1.39	1.35***	1.27-1.43
> 14 years	1.23***	1.16-1.29	1.25***	1.19-1.32	1.30***	1.23-1.37
1 - most deprived	1.32***	1.22-1.42	1.21***	1.12-1.29	1.27***	1.18-1.36
2	1.40***	1.30-1.52	1.26***	1.17-1.36	1.29***	1.19-1.39
3	1.54***	1.43-1.67	1.39***	1.29-1.49	1.40***	1.30-1.50
4	1.51***	1.39-1.65	1.40***	1.29-1.53	1.44***	1.32-1.57
5 - least deprived	1.45***	1.33-1.60	1.31***	1.20-1.43	1.33***	1.22-1.45

Exponentiated coefficients; 95% confidence intervals in brackets
* p<0.05, ** p<0.01, *** p<0.001
[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)
[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

Figure 7 Forest plot of the fully adjusted multivariate association between diabetes and absence from school by age and area deprivation



5.3.3 Association between asthma and number of days absent

The asthma analysis dataset contained 1,415,923 pupil census records after excluding 181,474 where ineligible asthma drugs had been prescribed (section 4.3.2.2.2). Asthma was significantly associated with the number of days absent from school on univariate analysis (IRR 1.30, 95% CI 1.28-1.31) and after adjusting for sociodemographic (IRR 1.25, 95% CI 1.24-1.26) and maternity (IRR 1.25, 95% CI 1.24-1.26) confounders. Pupils with asthma had a 1.25 fold risk of absence (25% increased risk) compared to pupils who did not have asthma. There was a significant univariate interaction with gender (p=0.003) whereby the association between asthma and absence was stronger in girls than boys; however the interaction was no longer significant after adjusting for sociodemographic and maternity confounders (p=0.447). There were significant univariate and multivariate interactions between asthma and age category (both p<0.001) and deprivation (both p<0.001) and significant trends whereby the associations between asthma and absenteeism was stronger for younger and less deprived pupils. Within the subgroup of pupils over 14 years of age, on fully adjusted analyses, children with asthma had a 15% increased risk of school absence compared to children without asthma. However, the risk of school absence

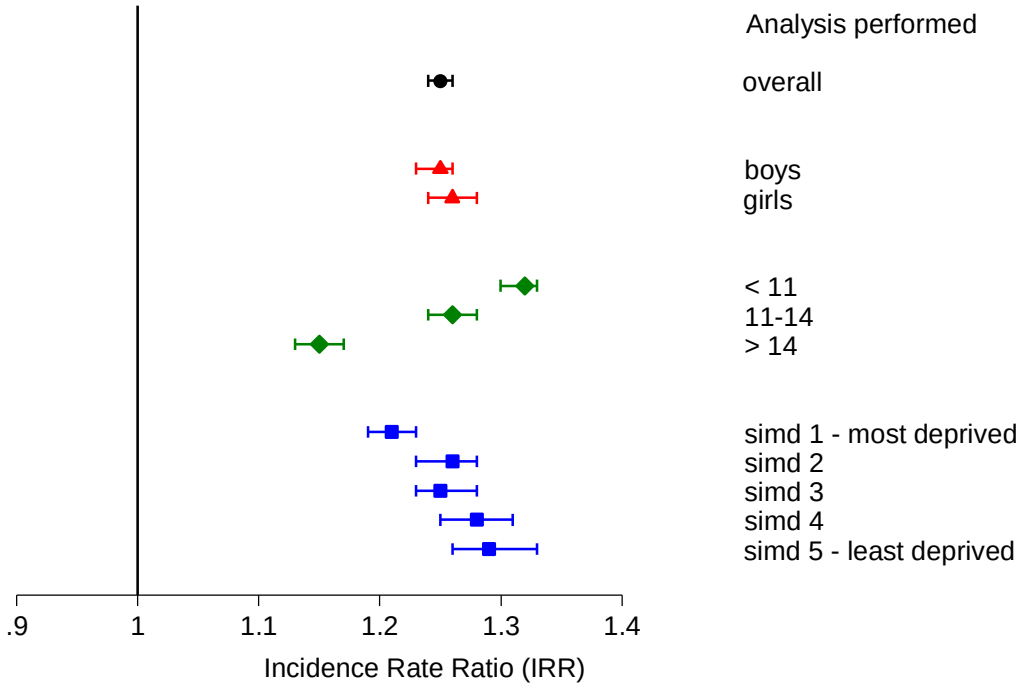
for asthmatic versus non asthmatic children was 26% higher and 32% higher in the 11-14 and <11 age groups respectively. Similarly, within the most deprived 20% of the population, asthmatic children had a 21% increased risk of school absence (IRR 1.21) compared to non-asthmatic children which was significantly lower than the corresponding increased risk found within the less deprived subgroups. The relative impact of asthma on absenteeism was smaller for pupils who lived in more socioeconomically deprived areas because, among unaffected children, they already experienced higher absenteeism (medians, 11.0 and 5.0 days in the most and least deprived quintiles, respectively). Among asthmatic children, absences were still more common in the most deprived quintile (medians, 14.0 vs. 7.0 days). Similarly, the relative impact was smaller for older children because, among unaffected pupils, they already had more absence. Table 30 presents incidence rate ratios and confidence intervals from univariate and multivariate analyses; overall and within sex, age and deprivation subgroups. The fully adjusted multivariate incidence rate ratios and confidence intervals are summarised visually in Figure 8.

Table 30 Association between asthma and absence from school by sex, age and area deprivation

	Univariate model		Multivariate[1]		Multivariate[2]	
	IRR	95% CI	IRR	95% CI	IRR	95% CI
Overall	1.30***	1.28-1.31	1.25***	1.24-1.26	1.25***	1.24-1.26
Boys	1.28***	1.27-1.30	1.24***	1.23-1.26	1.25***	1.23-1.26
Girls	1.32***	1.30-1.35	1.26***	1.24-1.28	1.26***	1.24-1.28
< 11 years	1.38***	1.36-1.40	1.32***	1.31-1.34	1.32***	1.30-1.33
11 - 14 years	1.29***	1.26-1.31	1.24***	1.22-1.26	1.26***	1.24-1.28
> 14 years	1.17***	1.15-1.19	1.14***	1.12-1.16	1.15***	1.13-1.17
1 - most deprived	1.19***	1.17-1.21	1.19***	1.17-1.21	1.21***	1.19-1.23
2	1.25***	1.23-1.28	1.26***	1.23-1.28	1.26***	1.23-1.28
3	1.26***	1.24-1.29	1.26***	1.23-1.29	1.25***	1.23-1.28
4	1.30***	1.27-1.33	1.29***	1.26-1.32	1.28***	1.25-1.31
5 - least deprived	1.32***	1.28-1.36	1.31***	1.27-1.35	1.29***	1.26-1.33

Exponentiated coefficients; 95% confidence intervals in brackets
* p<0.05, ** p<0.01, *** p<0.001
[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)
[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

Figure 8 Forest plot of the fully adjusted multivariate association between asthma and absence from school by sex, age and area deprivation



5.3.4 Association between epilepsy and number of days absent

The epilepsy analysis dataset contained 1,597,397 records. Children with epilepsy had significantly higher numbers of days absent from school on univariate analysis (IRR 1.62, 95% CI 1.56-1.67) and after adjusting for sociodemographic (IRR 1.50, 95% CI 1.45-1.56) and maternity (IRR 1.50, 95% CI 1.45-1.55) confounders. In the final multivariate analyses, pupils with epilepsy had a 1.5 fold risk of absence (50% increased risk) compared to pupils without epilepsy. There were significant univariate and multivariate interactions between epilepsy and gender (both $p<0.001$), age category (both $p<0.001$) and SIMD deprivation category (both $p<0.001$). On fully adjusted subgroup analyses, the association between epilepsy and absence was stronger in girls than boys (IRR 1.61 versus IRR 1.39). Girls with epilepsy experienced a 61% increased risk of absence compared to girls without epilepsy whereas the corresponding increased risk for epileptic boys was only 39%. Absolute absenteeism was higher in girls than boys both in unaffected children (medians, 8.0 vs. 7.5 days) and in those with epilepsy (medians, 12.0 vs. 11.0 days). There were clear trends whereby the association between epilepsy and school absence was stronger for younger and less deprived pupils. Within the subgroup of pupils younger than 11 years of age, children with epilepsy experienced a 66% increased risk of absence (IRR

1.66) compared to children without epilepsy. This was significantly greater than the corresponding increase in risk found within pupils older than 14 years of age (IRR 1.37). The increase in risk of absence associated with epilepsy was 23% (IRR 1.23) for children in the most deprived 20% of the population which was significantly lower than the corresponding increase in risk observed within less deprived subgroups. The relative impact of epilepsy on absenteeism was smaller for pupils who lived in more socioeconomically deprived areas because, within unaffected children, they already experienced higher absenteeism (medians, 11.0 and 5.5 days in the most and least deprived quintiles, respectively). Among epileptic children, absences were still more common in the most deprived quintile (medians, 13.5 vs. 8.5 days). Similarly, the relative impact was smaller in older pupils because, among unaffected children, they had higher absolute rates of absenteeism compared to younger pupils. Table 31 presents incidence rate ratios and confidence intervals from the univariate and multivariate analyses; overall and within sex, age and deprivation subgroups. The fully adjusted multivariate incidence rate ratios and confidence intervals are summarised visually in Figure 9.

Table 31 Association between epilepsy and absence from school by sex, age and area deprivation

	Univariate model		Multivariate[1]		Multivariate[2]	
	IRR	95% CI	IRR	95% CI	IRR	95% CI
Overall	1.62***	1.56-1.67	1.50***	1.45-1.56	1.50***	1.45-1.55
Boys	1.52***	1.45-1.59	1.40***	1.33-1.47	1.39***	1.33-1.46
Girls	1.72***	1.64-1.81	1.60***	1.53-1.69	1.61***	1.53-1.69
< 11 years	1.70***	1.61-1.79	1.67***	1.57-1.77	1.66***	1.56-1.76
11 - 14 years	1.49***	1.40-1.58	1.49***	1.41-1.58	1.51***	1.42-1.60
> 14 years	1.36***	1.29-1.43	1.37***	1.30-1.44	1.37***	1.30-1.45
1 - most deprived	1.29***	1.22-1.37	1.21***	1.14-1.28	1.23***	1.16-1.30
2	1.51***	1.40-1.62	1.42***	1.33-1.52	1.44***	1.35-1.54
3	1.77***	1.63-1.91	1.64***	1.51-1.78	1.61***	1.48-1.75
4	1.83***	1.68-1.99	1.65***	1.53-1.79	1.63***	1.50-1.77
5 - least deprived	2.01***	1.82-2.21	1.83***	1.66-2.02	1.81***	1.64-1.99

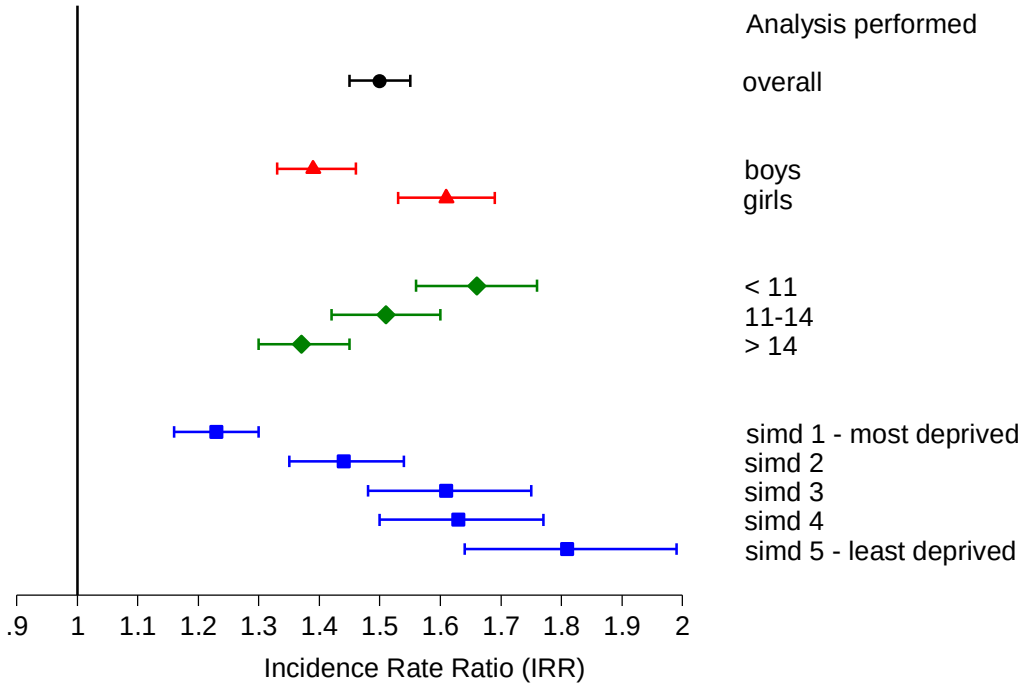
Exponentiated coefficients; 95% confidence intervals in brackets

* p<0.05, ** p<0.01, *** p<0.001

[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)

[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

Figure 9 Forest plot of the fully adjusted multivariate association between epilepsy and absence from school by sex, age and area deprivation



5.3.5 Association between attention deficit hyperactivity disorder and number of days absent

The ADHD analysis dataset contained 1,597,397 records. Children with ADHD had significantly higher numbers of days absent at school on univariate analysis (IRR 1.37, 95% CI 1.33-1.40) and after adjusting for sociodemographic (IRR 1.29, 95% CI 1.26-1.31) and maternity (IRR 1.18, 95% CI 1.16-1.21) confounders. In the final multivariate analyses, pupils with ADHD had a 1.18 fold risk of absence (18% increased risk) compared to pupils without ADHD. There was no significant univariate interaction with gender ($p=0.221$); however, there were significant univariate and multivariate interactions between ADHD and age (both $p<0.001$) and deprivation (both $p<0.001$). On fully adjusted subgroup analyses, within children younger than eleven years of age, those who had ADHD experienced an 8% increased risk of absence compared to children without ADHD. This was significantly lower than the corresponding increase in risk observed within older age groups. There was no clear trend across different categories of deprivation; however, the association between ADHD and absence was significantly weaker for children in the most deprived 20% of the population than for children in the least deprived 20% (10% versus 31% increased risk of absence respectively). The relative impact of ADHD on absenteeism was smaller for pupils who lived in more socioeconomically deprived areas

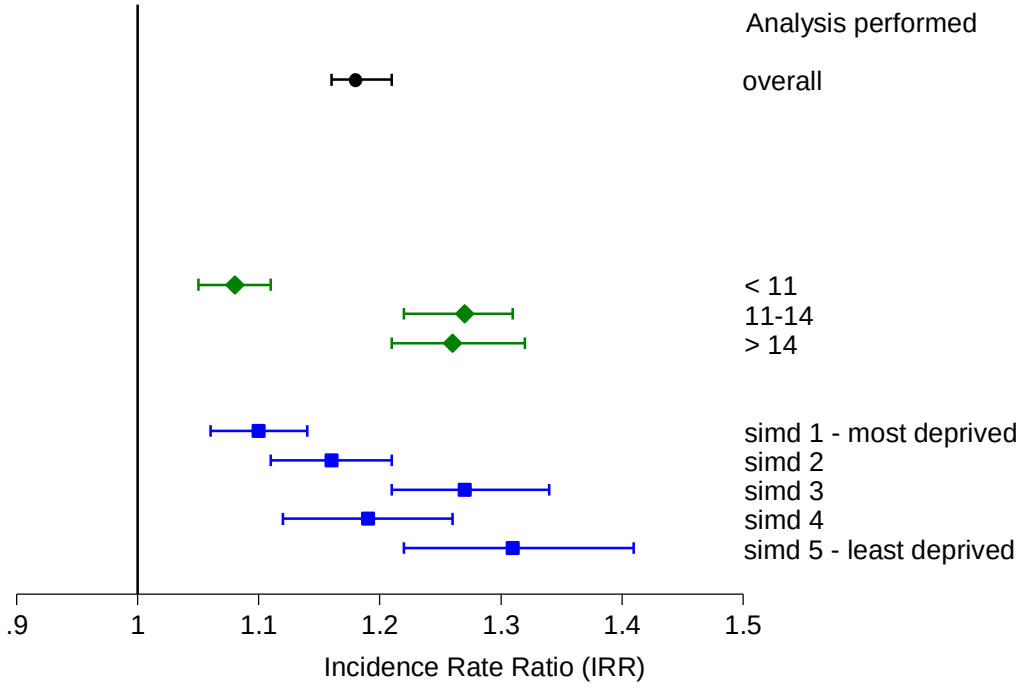
because, within unaffected children, they already experienced higher absenteeism (medians, 11.0 and 5.5 days in the most and least deprived quintiles, respectively). Among children with ADHD, absences were still more common in the most deprived quintile (medians, 12.0 vs. 6.5 days). One possible explanation of the greater relative impact of ADHD on absenteeism in older pupils may be that pupils with ADHD who are older are more likely to miss school through truancy or school refusal compared to younger pupils with ADHD whose attendance may be more closely monitored by their parents. Table 32 presents incidence rate ratios and confidence intervals from univariate and multivariate analyses; overall and within age and deprivation subgroups. The fully adjusted multivariate incidence rate ratios and confidence intervals are summarised visually in Figure 10.

Table 32 Association between attention deficit hyperactivity disorder and absence from school by age and area deprivation

	Univariate model		Multivariate[1]		Multivariate[2]	
	IRR	95% CI	IRR	95% CI	IRR	95% CI
Overall	1.37***	1.33-1.40	1.29***	1.26-1.31	1.18***	1.16-1.21
< 11 years	1.24***	1.21-1.28	1.18***	1.14-1.21	1.08***	1.05-1.11
11 - 14 years	1.49***	1.44-1.54	1.38***	1.33-1.43	1.27***	1.22-1.31
> 14 years	1.42***	1.36-1.48	1.37***	1.32-1.44	1.26***	1.21-1.32
1 - most deprived	1.16***	1.11-1.20	1.15***	1.11-1.19	1.10***	1.06-1.14
2	1.27***	1.21-1.33	1.27***	1.22-1.32	1.16***	1.11-1.21
3	1.42***	1.35-1.51	1.41***	1.34-1.48	1.27***	1.21-1.34
4	1.28***	1.20-1.36	1.33***	1.26-1.41	1.19***	1.12-1.26
5 - least deprived	1.47***	1.36-1.59	1.48***	1.38-1.59	1.31***	1.22-1.41

Exponentiated coefficients; 95% confidence intervals in brackets
* p<0.05, ** p<0.01, *** p<0.001
[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)
[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

Figure 10 Forest plot of the fully adjusted multivariate association between attention deficit hyperactivity disorder and absence from school by age and area deprivation



5.3.6 Association between depression and number of days absent

The depression analysis dataset contained 1,597,379 pupil census records after excluding 18 where ineligible anti-depressant drugs had been prescribed (section 4.3.2.2.5). Children with depression had significantly higher numbers of days absent at school on univariate analysis (IRR 2.23, 95% CI 2.17-2.29) and after adjusting for sociodemographic (IRR 1.91, 95% CI 1.86-1.96) and maternity (IRR 1.95, 95% CI 1.90-2.00) confounders. In the fully adjusted model, pupils with depression had a 1.95 fold risk of absence (95% increased risk) compared to pupils without depression. There was no significant univariate interaction between depression and gender (p=0.762); however, there were significant univariate and multivariate interactions with age (both p<0.001) and deprivation (both p<0.001). Subsequent subgroup analyses demonstrated trends whereby the association between depression and absence was stronger for older and less deprived pupils. In the fully adjusted analyses children with depression in the <11 age group experienced a 61% increased risk of absence compared to similarly aged children without depression and this was significantly lower than the corresponding increased risk observed within older age groups. Within the least deprived 20% of the population, depressed children experienced a

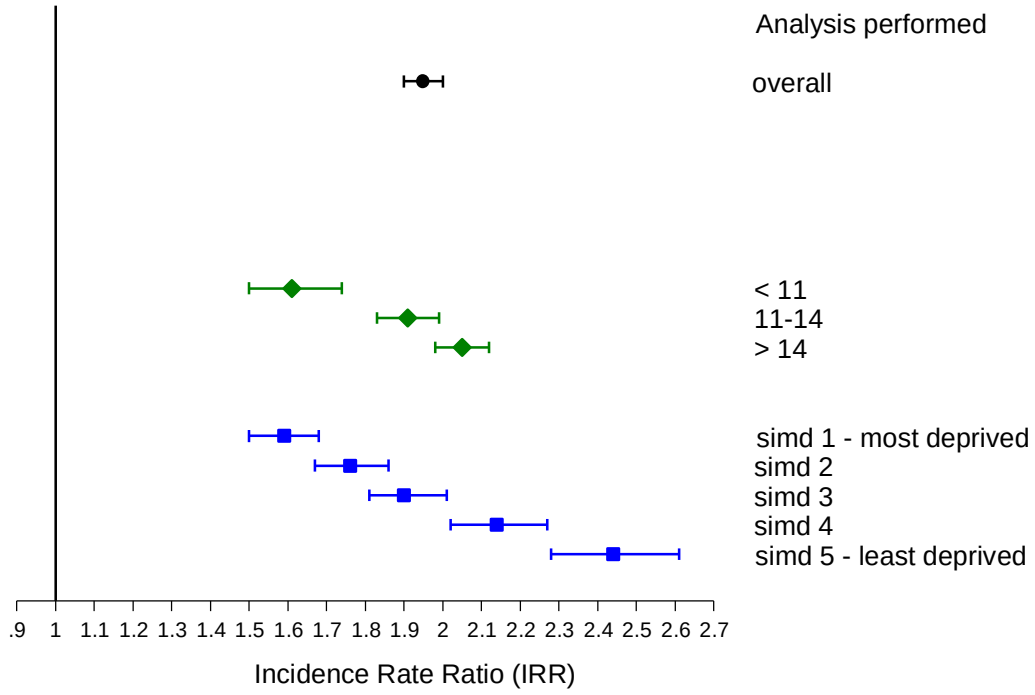
2.44 times greater risk of absence compared to children without depression which was significantly higher than that observed within more deprived subgroups. The relative impact of depression on absenteeism was smaller for pupils who lived in more socioeconomically deprived areas because, among unaffected children, they already experienced higher absenteeism (medians, 11.0 and 5.5 days in the most and least deprived quintiles, respectively). Among depressed children, absences were still more common in the most deprived quintile (medians, 20.0 vs. 12.5 days). One possible explanation of the greater relative impact of depression on absenteeism in older pupils may be that older depressed pupils are more likely to miss school through truancy or school refusal compared to younger depressed pupils whose attendance may be more closely monitored by their parents. Table 33 presents incidence rate ratios and confidence intervals from univariate and multivariate analyses; overall and within age and deprivation subgroups. The fully adjusted multivariate incidence rate ratios and confidence intervals are summarised visually in Figure 11.

Table 33 Association between depression and absence from school by age and area deprivation

	Univariate model		Multivariate[1]		Multivariate[2]	
	IRR	95% CI	IRR	95% CI	IRR	95% CI
Overall	2.23***	2.17-2.29	1.91***	1.86-1.96	1.95***	1.90-2.00
< 11 years	1.56***	1.45-1.68	1.60***	1.49-1.73	1.61***	1.50-1.74
11 - 14 years	1.80***	1.73-1.87	1.89***	1.81-1.96	1.91***	1.83-1.99
> 14 years	1.90***	1.84-1.96	1.98***	1.92-2.05	2.05***	1.98-2.12
1 - most deprived	1.85***	1.76-1.95	1.53***	1.45-1.61	1.59***	1.50-1.68
2	2.12***	2.01-2.24	1.71***	1.62-1.80	1.76***	1.67-1.86
3	2.29***	2.17-2.41	1.87***	1.77-1.97	1.90***	1.81-2.01
4	2.59***	2.45-2.74	2.12***	2.00-2.24	2.14***	2.02-2.27
5 - least deprived	2.97***	2.78-3.18	2.41***	2.25-2.58	2.44***	2.28-2.61

Exponentiated coefficients; 95% confidence intervals in brackets
* p<0.05, ** p<0.01, *** p<0.001
[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)
[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

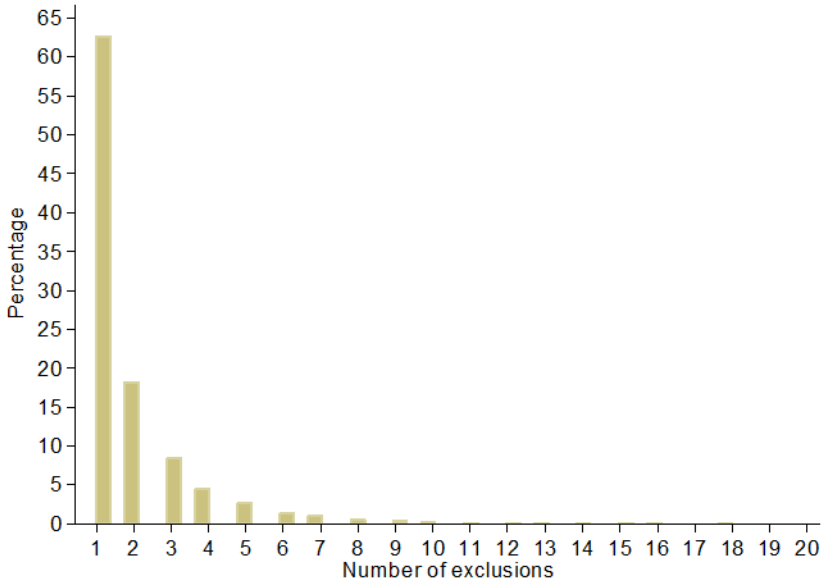
Figure 11 Forest plot of the fully adjusted multivariate association between depression and absence from school by age and area deprivation



5.4 Number of school exclusions

The final exclusion dataset contained 1,597,397 pupil census records after retaining only census years 2009, 2010 and 2012. As expected, the vast majority of children (97.9%) had no exclusions; therefore, to produce a clearer distribution of numbers of exclusions a histogram was produced only for those children who had one or more exclusions in a census year (Figure 12). The distribution was skewed with 60% of children experiencing only one exclusion episode within a school census year. The number of exclusions ranged from 1 to 18 per year. The mean was 1.82 exclusions per year (SD 1.49) and the median was 1 per year (IQR 1). The exclusion distribution was very over dispersed because the variance was higher than would be expected for a Poisson distribution where the mean and variance should be roughly equal. The mean number of exclusions was 1.82 per year and the variance was 2.22 per year therefore exclusions were most accurately modelled using a negative binomial distribution.

Figure 12 Histogram of number of exclusions for children excluded from school



5.4.1 Characteristics of children by whether or not they were excluded

Each chronic condition, with the exception of diabetes, was significantly associated with exclusion from school. Children with asthma, epilepsy, ADHD or depression were more likely to be excluded from school (Table 34). Significantly more boys than girls were excluded from school. Children who were excluded from school were significantly older, more deprived and significantly more likely to be of white ethnicity. They had significantly lower sex-gestation-specific birthweight centiles, 5-minute Apgar scores and gestational ages and their mothers were significantly more likely to have been younger than 25 years of age, multiparous and to have smoked during pregnancy. Children born via spontaneous vaginal delivery were also more likely to be excluded (Table 35).

Table 34 Exclusion from school by presence or absence of chronic conditions

		Not Excluded N=1,564,143		Excluded N=33,254		Total N=1,597,397		
		n	%	n	%	n	%	P value
Diabetes								
	No	1,558,292	99.6	33,119	99.6	1,591,411	99.6	0.407
	Yes	5,851	0.4	135	0.4	5,986	0.4	
	Missing	0		0		0		
Asthma								
	No	1,486,578	95.0	31,274	94.0	1,517,852	95.0	<0.001
	Yes	77,565	5.0	1,980	6.0	79,545	5.0	
	Missing	0		0		0		
Epilepsy								
	No	1,556,120	99.5	33,044	99.4	1,589,164	99.5	0.010
	Yes	8,023	0.5	210	0.6	8,233	0.5	
	Missing	0		0		0		
ADHD								
	No	1,549,572	99.1	30,565	91.9	1,580,137	98.9	<0.001
	Yes	14,571	0.9	2,689	8.1	17,260	1.1	
	Missing	0		0		0		
Depression								
	No	1,552,134	99.2	32,782	98.6	1,584,916	99.2	<0.001
	Yes	12,009	0.8	472	1.4	12,481	0.8	
	Missing	0		0		0		

Table 35 Characteristics of children by whether or not they were excluded from school

		Not Excluded N=1,564,143		Excluded N=33,254		Total N=1,597,397		
		n	%	n	%	n	%	P value
Sociodemographic factors								
Sex								
	male	787,200	50.3	25,366	76.3	812,566	50.9	<0.001
	female	776,943	49.7	7,888	23.7	784,831	49.1	
	Missing	0		0		0		
Age category								
	<11	794,993	50.8	4,231	12.7	799,224	50.0	<0.001
	11-14	375,346	24.0	15,104	45.4	390,450	24.4	
	>14	393,804	25.2	13,919	41.9	407,723	25.5	
	Missing	0		0		0		
Deprivation quintile								
	1 (most deprived)	351,131	22.5	14,343	43.2	365,474	22.9	<0.001
	2	309,692	19.8	8,122	24.5	317,814	19.9	
	3	301,994	19.3	5,134	15.5	307,128	19.3	
	4	309,255	19.8	3,571	10.8	312,826	19.6	
	5 (least deprived)	289,860	18.6	2,019	6.1	291,879	18.3	
	Missing	2,211		65		2,276		
Ethnic group								
	White	1,477,589	96.4	32,090	98.1	1,509,679	96.4	<0.001
	Asian	34,429	2.2	303	0.9	34,732	2.2	
	Black	2,511	0.2	24	0.1	2,535	0.2	
	Mixed	13,146	0.9	195	0.6	13,341	0.9	
	Other	5,058	0.3	104	0.3	5,162	0.3	
	Missing	31,410		538		31,948		
Maternity factors								
Maternal age (years)								
	<=24	414,665	26.5	16,126	48.5	430,791	27.0	<0.001
	25-29	458,594	29.3	8,936	26.9	467,530	29.3	
	30-34	458,592	29.3	5,737	17.3	464,329	29.1	
	>=35	232,260	14.8	2,455	7.4	234,715	14.7	
	Missing	32		0		32		
Maternal smoking								
	No	1,018,238	72.4	12,348	42.8	1,030,586	71.8	<0.001
	Yes	387,761	27.6	16,475	57.2	404,236	28.2	
	Missing	158,144		4,431		162,575		
Parity								
	0	703,458	45.2	12,559	37.9	716,017	45.1	<0.001
	1	543,004	34.9	10,750	32.4	553,754	34.9	
	>1	309,100	19.9	9,865	29.7	318,965	20.1	
	Missing	8,581		80		8,661		

	Not Excluded N=1,564,143		Excluded N=33,254		Total N=1,597,397		
	n	%	n	%	n	%	P value
Maternity factors							
Mode of delivery							
SVD	1,019,619	65.2	24,462	73.6	1,044,081	65.4	<0.001
Cephalic	37,548	2.4	909	2.7	38,457	2.4	
Assisted VD	187,057	12.0	2,877	8.7	189,934	11.9	
Breech Delivery	4,513	0.3	110	0.3	4,623	0.3	
Elective Caesarean	116,542	7.5	1,751	5.3	118,293	7.4	
Emergency Caesarean	198,590	12.7	3,138	9.4	201,728	12.6	
Other Delivery	272	0.0	7	0.0	279	0.0	
Missing	2		0		2		
Gestation (weeks)							
<24	49	0.0	2	0.0	51	0.0	<0.001
24-27	2,308	0.1	29	0.1	2,337	0.1	
28-32	14,477	0.9	314	0.9	14,791	0.9	
33-36	72,317	4.6	1,825	5.5	74,142	4.6	
37	76,851	4.9	1,962	5.9	78,813	4.9	
38	197,961	12.7	4,489	13.5	202,450	12.7	
39	320,643	20.5	6,696	20.1	327,339	20.5	
40	470,625	30.1	10,105	30.4	480,730	30.1	
41	351,773	22.5	6,738	20.3	358,511	22.5	
42	54,851	3.5	1,049	3.2	55,900	3.5	
43	1,155	0.1	17	0.1	1,172	0.1	
>43	279	0.0	6	0.0	285	0.0	
Missing	854		22		876		
Sex-gestation-specific birthweight centile							
1-3	64,631	4.1	1,797	5.4	66,428	4.2	<0.001
4-10	140,735	9.0	3,469	10.4	144,204	9.0	
11-20	186,706	11.9	4,382	13.2	191,088	12.0	
21-80	918,234	58.8	19,056	57.4	937,290	58.7	
81-90	133,093	8.5	2,457	7.4	135,550	8.5	
91-97	84,479	5.4	1,495	4.5	85,974	5.4	
98-100	34,715	2.2	563	1.7	35,278	2.2	
Missing	1,550		35		1,585		
5 min Apgar							
1-3	7,168	0.5	207	0.6	7,375	0.5	0.001
4-6	14,598	0.9	324	1.0	14,922	0.9	
7-10	1,527,265	98.6	32,551	98.4	1,559,816	98.6	
Missing	15,112		172		15,284		

5.4.2 Association between diabetes and number of exclusions

The diabetes analysis dataset contained 1,596,954 pupil census records after excluding 443 where ineligible diabetes drugs had been prescribed (section 4.3.2.2.1). Diabetes was not significantly associated with number of exclusions from school on univariate analysis (IRR 0.97, 95% CI 0.78-1.22) or after adjusting for sociodemographic (IRR 0.78, 95% CI 0.62-0.97) or maternity (IRR 0.89, 95% CI 0.71-1.11) confounders.

5.4.3 Association between asthma and number of exclusions

The asthma analysis dataset contained 1,415,923 pupil census records after excluding 181,474 where ineligible asthma drugs had been prescribed (section 4.3.2.2.2). Asthma was significantly associated with number of exclusions at school on univariate analysis (IRR 1.16, 95% CI 1.08-1.24) and after adjusting for sociodemographic (IRR 0.90, 95% CI 0.84-0.96) and maternity (IRR 0.91, 95% CI 0.85-0.97) confounders. Asthma seemed to have a protective effect whereby asthmatic pupils were 9% less likely to be excluded from school compared to children without asthma after adjusting for sociodemographic and maternity confounders. There was no significant univariate interaction between asthma and either gender (p=0.510) or age category (p=0.317); however, there was a significant univariate interaction between asthma and deprivation (p=0.012). Only children in the bottom two deprivation categories experienced a significant association; however, the interaction of asthma with deprivation disappeared after further adjustment for sociodemographic and maternity confounders (p=0.056). It was unclear why children with asthma had fewer exclusions compared to healthy peers even after adjusting for sociodemographic and maternal confounders; however, on subgroup analyses, the findings were only significant within boys and those aged between 11 and 14 years and within the most deprived 40% of the population. Unaffected children within these groups already had the highest rates of exclusion; therefore it is possible that small differences, large sample size and residual confounding played parts in these findings. Table 36 shows incidence rate ratios and confidence intervals from univariate and multivariate analyses; overall and within sex, age and deprivation subgroups. The fully adjusted multivariate incidence rate ratios and confidence intervals are summarised visually in Figure 13.

Table 36 Association between asthma and exclusion from school by sex, age and area deprivation

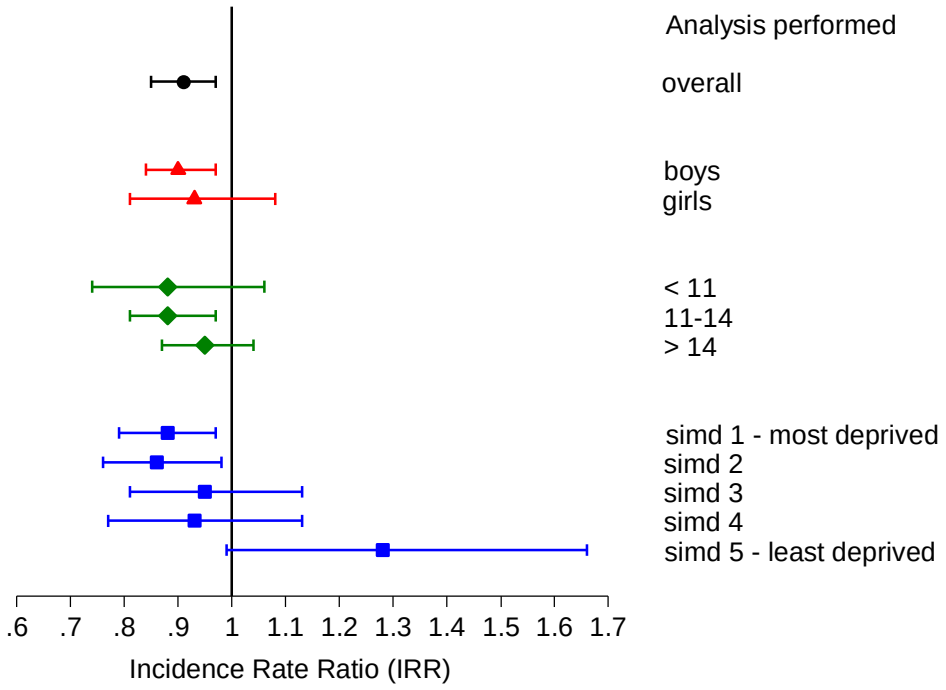
	Univariate model		Multivariate[1]		Multivariate[2]	
	IRR	95% CI	IRR	95% CI	IRR	95% CI
Overall	1.16***	1.08,1.24	0.90**	0.84,0.96	0.91**	0.85,0.97
Boys	1.04	0.96,1.12	0.90**	0.84,0.97	0.90**	0.84,0.97
Girls	1.10	0.95,1.27	0.89	0.77,1.02	0.94	0.82,1.09
< 11 years	1.18	0.99,1.42	0.91	0.76,1.09	0.87	0.72,1.05
11 - 14 years	1.03	0.94,1.13	0.85***	0.78,0.94	0.89**	0.81,0.97
> 14 years	1.10*	1.00,1.21	0.96	0.87,1.05	0.95	0.87,1.05
1 - most deprived	0.98	0.89,1.09	0.84***	0.76,0.93	0.87**	0.79,0.96
2	1.05	0.92,1.20	0.90	0.78,1.03	0.87*	0.77,0.99
3	1.13	0.95,1.34	0.96	0.81,1.14	0.96	0.81,1.13
4	1.12	0.92,1.36	0.93	0.77,1.13	0.93	0.76,1.13
5 - least deprived	1.59***	1.23,2.05	1.32*	1.02,1.70	1.29	1.00,1.67

Exponentiated coefficients; 95% confidence intervals in brackets
* p<0.05, ** p<0.01, *** p<0.001

[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)

[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

Figure 13 Forest plot of the fully adjusted multivariate association between asthma and exclusion by sex, age and area deprivation



5.4.4 Association between epilepsy and number of exclusions

The epilepsy analysis dataset contained 1,597,397 records. Epilepsy was not significantly associated with number of exclusions from school on univariate analysis (IRR 1.19, 95% CI 0.98-1.44) or after adjusting for sociodemographic (IRR 0.90, 95% CI 0.74-1.09) and maternity (IRR 0.89, 95% CI 0.73-1.09) confounders.

5.4.5 Association between attention deficit hyperactivity disorder and number of exclusions

The ADHD analysis dataset contained 1,597,397 records. ADHD was significantly associated with number of exclusions from school on univariate analysis (IRR 10.30, 95% CI 9.73-10.90) and after adjusting for sociodemographic (IRR 7.01, 95% CI 6.61-7.43) and maternity (IRR 5.82, 95% CI 5.47-6.18) confounders. Pupils with ADHD had a 5.82 times greater risk of exclusion compared to pupils who did not have ADHD. There were significant univariate and multivariate interactions between ADHD and gender ($p<0.001$), age category ($p<0.001$) and deprivation category ($p<0.001$). On fully adjusted subgroup analyses the association between ADHD and exclusion from school was stronger in girls than boys (IRR 9.52 versus 5.42). Girls with ADHD experienced a 9.52 times greater risk of exclusion compared to girls without ADHD whereas the corresponding increased risk for boys with ADHD was 5.42. There were trends whereby the multivariate association between ADHD and exclusion was stronger for younger and less deprived children. Within the >14 age group, children with ADHD had a 3.88 times greater risk of exclusion compared to children without ADHD. However, the increased risk of exclusion for children with ADHD was 5 fold and 11 fold within the 11-14 year and <11 year age groups respectively. The risk of exclusion steadily increased with each decreasing category of deprivation. Within the least deprived 20% of the population (category 5), children with ADHD experienced a 10.49 times greater risk of exclusion compared to children without ADHD and this was significantly higher than the corresponding increased risk found within any of the categories representing the most deprived 60% of the population (categories 1-3). The relatively smaller impact of ADHD on risk of exclusion from school for boys and pupils from more socioeconomically deprived areas can be explained because, among unaffected children, they had higher absolute rates of exclusion. Among unaffected children, 10,569 (6.7%) vs. 1,578 (1.2%) in the most and least deprived quintiles, respectively, were excluded. Among children with ADHD, 780 (33.3%) vs. 124 (17.4%) in the most and least deprived quintiles, respectively, were excluded.

The greater relative impact of ADHD on exclusion for younger pupils may be due to teachers being less aware of the condition early on. They may more frequently exclude younger children with ADHD because they label them as bad pupils rather than pupils with a chronic condition. As these pupils pass through school, their condition may become more apparent to school services resulting in less exclusion. Younger children with ADHD may also possibly have less truancy or school refusal and therefore be in school more than older students with ADHD and have more opportunity for exclusion. Table 37 presents incidence rate ratios and confidence intervals from univariate and multivariate analyses; overall and within sex, age and deprivation subgroups. The fully adjusted multivariate incidence rate ratios and confidence intervals are summarised visually in Figure 14.

Table 37 Association between attention deficit hyperactivity disorder and exclusion from school by sex, age and area deprivation

	Univariate model		Multivariate[1]		Multivariate[2]	
	IRR	95% CI	IRR	95% CI	IRR	95% CI
Overall	10.30***	9.73,10.90	7.01***	6.61,7.43	5.82***	5.47,6.18
Boys	7.20***	6.77,7.65	6.47***	6.09,6.89	5.42***	5.08,5.78
Girls	14.49***	12.22,17.18	11.99***	10.12,14.19	9.52***	7.97,11.3
< 11 years	23.97***	21.41,26.84	13.08***	11.64,14.69	10.84***	9.61,12.23
11 - 14 years	8.76***	8.11,9.46	5.91***	5.45,6.40	5.02***	4.62,5.46
> 14 years	7.23***	6.57,7.95	4.80***	4.35,5.28	3.88***	3.51,4.29
1 - most deprived	6.66***	6.10,7.28	5.28***	4.82,5.79	4.78***	4.34,5.27
2	8.90***	8.00,9.90	6.51***	5.85,7.25	5.33***	4.78,5.96
3	11.52***	10.03,13.22	9.10***	7.89,10.50	6.82***	5.87,7.92
4	13.67***	11.66,16.03	10.69***	9.07,12.60	7.51***	6.31,8.95
5 - least deprived	23.09***	18.41,28.96	16.56***	13.15,20.86	10.49***	8.24,13.36

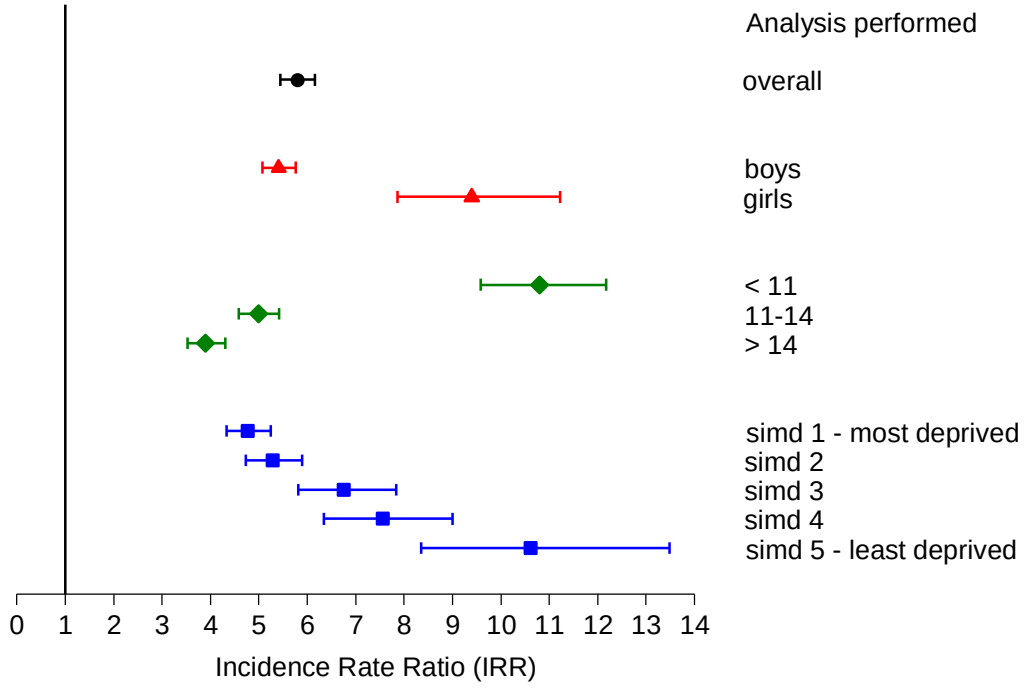
Exponentiated coefficients; 95% confidence intervals in brackets

* p<0.05, ** p<0.01, *** p<0.001

[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)

[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

Figure 14 Forest plot of the fully adjusted multivariate association between attention deficit hyperactivity disorder and exclusion from school by sex, age and area deprivation



5.4.6 Association between depression and number of exclusions

The depression analysis dataset contained 1,597,379 pupil census records after excluding 18 where ineligible anti-depressant drugs had been prescribed (section 4.3.2.2.5). Depression was significantly associated with number of exclusions from school on univariate analysis (IRR 1.86, 95% CI 1.63-2.13) and after adjusting for sociodemographic (IRR 1.63, 95% CI 1.43-1.86) and maternity (IRR 1.65, 95% CI 1.45-1.88) confounders. Pupils on anti-depressant medication had a 1.65 fold risk of exclusion (65% increased risk) compared to pupils who were not on anti-depressant medication. There were no significant univariate interactions with deprivation (p=0.068) or gender (p=0.776). However there were significant univariate and multivariate interactions with age (p<0.001 and p=0.002 respectively). On multivariate subgroup analyses, the association between depression and exclusion was stronger for children under 11 years of age. Depressed children in this age category experienced a 3.5 times greater risk of exclusion compared to unaffected peers and this was significantly higher than the corresponding increased risk observed within older children. For unaffected pupils, children younger than 11 years of age had fewer exclusions compared to older children which explains the greater relative impact of depression on risk of exclusion. Table 38 presents incidence rate ratios and confidence intervals from univariate and multivariate analyses; overall and by age category. The fully

adjusted multivariate incidence rate ratios and confidence intervals are summarised visually in Figure 15.

Table 38 Association between depression and exclusion from school by age

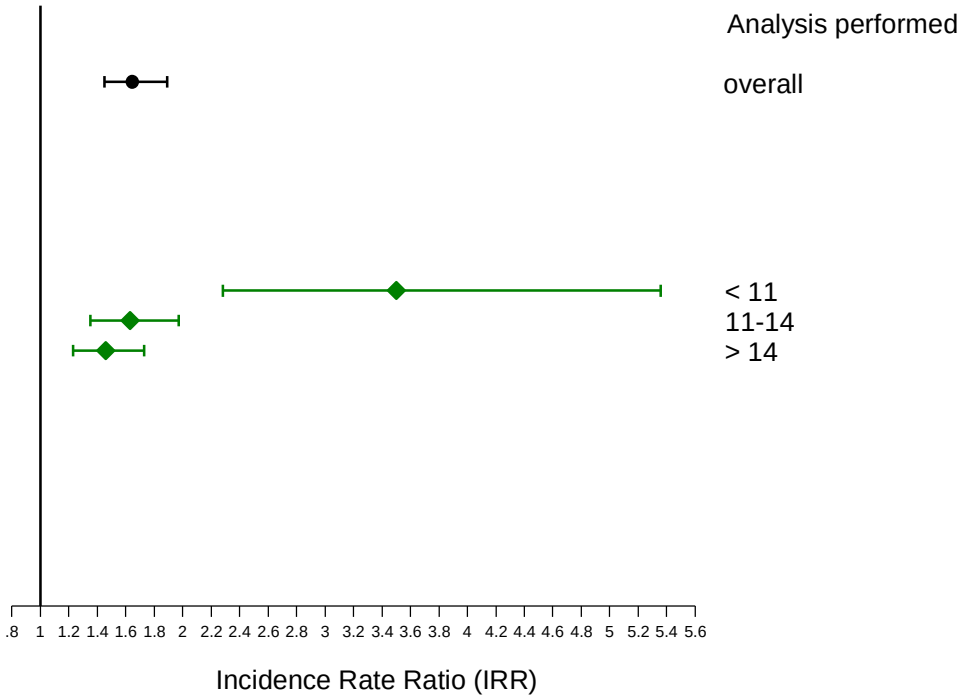
	Univariate model		Multivariate[1]		Multivariate[2]	
	IRR	95% CI	IRR	95% CI	IRR	95% CI
Overall	1.86***	1.63,2.13	1.63***	1.43,1.86	1.65***	1.45,1.88
< 11 years	3.60***	2.33,5.56	3.64***	2.37,5.58	3.50***	2.29,5.37
11 - 14 years	1.25*	1.04,1.52	1.61***	1.33,1.95	1.63***	1.35,1.98
> 14 years	1.07	0.90,1.28	1.42***	1.20,1.70	1.44***	1.21,1.72

Exponentiated coefficients; 95% confidence intervals in brackets
* p<0.05, ** p<0.01, *** p<0.001

[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)

[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

Figure 15 Forest plot of the fully adjusted multivariate association between depression and exclusion from school by age



5.5 Special educational need

5.5.1 Characteristics of children by record of special educational need

The final SEN dataset contained 2,793,185 pupil records across census years 2009 through to 2013. Diabetes, asthma, epilepsy, ADHD and depression were all associated with increased SEN (Table 39). Boys were significantly more likely than girls to have a SEN. Children with a SEN were significantly more deprived and had significantly lower Apgar scores, sex-gestation-specific birthweight centiles and gestational ages. Their mothers were significantly more likely to have been younger than 25 years of age, multiparous and to have smoked during pregnancy. Children with SEN were more likely to be aged 11 to 14 years, more likely to be white and less likely to be of Asian descent and more likely to have been delivered by emergency caesarean section (Table 40).

Table 39 Special educational need by presence or absence of chronic conditions

		No SEN N=2,489,946		**SEN** N=303,239		**Total** N=2,793,185		
		n	%	n	%	n	%	P value
Diabetes								
	No	2,481,704	99.7	300,877	99.2	2,782,581	99.6	<0.001
	Yes	8,242	0.3	2,362	0.8	10,604	0.4	
	Missing	0		0		0		
Asthma								
	No	2,363,085	94.9	281,373	92.8	2,644,458	94.7	<0.001
	Yes	126,861	5.1	21,866	7.2	148,727	5.3	
	Missing	0		0		0		
Epilepsy								
	No	2,482,891	99.7	294,999	97.3	2,777,890	99.5	<0.001
	Yes	7,055	0.3	8,240	2.7	15,295	0.5	
	Missing	0		0		0		
ADHD								
	No	2,477,512	99.5	285,505	94.2	2,763,017	98.9	<0.001
	Yes	12,434	0.5	17,734	5.8	30,168	1.1	
	Missing	0		0		0		
Depression								
	No	2,473,562	99.3	299,295	98.7	2,772,857	99.3	<0.001
	Yes	16,384	0.7	3,944	1.3	20,328	0.7	
	Missing	0		0		0		

Table 40 Characteristics of children by whether or not they had a special educational need

		No SEN		SEN		Total		
		N=2,489,946		N=303,239		N=2,793,185		
		n	%	n	%	n	%	P value
Sociodemographic factors								
Sex								
	male	1,218,591	48.9	203,206	67.0	1,421,797	50.9	<0.001
	female	1,271,355	51.1	100,033	33.0	1,371,388	49.1	
	Missing	0		0		0		
Age category								
	<11	1,259,384	50.6	139,012	45.8	1,398,396	50.1	<0.001
	11-14	588,486	23.6	86,006	28.4	674,492	24.1	
	>14	642,076	25.8	78,221	25.8	720,297	25.8	
	Missing	0		0		0		
Deprivation quintile								
	1 (most deprived)	537,546	21.6	88,611	29.3	626,157	22.4	<0.001
	2	492,494	19.8	66,043	21.8	558,537	20.0	
	3	482,666	19.4	58,034	19.2	540,700	19.4	
	4	495,713	19.9	51,891	17.1	547,604	19.6	
	5 (least deprived)	478,208	19.2	38,125	12.6	516,333	18.5	
	Missing	3,319		535		3,854		
Ethnic group								
	White	2,353,839	96.3	290,301	97.3	2,644,140	96.4	<0.001
	Asian	56,702	2.3	4,912	1.6	61,614	2.2	
	Black	4,827	0.2	546	0.2	5,373	0.2	
	Mixed	21,887	0.9	2,060	0.7	23,947	0.9	
	Other	7,124	0.3	655	0.2	7,779	0.3	
	Missing	45,567		4,765		50,332		
Maternity factors								
Maternal age (years)								
	<=24	648,462	26.0	103,142	34.0	751,604	26.9	<0.001
	25-29	721,552	29.0	82,869	27.3	804,421	28.8	
	30-34	738,258	29.7	74,390	24.5	812,648	29.1	
	>=35	381,628	15.3	42,831	14.1	424,459	15.2	
	Missing	46		7		53		
Maternal smoking								
	No	1,648,602	73.5	168,336	61.4	1,816,938	72.2	<0.001
	Yes	593,403	26.5	106,014	38.6	699,417	27.8	
	Missing	247,941		28,889		276,830		
Parity								
	0	1,131,728	45.7	123,674	41.0	1,255,402	45.2	<0.001
	1	864,758	34.9	100,516	33.3	965,274	34.8	
	>1	479,108	19.4	77,219	25.6	556,327	20.0	
	Missing	14,352		1,830		16,182		

		No SEN		SEN		Total		
		N=2,489,946		N=303,239		N=2,793,185		
		n	%	n	%	n	%	P value
Maternity factors								
Mode of delivery								
	SVD	1,612,033	64.7	197,454	65.1	1,809,487	64.8	0.003
	Cephalic	61,194	2.5	8,208	2.7	69,402	2.5	
	Assisted VD	302,126	12.1	31,795	10.5	333,921	12.0	
	Breech Delivery	6,766	0.3	1,211	0.4	7,977	0.3	
	Elective Caesarean	188,581	7.6	22,572	7.4	211,153	7.6	
	Emergency Caesarean	318,766	12.8	41,955	13.8	360,721	12.9	
	Other Delivery	475	0.0	44	0.0	519	0.0	
	Missing	5		0		5		
Gestation (weeks)								
	<24	82	0.0	18	0.0	100	0.0	<0.001
	24-27	2,476	0.1	1,658	0.5	4,134	0.1	
	28-32	20,398	0.8	5,610	1.9	26,008	0.9	
	33-36	110,836	4.5	19,359	6.4	130,195	4.7	
	37	119,761	4.8	17,836	5.9	137,597	4.9	
	38	310,865	12.5	42,041	13.9	352,906	12.6	
	39	514,884	20.7	61,851	20.4	576,735	20.7	
	40	752,191	30.2	84,198	27.8	836,389	30.0	
	41	568,919	22.9	60,945	20.1	629,864	22.6	
	42	86,095	3.5	9,303	3.1	95,398	3.4	
	43	1,816	0.1	189	0.1	2,005	0.1	
	>43	424	0.0	54	0.0	478	0.0	
	Missing	1,199		177		1,376		
Sex-gestation-specific birthweight centile								
	1-3	97,870	3.9	17,430	5.8	115,300	4.1	<0.001
	4-10	219,720	8.8	31,136	10.3	250,856	9.0	
	11-20	295,663	11.9	37,444	12.4	333,107	11.9	
	21-80	1,470,084	59.1	169,775	56.1	1,639,859	58.8	
	81-90	213,735	8.6	24,341	8.0	238,076	8.5	
	91-97	135,079	5.4	15,804	5.2	150,883	5.4	
	98-100	55,498	2.2	6,920	2.3	62,418	2.2	
	Missing	2,297		389		2,686		
5 min Apgar								
	1-3	10,466	0.4	2,157	0.7	12,623	0.5	<0.001
	4-6	21,887	0.9	4,191	1.4	26,078	0.9	
	7-10	2,431,875	98.7	293,096	97.9	2,724,971	98.6	
	Missing	25,718		3,795		29,513		

5.5.2 Association between diabetes and special educational need

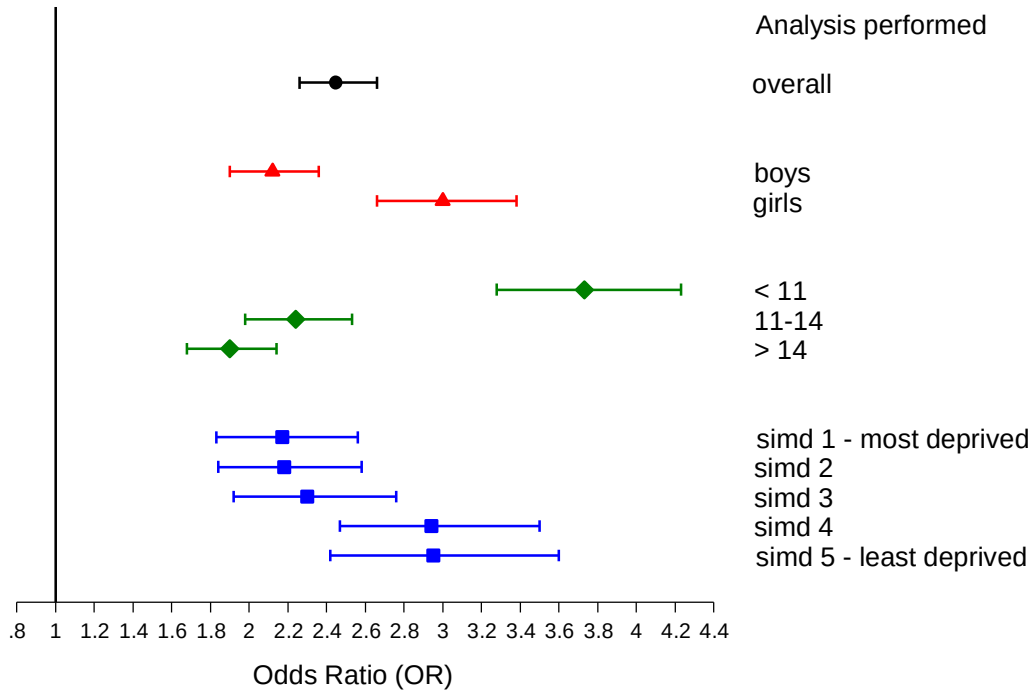
The diabetes analysis dataset contained 2,792,469 pupil census records after excluding 716 where ineligible diabetes drugs had been prescribed (section 4.3.2.2.1). Diabetes was significantly associated with SEN on univariate analysis (OR 2.36, 95% CI 2.19-2.55) and after adjusting for sociodemographic (OR 2.36, 95% CI 2.18-2.55) and maternity (OR 2.45, 95% CI 2.26-2.66) confounders. Pupils with diabetes had a 2.45 times greater odds of SEN compared to pupils who did not have diabetes. There were significant univariate and multivariate interactions between diabetes and gender (both $p<0.001$), age (both $p<0.001$) and deprivation (univariate $p=0.013$, multivariate $p=0.043$). On fully adjusted subgroup analyses the association between diabetes and SEN was stronger for girls than boys (OR 3.00 versus 2.12). Girls with diabetes experienced three-fold odds of SEN compared to girls without diabetes whereas the corresponding increase in odds was only two-fold for boys. There was a visible trend whereby the association between diabetes and SEN was stronger within younger age categories. For children younger than 11 years of age, those with diabetes experienced a 3.73 times greater odds of SEN compared to children without diabetes and this was significantly higher than the corresponding increased odds observed within the older age groups. The odds of SEN was higher in the least deprived 40% of the population compared to the most deprived 40% after adjusting for sociodemographic and maternity confounders. Among non-diabetic children, 32,045 (19.7%) in the most deprived quintile had SEN compared with 14,548 (10.4%) in the least-deprived quintile. Among diabetic children, 235 (35.1%) in the most deprived quintile had SEN compared with 165 (26.7%) in the least deprived quintile. The relative impact of diabetes on the odds of SEN was greater for girls and children younger than 11 years of age because, among unaffected children, they had lower absolute rates of SEN than boys and older pupils. Table 41 presents odds ratios and confidence intervals from univariate and multivariate analyses; overall and within sex, age and deprivation subgroups. The fully adjusted multivariate odds ratios and confidence intervals are summarised visually in Figure 16.

Table 41 Association between diabetes and special educational need by sex, age and area deprivation

	Univariate model		Multivariate[1]		Multivariate[2]	
	OR	95% CI	OR	95% CI	OR	95% CI
Overall	2.36***	2.19-2.55	2.36***	2.18-2.55	2.45***	2.26-2.66
Boys	2.10***	1.89-2.33	2.06***	1.85-2.29	2.12***	1.90-2.36
Girls	2.94***	2.62-3.29	2.85***	2.54-3.20	3.00***	2.66-3.38
< 11 years	3.42***	3.04-3.84	3.58***	3.18-4.05	3.73***	3.28-4.23
11 - 14 years	2.06***	1.83-2.31	2.13***	1.89-2.40	2.24***	1.98-2.53
> 14 years	1.82***	1.62-2.04	1.83***	1.62-2.06	1.90***	1.68-2.14
1 - most deprived	2.07***	1.76-2.42	2.03***	1.72-2.39	2.17***	1.83-2.56
2	2.19***	1.87-2.56	2.12***	1.80-2.50	2.18***	1.84-2.58
3	2.31***	1.95-2.75	2.28***	1.91-2.72	2.30***	1.92-2.76
4	2.80***	2.37-3.32	2.80***	2.36-3.33	2.94***	2.47-3.50
5 - least deprived	2.93***	2.44-3.53	2.88***	2.38-3.48	2.95***	2.42-3.60

Exponentiated coefficients; 95% confidence intervals in brackets
* p<0.05, ** p<0.01, *** p<0.001
[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)
[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

Figure 16 Forest plot of the fully adjusted multivariate association between diabetes and special educational need by sex, age and area deprivation



5.5.3 Association between asthma and special educational need

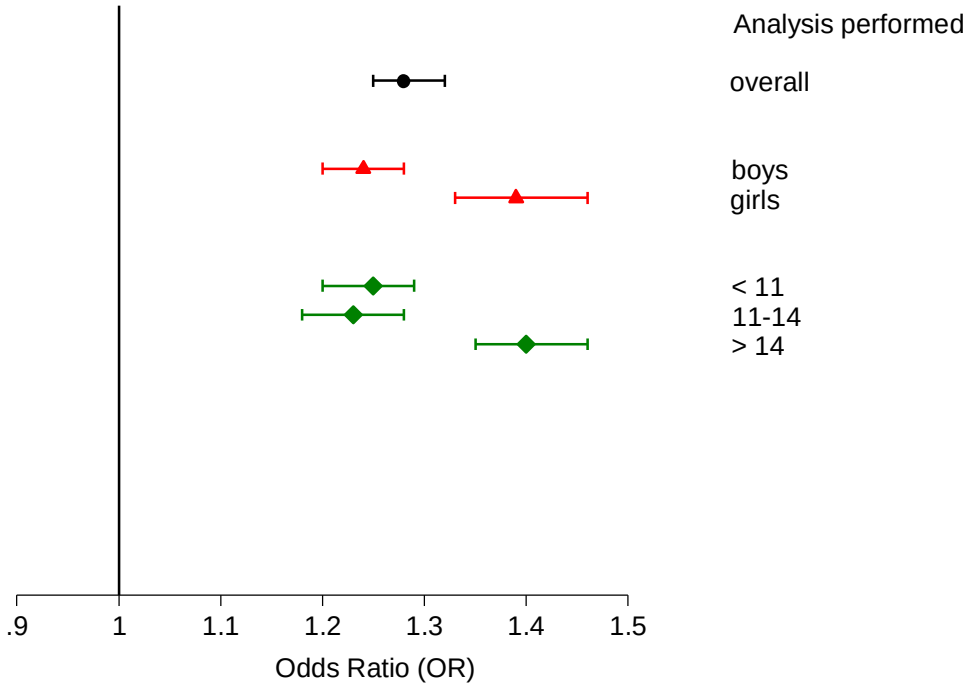
The asthma analysis dataset contained 2,472,798 pupil census records after excluding 320,387 where ineligible asthma drugs had been prescribed (section 4.3.2.2.2). Asthma was significantly associated with SEN at school on univariate analysis (OR 1.47, 95% CI 1.43-1.50) and after adjusting for sociodemographic (OR 1.31, 95% CI 1.28-1.35) and maternity (OR 1.28, 95% CI 1.25-1.32) confounders. Pupils with asthma had 1.28 times greater odds of SEN (28% increased odds) compared to pupils who did not have asthma. There was no significant univariate interaction between asthma and deprivation (p=0.097); however, there were significant univariate and multivariate interactions between asthma and sex (both p<0.001) and age (both p<0.001). On fully adjusted subgroup analyses the association between asthma and SEN was stronger for girls than boys (OR 1.39 versus 1.24). Girls with asthma experienced a 39% increase in odds of SEN compared to girls without asthma. The corresponding increase in odds was only 24% for boys. Over 14 years of age children with asthma experienced a 40% increase in the odds of SEN compared to children without asthma. This was significantly higher than the corresponding increased odds observed within younger age groups. The relative impact of asthma on the odds of SEN was greater for girls and children older than 14 years of age because, among unaffected children, they had lower absolute rates of SEN compared to boys and younger children. Table 42 presents odds ratios and confidence intervals from univariate and multivariate analyses; overall and within sex and age subgroups. The fully adjusted multivariate odds ratios and confidence intervals are summarised visually in Figure 17.

Table 42 Association between asthma and special educational need by sex and age

	Univariate model		Multivariate[1]		Multivariate[2]	
	OR	95% CI	OR	95% CI	OR	95% CI
Overall	1.47***	1.43-1.50	1.31***	1.28-1.35	1.28***	1.25-1.32
Boys	1.33***	1.29-1.37	1.27***	1.23-1.31	1.24***	1.20-1.28
Girls	1.50***	1.43-1.57	1.42***	1.36-1.49	1.39***	1.33-1.46
< 11 years	1.44***	1.39-1.49	1.29***	1.25-1.34	1.25***	1.20-1.29
11 - 14 years	1.38***	1.33-1.43	1.25***	1.20-1.30	1.23***	1.18-1.28
> 14 years	1.54***	1.48-1.60	1.43***	1.37-1.49	1.40***	1.35-1.46

Exponentiated coefficients; 95% confidence intervals in brackets
* p<0.05, ** p<0.01, *** p<0.001
[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)
[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

Figure 17 Forest plot of the fully adjusted multivariate association between asthma and special educational need by sex and age



5.5.4 Association between epilepsy and special educational need

The epilepsy analysis dataset contained 2,793,185 records. Epilepsy was significantly associated with SEN at school on univariate analysis (OR 9.83, 95% CI 9.29-10.40) and after adjusting for sociodemographic (OR 10.25, 95% CI 9.65-10.88) and maternity (OR 10.11, 95% CI 9.51-10.75) confounders. Pupils with epilepsy had a ten-fold odds of SEN compared to pupils who did not have epilepsy. There were significant univariate and multivariate interactions between epilepsy and gender (both $p<0.001$), age (both $p<0.001$) and deprivation (both $p<0.001$). On fully adjusted subgroup analyses the association between epilepsy and SEN was stronger for girls than boys (OR 11.68 versus 8.76). Girls with epilepsy experienced 11.68 times greater odds of SEN compared to girls without epilepsy. The corresponding increase in odds was only 8.76 fold for boys. The association between epilepsy and SEN was stronger within younger and less deprived groups of children. Children younger than 11 years of age with epilepsy experienced 13.5 times greater odds of SEN compared to children in the same age group without epilepsy. This was significantly higher than the corresponding increased odds observed within older age groups. Similarly, within the least deprived 20% of the population, children with epilepsy

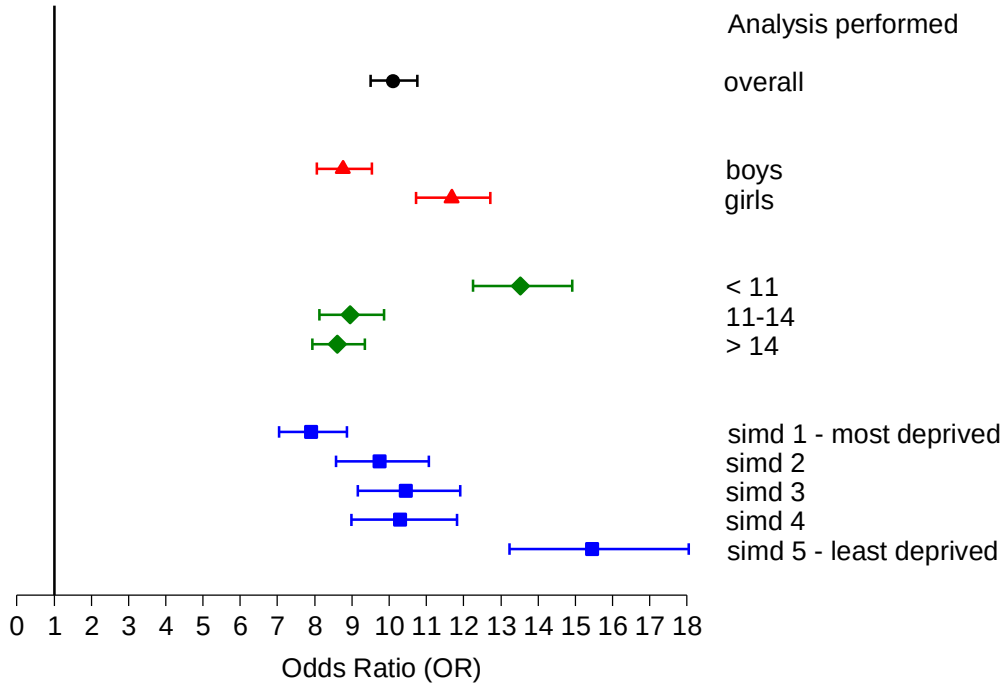
had 15.45 times greater odds of SEN compared to children without epilepsy. This was significantly higher than the corresponding odds within the more deprived subgroups. The relative impact of epilepsy on the odds of SEN was greater for girls, children younger than 11 years of age and children in the least deprived 20% of the population. This can be explained because, among unaffected pupils, these groups of children had lower absolute rates of SEN compared to boys, more deprived children and pupils aged between 11 and 14 years of age. Among children without epilepsy, 33,475 (19.4%) in the most deprived quintile had SEN compared with 14,325 (10.2%) in the least-deprived quintile. Among children with epilepsy, 827 (60.3%) in the most deprived quintile had SEN compared with 395 (49.9%) in the least deprived quintile. Epilepsy can be more severe in younger children [271] and this may also explain the greater relative impact of epilepsy on SEN in children younger than 11 years of age. Table 43 presents odds ratios and confidence intervals from univariate and multivariate analyses; overall and within sex, age and deprivation subgroups. The fully adjusted multivariate odds ratios and confidence intervals are summarised visually in Figure 18.

Table 43 Association between epilepsy and special educational need by sex, age and area deprivation

	Univariate model		Multivariate[1]		Multivariate[2]	
	OR	95% CI	OR	95% CI	OR	95% CI
Overall	9.83***	9.29-10.40	10.25***	9.65-10.88	10.11***	9.51-10.75
Boys	9.10***	8.40-9.86	8.81***	8.11-9.56	8.76***	8.05-9.53
Girls	11.99***	11.07-12.99	11.92***	10.98-12.95	11.68***	10.73-12.71
< 11 years	13.16***	12.0-14.3	13.71***	12.46-15.08	13.52***	12.25-14.91
11 - 14 years	8.44***	7.71-9.23	8.91***	8.11-9.80	8.96***	8.13-9.87
> 14 years	8.20***	7.60-8.86	8.75***	8.07-9.47	8.61***	7.94-9.34
1 - most deprived	7.90***	7.10-8.79	7.88***	7.05-8.81	7.91***	7.05-8.87
2	9.45***	8.38-10.65	9.94***	8.77-11.28	9.74***	8.57-11.06
3	9.90***	8.75-11.21	10.64***	9.34-12.11	10.44***	9.15-11.91
4	9.90***	8.70-11.27	10.73***	9.39-12.26	10.30***	8.98-11.83
5 - least deprived	14.33***	12.39-16.57	15.30***	13.16-17.80	15.45***	13.23-18.05

Exponentiated coefficients; 95% confidence intervals in brackets
* p<0.05, ** p<0.01, *** p<0.001
[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)
[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

Figure 18 Forest plot of the fully adjusted multivariate association between epilepsy and special educational need by sex, age and area deprivation



5.5.5 Association between attention deficit hyperactivity disorder and special educational need

The ADHD analysis dataset contained 2,793,185 records. ADHD was significantly associated with SEN at school on univariate analysis (OR 12.38, 95% CI 11.91-12.86) and after adjusting for sociodemographic (OR 9.63, 95% CI 9.25-10.00) and maternity (OR 8.77, 95% CI 8.40-9.15) confounders. Pupils with ADHD had 8.77 times greater odds of SEN compared to pupils who did not have ADHD. There were significant univariate and multivariate interactions between ADHD and gender (both $p<0.001$), age (both $p<0.001$) and deprivation (both $p<0.001$). On fully adjusted subgroup analyses the association between ADHD and SEN was stronger for girls than boys (OR 12.85 versus 8.15). Girls with ADHD experienced 12.85 times greater odds of SEN compared to girls without ADHD. The corresponding increase in odds was only 8.15 fold for boys. There were visible trends whereby the association between ADHD and SEN was stronger for older and less deprived pupils. In children over 14 years of age, those with ADHD had 10.85 times greater odds of SEN compared to children without ADHD. The odds ratio decreased to 8.80 within the 11-14 years group and again to 7.82 within children younger than 11 years of age. Within the most deprived 20% of the population, children with ADHD had 7.17 times greater odds of SEN compared to children without ADHD which was significantly

lower than the corresponding increased odds found within the other deprivation subgroups. For unaffected children, girls and pupils from less deprived backgrounds had lower absolute rates of SEN which explains the greater relative impact of ADHD on SEN within these subgroups. Among children without ADHD, 32,461 (18.9%) in the most deprived quintile had SEN compared with 14,185 (10.1%) in the least-deprived quintile. Among children with ADHD, 1,841 (76.9%) in the most deprived quintile had SEN compared with 535 (70.8%) in the least deprived quintile. The relative impact of ADHD on SEN was greater in older children because, for those unaffected by ADHD, older children had lower absolute rates of SEN compared to children aged between 11 and 14 years of age. Children start secondary school between 11 and 14 years of age and this transition could theoretically be one reason why absolute rates of SEN are higher in this age group. Table 44 presents odds ratios and confidence intervals from univariate and multivariate analyses; overall and within sex, age and deprivation subgroups. The fully adjusted multivariate odds ratios and confidence intervals are summarised visually in Figure 19.

Table 44 Association between attention deficit hyperactivity disorder and special educational need by sex, age and area deprivation

	Univariate model		Multivariate[1]		Multivariate[2]	
	OR	95% CI	OR	95% CI	OR	95% CI
Overall	12.38***	11.91-12.86	9.63***	9.25-10.0	8.77***	8.40-9.15
Boys	9.52***	9.12-9.93	8.93***	8.56-9.33	8.15***	7.79-8.53
Girls	15.14***	13.73-16.70	14.20***	12.84-15.71	12.85***	11.54-14.29
< 11 years	11.06***	10.51-11.63	8.62***	8.18-9.10	7.82***	7.38-8.28
11 - 14 years	12.25***	11.54-13.00	9.60***	9.03-10.22	8.80***	8.25-9.39
> 14 years	15.08***	14.08-16.15	12.00***	11.18-12.88	10.85***	10.09-11.68
1 - most deprived	9.24***	8.65-9.88	7.49***	7.00-8.01	7.17***	6.68-7.70
2	11.79***	10.92-12.72	9.54***	8.82-10.32	8.70***	8.02-9.45
3	12.23***	11.21-13.36	10.12***	9.25-11.08	8.85***	8.05-9.74
4	14.02***	12.74-15.44	11.52***	10.44-12.72	10.05**	9.05-11.15
5 - least deprived	17.74***	15.73-20.02	14.55***	12.85-16.48	12.39***	10.83-14.18

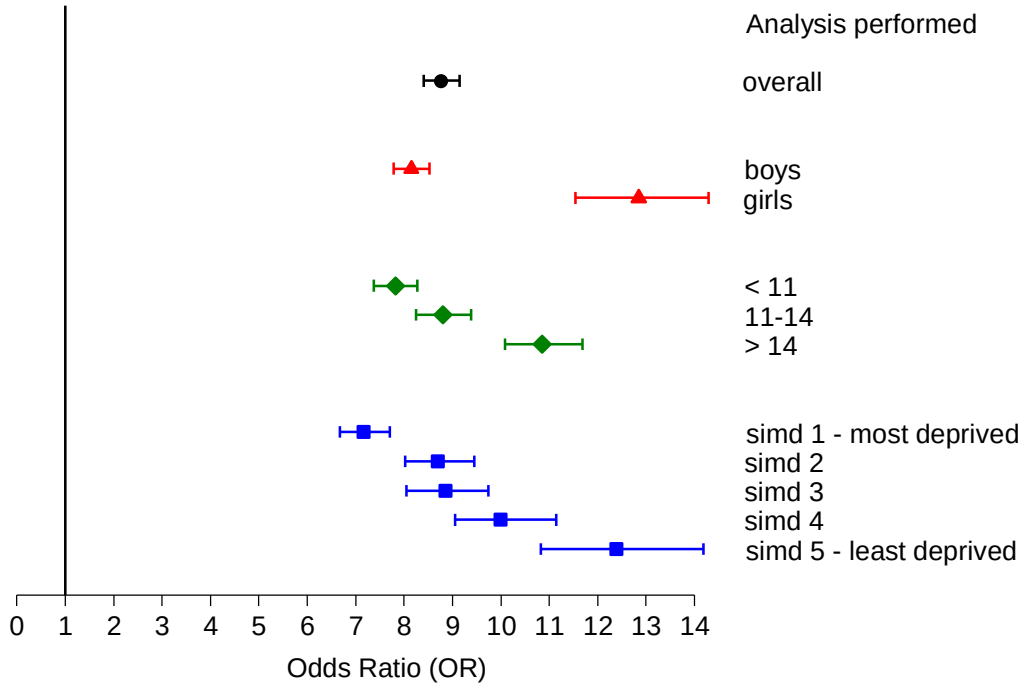
Exponentiated coefficients; 95% confidence intervals in brackets

* p<0.05, ** p<0.01, *** p<0.001

[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)

[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

Figure 19 Forest plot of the fully adjusted multivariate association between attention deficit hyperactivity disorder and special educational need by sex, age and area deprivation



5.5.6 Association between depression and special educational need

The depression analysis dataset contained 2,793,157 pupil census records after excluding 28 where ineligible anti-depressant drugs had been prescribed (section 4.3.2.2.5). Depression was significantly associated with SEN on univariate analysis (OR 1.99, 95% CI 1.87-2.12) and after adjusting for sociodemographic (OR 2.23, 95% CI 2.09-2.38) and maternity (OR 2.24, 95% CI 2.10-2.39) confounders. Pupils with depression had a 2.24 fold odds of SEN compared to pupils who did not have depression. There were significant univariate and multivariate interactions between depression and gender (both $p<0.001$), age (both $p<0.001$) and deprivation (both $p<0.001$). On fully adjusted subgroup analyses the association between depression and SEN was stronger for boys than girls (OR 2.66 versus 1.93). Boys with depression experienced 2.66 times greater odds of SEN compared to boys without depression. The corresponding increase in odds was only 1.93 fold in girls. Among children without depression, 75,413 (19.4%) boys had SEN compared with 39,368 (10.6%) girls. Among depressed children, 709 (40.5%) boys had SEN compared with 754 (21.0%) girls. On fully adjusted subgroup analysis by age category the odds ratio was significantly lower in the middle age group (11-14 years) compared to the other two groups. Children in the 11 to 14 year age group with depression experienced a 78% increase in odds of SEN

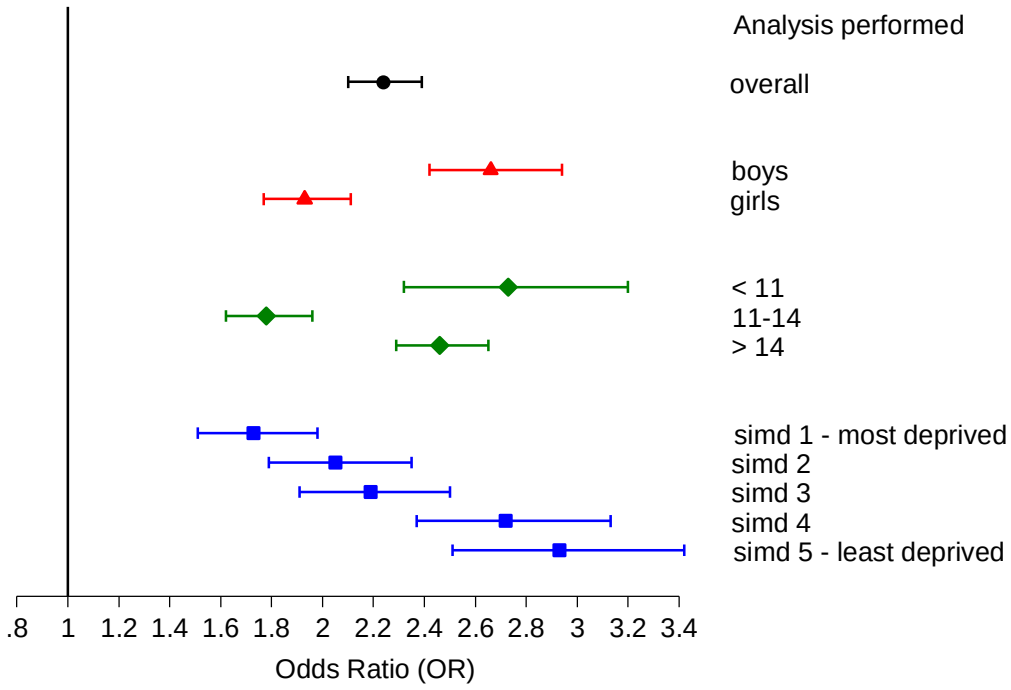
compared to children in the same age group without depression. This was significantly lower than the corresponding odds ratios within the younger and older age groups. There was a visible trend whereby the strength of the association between depression and SEN steadily increased with decreasing deprivation. Within the least deprived 20% of the population, depressed children had 2.93 times greater odds of SEN compared to children without depression. This was significantly higher than the corresponding odds within categories representing the most deprived 60%. The relative impact of depression on odds of SEN was less for more deprived pupils and children aged between 11 and 14 years of age because, among unaffected children, these subgroups of children already experienced higher absolute rates of SEN. Among children without depression, 33,990 (19.7%) in the most deprived quintile had SEN compared with 14,489 (10.4%) in the least-deprived quintile. Among depressed children, 312 (30.7%) in the most deprived quintile had SEN compared with 231 (23.3%) in the least deprived quintile. Depressed boys may present more obvious symptoms than depressed girls perhaps due to presence of more comorbid conditions, increased disease severity or differences in behaviour patterns. Therefore, girls may be less likely to be identified as depressed by their school and receive less help. This may explain the greater relative impact of depression on risk of SEN observed in boys. Table 45 presents odds ratios and confidence intervals from univariate and multivariate analyses; overall and within sex, age and deprivation subgroups. The fully adjusted multivariate odds ratios and confidence intervals are summarised visually in Figure 20.

Table 45 Association between depression and special educational need by sex, age and area deprivation

	Univariate model		Multivariate[1]		Multivariate[2]	
	OR	95% CI	OR	95% CI	OR	95% CI
Overall	1.99***	1.87-2.12	2.23***	2.09-2.38	2.24***	2.10-2.39
Boys	2.71***	2.47-2.97	2.64***	2.40-2.91	2.66***	2.42-2.94
Girls	2.02***	1.86-2.20	1.93***	1.77-2.11	1.93***	1.77-2.11
< 11 years	2.74***	2.36-3.19	2.74***	2.35-3.20	2.73***	2.32-3.20
11 - 14 years	1.52***	1.39-1.67	1.77***	1.61-1.94	1.78***	1.62-1.96
> 14 years	2.00***	1.86-2.15	2.45***	2.28-2.63	2.46***	2.29-2.65
1 - most deprived	1.60***	1.40-1.83	1.69***	1.48-1.93	1.73***	1.51-1.98
2	1.86***	1.63-2.13	2.06***	1.80-2.35	2.05***	1.79-2.35
3	1.94***	1.70-2.21	2.17***	1.90-2.48	2.19***	1.91-2.50
4	2.40***	2.09-2.75	2.76***	2.41-3.16	2.72***	2.37-3.13
5 - least deprived	2.70***	2.32-3.13	2.98***	2.57-3.46	2.93***	2.51-3.42

Exponentiated coefficients; 95% confidence intervals in brackets
* p<0.05, ** p<0.01, *** p<0.001
[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)
[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

Figure 20 Forest plot of the fully adjusted multivariate association between depression and special educational need by sex, age and area deprivation



5.5.7 Association between each chronic condition and specific type of special educational need

Table 46 provides univariate and multivariate odds ratios for analyses of each of the chronic conditions against each of the SEN categories. There was a significant association between diabetes and some but not all types of SEN. Those significantly associated with diabetes included learning difficulty (adjusted OR 1.19, 95% CI 1.03-1.38, $p<0.05$), physical motor disability (adjusted OR 1.55, 95% CI 1.13-2.12, $p<0.01$) and physical health condition (adjusted OR 24.08, 95% CI 21.83-26.57, $p<0.001$). The strongest association occurred between diabetes and SEN attributed to a physical health condition. Children with diabetes experienced 24-fold odds of SEN due to a physical health condition compared to non-diabetic children.

All of the other chronic conditions were significantly associated with all types of SEN. Asthma (adjusted OR 2.76, 95% CI 2.57-2.95, $p<0.001$) and epilepsy (adjusted OR 60.15, 95% CI 55.44-65.26, $p<0.001$) were also most strongly associated with SEN attributed to a physical health condition. The odds of having a SEN at school in that particular category increased almost 3-fold for children with asthma compared to children without asthma and by 60-fold for children with epilepsy compared to children without epilepsy. ADHD (adjusted OR 59.15, 95% CI 51.68-67.71, $p<0.001$) and depression (adjusted OR 17.53, 95% CI 15.05-20.41, $p<0.001$) were most strongly associated with SEN attributed to a mental health condition. The odds of having that particular type of SEN increased almost 59 fold for pupils with ADHD compared to children without ADHD and 17.5 fold for depressed versus non depressed children.

There were two other strong associations of note. Firstly, asthmatic children experienced 76% increased odds (adjusted OR 1.76, 95% CI 1.49-2.08, $p<0.001$) of SEN attributed to a mental health condition compared to non-asthmatic children. Secondly, epileptic children experienced 57.5 times greater odds (adjusted OR 57.51, 95% CI 52.57-62.91, $p<0.001$) of SEN attributed to a physical motor disability compared to children without epilepsy.

Table 46 Association between chronic disease and type of special educational need

	Univariate model		Multivariate[1]		Multivariate[2]	
	OR	95% CI	OR	95% CI	OR	95% CI
learning disability						
Diabetes	1.16	0.91-1.48	1.13	0.89-1.45	1.18	0.92-1.51
Asthma	1.43***	1.35-1.51	1.27***	1.19-1.34	1.21***	1.14-1.29
Epilepsy	28.08***	26.11-30.19	28.46***	26.35-30.74	27.92***	25.78-30.25
ADHD	12.49***	11.64-13.41	9.66***	8.97-10.39	8.35***	7.72-9.02
Depression	1.79***	1.55-2.08	1.97***	1.70-2.29	1.96***	1.69-2.28
learning difficulty						
Diabetes	1.23**	1.06-1.42	1.15	1.00-1.34	1.19*	1.03-1.38
Asthma	1.43***	1.37-1.48	1.29***	1.25-1.34	1.28***	1.23-1.33
Epilepsy	5.45***	5.00-5.94	5.35***	4.90-5.85	5.30***	4.85-5.81
ADHD	7.91***	7.48-8.38	6.25***	5.89-6.62	5.55***	5.22-5.89
Depression	1.25***	1.12-1.40	1.30***	1.16-1.45	1.29***	1.15-1.44
sensory impairment						
Diabetes	1.26	0.83-1.90	1.21	0.80-1.84	1.28	0.84-1.94
Asthma	1.76***	1.59-1.94	1.65***	1.49-1.82	1.49***	1.35-1.65
Epilepsy	33.76***	30.30-37.62	33.01***	29.56-36.87	30.59***	27.26-34.33
ADHD	5.44***	4.54-6.52	4.77***	3.96-5.73	4.01***	3.31-4.86
Depression	1.77***	1.36-2.29	1.78***	1.37-2.32	1.72***	1.32-2.24
physical motor disability						
Diabetes	1.49*	1.09-2.03	1.49*	1.09-2.03	1.55**	1.13-2.12
Asthma	1.54***	1.41-1.68	1.43***	1.31-1.57	1.28***	1.16-1.40
Epilepsy	57.11***	52.61-61.99	60.59***	55.63-66.00	57.51***	52.57-62.91
ADHD	9.40***	8.34-10.59	7.66***	6.78-8.66	6.72***	5.88-7.68
Depression	2.38***	1.94-2.91	2.73***	2.23-3.35	2.66***	2.16-3.28
communication problems						
Diabetes	0.97	0.71-1.32	1.18	0.87-1.60	1.21	0.89-1.64
Asthma	1.33***	1.24-1.42	1.23***	1.15-1.32	1.19***	1.11-1.28
Epilepsy	20.55***	18.81-22.45	25.76***	23.40-28.37	24.98***	22.65-27.55
ADHD	7.50***	6.76-8.31	5.56***	5.00-6.18	5.05***	4.52-5.64
Depression	1.05	0.84-1.32	1.75***	1.39-2.20	1.72***	1.37-2.17
ASD						
Diabetes	1.39*	1.03-1.89	1.32	0.97-1.78	1.30	0.95-1.76
Asthma	1.35***	1.25-1.47	1.15***	1.06-1.25	1.11*	1.03-1.21
Epilepsy	12.13***	10.74-13.69	12.41***	10.92-14.10	12.20***	10.72-13.89
ADHD	21.26***	19.71-22.95	14.18***	13.09-15.35	14.57***	13.42-15.82
Depression	4.29***	3.76-4.91	5.63***	4.91-6.46	5.62***	4.89-6.45
social, emotional or behavioral difficulty						
Diabetes	0.90	0.72-1.12	0.86	0.69-1.07	0.94	0.75-1.18
Asthma	1.42***	1.35-1.48	1.21***	1.15-1.27	1.18***	1.13-1.24
Epilepsy	7.16***	6.50-7.88	7.10***	6.40-7.88	7.14***	6.41-7.96
ADHD	33.06***	31.61-34.57	23.72***	22.60-24.90	20.48***	19.44-21.58
Depression	2.80***	2.55-3.07	3.32***	3.02-3.65	3.34***	3.03-3.69
physical health condition						
Diabetes	24.27***	22.16-26.59	23.19***	21.11-25.48	24.08***	21.83-26.57
Asthma	3.14***	2.94-3.36	2.94***	2.75-3.14	2.76***	2.57-2.95
Epilepsy	63.98***	59.29-69.05	61.54***	56.88-66.58	60.15***	55.44-65.26
ADHD	9.36***	8.33-10.51	8.22***	7.30-9.25	7.41***	6.55-8.39
Depression	4.39***	3.80-5.07	4.16***	3.59-4.81	4.10***	3.54-4.76
mental health condition						
Diabetes	1.14	0.56-2.31	0.93	0.46-1.88	0.99	0.49-2.00
Asthma	2.01***	1.71-2.37	1.79***	1.51-2.11	1.76***	1.49-2.08
Epilepsy	30.61***	25.36-36.94	25.91***	21.37-31.41	25.75***	21.16-31.33
ADHD	73.18***	65.11-82.26	65.54***	57.68-74.47	59.15***	51.68-67.71
Depression	19.88***	17.21-22.97	17.58***	15.12-20.44	17.53***	15.05-20.41

Exponentiated coefficients; 95% confidence intervals in brackets

* p<0.05, ** p<0.01, *** p<0.001

[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)

[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

5.6 Poor academic attainment

5.6.1 Characteristics of children by poor academic attainment

The final attainment dataset contained 139,205 pupil records across census years 2009 through to 2013 with one summary record present per pupil. Asthma, epilepsy and ADHD were significantly associated with poorer academic attainment; however, there was no significant association between diabetes or depression and poorer academic attainment (Table 47). Boys had significantly poorer academic attainment than girls. Children with poorer academic attainment were significantly younger and more deprived and had significantly lower sex-gestation-specific birthweight centiles, gestational ages and 5-minute Apgar scores. Their mothers were significantly younger (specifically <=24), multiparous and were significantly more likely to have smoked during pregnancy. Children with low attainment were significantly more likely to have been born by spontaneous vaginal delivery and significantly less likely to have been delivered by caesarean section or assisted vaginal delivery. They were significantly more likely to be white and significantly less likely to be of Asian descent (Table 48).

Table 47 Poor academic attainment by presence or absence of chronic conditions

		High N=58,493		General N=30,321		Basic N=43,857		Low N=6,534		Total N=139,205		
		n	%	n	%	n	%	n	%	n	%	P value
Diabetes												
	No	58,131	99.4	30,102	99.3	43,598	99.4	6,505	99.6	138,336	99.4	0.212
	Yes	362	0.6	219	0.7	259	0.6	29	0.4	869	0.6	
	Missing	0		0		0		0		0		
Asthma												
	No	55,214	94.4	28,493	94.0	41,159	93.8	6,120	93.7	130,986	94.1	<0.001
	Yes	3,279	5.6	1,828	6.0	2,698	6.2	414	6.3	8,219	5.9	
	Missing	0		0		0		0		0		
Epilepsy												
	No	58,151	99.4	30,102	99.3	43,519	99.2	6,424	98.3	138,196	99.3	<0.001
	Yes	342	0.6	219	0.7	338	0.8	110	1.7	1,009	0.7	
	Missing	0		0		0		0		0		
ADHD												
	No	58,385	99.8	30,134	99.4	43,077	98.2	6,220	95.2	137,816	99.0	<0.001
	Yes	108	0.2	187	0.6	780	1.8	314	4.8	1,389	1.0	
	Missing	0		0		0		0		0		
Depression												
	No	57,549	98.4	29,742	98.1	43,218	98.5	6,356	97.3	136,865	98.3	0.093
	Yes	944	1.6	579	1.9	639	1.5	178	2.7	2,340	1.7	
	Missing	0		0		0		0		0		

Table 48 Characteristics of children by whether or not they had poor academic attainment

		High N=58,493		General N=30,321		Basic N=43,857		Low N=6,534		Total N=139,205		
		n	%	n	%	n	%	n	%	n	%	P value
Sociodemographic factors												
Sex												
	male	25,696	43.9	15,987	52.7	24,886	56.7	3,630	55.6	70,199	50.4	<0.001
	female	32,797	56.1	14,334	47.3	18,971	43.3	2,904	44.4	69,006	49.6	
	Missing	0		0		0		0		0		
Age	14	0	0.0	121	0.4	897	2.0	290	4.4	1,308	0.9	<0.001
	15	1,123	1.9	4,804	15.8	27,389	62.5	5,382	82.4	38,698	27.8	
	16	23,801	40.7	14,135	46.6	12,061	27.5	653	10.0	50,650	36.4	
	17	33,397	57.1	11,140	36.7	3,463	7.9	180	2.8	48,180	34.6	
	18	172	0.3	121	0.4	47	0.1	29	0.4	369	0.3	
	Missing	0		0		0		0		0		
Deprivation quintile												
	1 (most deprived)	6,549	11.2	6,832	22.5	13,935	31.8	2,882	44.2	30,198	21.7	<0.001
	2	9,017	15.4	6,725	22.2	10,827	24.7	1,704	26.2	28,273	20.3	
	3	11,615	19.9	6,325	20.9	8,538	19.5	1,046	16.1	27,524	19.8	
	4	14,325	24.5	5,768	19.0	6,472	14.8	608	9.3	27,173	19.5	
	5 (least deprived)	16,948	29.0	4,648	15.3	4,012	9.2	275	4.2	25,883	18.6	
	Missing	39		23		73		19		154		
Ethnic group												
	White	56,116	96.7	29,256	97.4	42,471	98.1	6,301	98.2	134,144	97.3	<0.001
	Asian	1,359	2.3	519	1.7	473	1.1	41	0.6	2,392	1.7	
	Black	50	0.1	19	0.1	18	0.0	4	0.1	91	0.1	
	Mixed	409	0.7	179	0.6	205	0.5	39	0.6	832	0.6	
	Other	122	0.2	58	0.2	126	0.3	32	0.5	338	0.2	
	Missing	437		290		564		117		1,408		
Maternity factors												
Maternal age (years)												
	<=24	9,151	15.6	8,912	29.4	17,342	39.5	3,059	46.8	38,464	27.6	<0.001
	25-29	21,020	35.9	10,889	35.9	14,411	32.9	1,949	29.8	48,269	34.7	
	30-34	20,310	34.7	7,794	25.7	8,967	20.4	1,055	16.1	38,126	27.4	
	>=35	8,012	13.7	2,726	9.0	3,137	7.2	471	7.2	14,346	10.3	
	Missing	0		0		0		0		0		
Maternal smoking												
	No	39,993	82.6	16,931	68.0	20,131	55.5	2,252	41.6	79,307	69.0	<0.001
	Yes	8,445	17.4	7,959	32.0	16,153	44.5	3,157	58.4	35,714	31.0	
	Missing	10,055		5,431		7,573		1,125		24,184		
Parity												
	0	28,638	49.0	13,634	45.0	17,320	39.5	2,163	33.1	61,755	44.4	<0.001
	1	20,794	35.6	10,935	36.1	15,849	36.1	2,153	33.0	49,731	35.7	
	>1	9,060	15.5	5,750	19.0	10,687	24.4	2,218	33.9	27,715	19.9	
	Missing	1		2		1		0		4		
Mode of delivery												
	SVD	39,972	68.3	21,370	70.5	32,510	74.1	4,987	76.3	98,839	71.0	<0.001
	Cephalic	786	1.3	427	1.4	622	1.4	96	1.5	1,931	1.4	
	Assisted VD	7,689	13.1	3,385	11.2	4,075	9.3	525	8.0	15,674	11.3	
	Breech Delivery	198	0.3	102	0.3	190	0.4	25	0.4	515	0.4	
	Elective Caesarean	3,719	6.4	1,862	6.1	2,378	5.4	343	5.2	8,302	6.0	
	Emergency Caesarean	6,113	10.5	3,171	10.5	4,075	9.3	557	8.5	13,916	10.0	
	Other Delivery	16	0.0	4	0.0	7	0.0	1	0.0	28	0.0	
	Missing	0		0		0		0		0		

	High		General		Basic		Low		Total		
	N=58,493		N=30,321		N=43,857		N=6,534		N=139,205		
	n	%	n	%	n	%	n	%	n	%	P value
Maternity factors											
Gestation (weeks)											
<24	2	0.0	1	0.0	1	0.0	1	0.0	5	0.0	<0.001
24-27	46	0.1	34	0.1	56	0.1	15	0.2	151	0.1	
28-32	435	0.7	271	0.9	418	1.0	63	1.0	1,187	0.9	
33-36	2,332	4.0	1,297	4.3	2,121	4.8	337	5.2	6,087	4.4	
37	2,646	4.5	1,469	4.9	2,309	5.3	384	5.9	6,808	4.9	
38	7,221	12.4	3,889	12.8	5,518	12.6	921	14.1	17,549	12.6	
39	11,482	19.6	5,817	19.2	8,504	19.4	1,274	19.5	27,077	19.5	
40	18,723	32.0	9,545	31.5	13,592	31.0	1,973	30.2	43,833	31.5	
41	12,869	22.0	6,536	21.6	9,349	21.3	1,267	19.4	30,021	21.6	
42	2,629	4.5	1,387	4.6	1,897	4.3	287	4.4	6,200	4.5	
43	48	0.1	30	0.1	25	0.1	6	0.1	109	0.1	
>43	15	0.0	10	0.0	7	0.0	0	0.0	32	0.0	
Missing	45		35		60		6		146		
Sex-gestation-specific birthweight centile											
1-3	1,864	3.2	1,348	4.5	2,272	5.2	454	7.0	5,938	4.3	<0.001
4-10	4,615	7.9	2,890	9.5	4,615	10.5	829	12.7	12,949	9.3	
11-20	6,664	11.4	3,711	12.3	5,852	13.4	929	14.2	17,156	12.3	
21-80	35,508	60.8	17,695	58.5	25,127	57.4	3,550	54.4	81,880	58.9	
81-90	5,249	9.0	2,454	8.1	3,155	7.2	401	6.1	11,259	8.1	
91-97	3,236	5.5	1,566	5.2	1,987	4.5	265	4.1	7,054	5.1	
98-100	1,293	2.2	608	2.0	780	1.8	99	1.5	2,780	2.0	
Missing	64		49		69		7		189		
5 min Apgar											
1-3	438	0.7	249	0.8	358	0.8	71	1.1	1,116	0.8	0.002
4-6	537	0.9	287	0.9	461	1.1	69	1.1	1,354	1.0	
7-10	57,516	98.3	29,783	98.2	43,037	98.1	6,393	97.9	136,729	98.2	
Missing	2		2		1		1		6		

5.6.2 Association between diabetes and academic attainment

The diabetes analysis dataset contained 139,131 pupil census records after excluding 74 where ineligible diabetes drugs had been prescribed (section 4.3.2.2.1). The assumption of parallel odds was violated on univariate analysis (p=0.024). Therefore, the association between diabetes and academic attainment was described using different odds ratios within each category of attainment: OR1 - odds of general, basic or low attainment; OR2 - odds of basic or low attainment; OR3 - odds of low attainment. However, the parallel odds assumption was met after adjustment for sociodemographic and maternity confounders (p=0.163). Therefore, OR1, OR2 and OR3 described above were the same and the association was accurately described via one odds ratio (OR) across all attainment categories. Diabetes was not significantly associated with academic attainment level on univariate analysis (OR1 1.02, 95% CI 0.89-1.16; OR2 0.87, 95% CI 0.76-1.01; OR3 0.70, 95% CI 0.48-1.01) or after adjusting for sociodemographic (OR 1.04, 95% CI 0.91-1.21) and maternity (OR 1.14, 95% CI 0.99 -1.31) confounders.

5.6.3 Association between asthma and academic attainment

The asthma analysis dataset contained 125,211 pupil census records after excluding 13,994 where ineligible asthma drugs had been prescribed (section 4.3.2.2.2). The parallel odds assumption was met on univariate and multivariate analyses (p=0.524 and p=0.430 respectively); therefore, the association between asthma and academic attainment was adequately described via one odds ratio (OR) across all attainment categories. Asthma was significantly associated with academic attainment level on univariate analysis (OR 1.09, 95% CI 1.04-1.13) and after adjusting for sociodemographic (OR 1.10, 95% CI 1.05-1.16) and maternity (OR 1.11, 95% CI 1.06 -1.16) confounders. Children with asthma had an 11% increased odds of poorer academic attainment compared to unaffected peers. There was no significant univariate interaction between asthma and gender (p=0.504) or deprivation (p=0.633).

5.6.4 Association between epilepsy and academic attainment

The epilepsy analysis dataset contained 139,205 records. The assumption of parallel odds was violated on univariate and multivariate analyses (p<0.001). Therefore, the association between epilepsy and academic attainment was described using different odds ratios for each category of attainment: OR1 - odds of general, basic or low attainment; OR2 - odds of basic or low attainment; OR3 - odds of low attainment. Epilepsy was significantly

associated with academic attainment level on univariate analysis (OR1 1.42, 95% CI 1.25-1.62; OR2 1.42, 95% CI 1.25-1.61; OR3 2.58, 95% CI 2.12-3.14) and after adjusting for sociodemographic (OR1 1.96, 95% CI 1.68-2.29; OR2 2.44, 95% CI 2.07-2.87; OR3 3.80, 95% CI 3.05-4.74) and maternity (OR1 1.98, 95% CI 1.68-2.32; OR2 2.48, 95% CI 2.11-2.93; OR3 4.07, 95% CI 3.26-5.08) confounders. Children with epilepsy had 98% increased odds of having attainment classed as general, basic or low compared to non-epileptic children. These same children had 2.48 times greater odds of having attainment classed specifically as basic or low compared to non-epileptic peers. Finally, compared to non-epileptic peers, children with epilepsy had 4 times greater odds of specifically low attainment. There was no significant univariate interaction between epilepsy and gender (p=0.388). Whilst there was an initial univariate interaction between epilepsy and deprivation (p=0.008) this was borderline insignificant after adjustment for sociodemographic and maternity confounders (p=0.048). Table 49 presents odds ratios and confidence intervals from univariate and multivariate analyses; overall and within deprivation subgroups. The fully adjusted multivariate odds ratios and confidence intervals are visually summarised in Figure 21.

Table 49 Association between epilepsy and poor academic attainment level by area deprivation

		Univariate model		Multivariate[1]		Multivariate[2]	
		OR	95% CI	OR	95% CI	OR	95% CI
Overall	general, basic or low[3]	1.42***	1.25-1.62	1.96***	1.68-2.29	1.98***	1.68-2.32
	basic or low[4]	1.42***	1.25-1.61	2.44***	2.07-2.87	2.48***	2.11-2.93
	low[5]	2.58***	2.12-3.14	3.80***	3.05-4.74	4.07***	3.26-5.08
1 - most deprived	general, basic or low[3]	1.38*	1.07-1.78	2.08***	1.58-2.73	2.10***	1.59-2.77
	basic or low[4]	"	"	"	"	"	"
	low[5]	"	"	"	"	"	"
2	general, basic or low[3]	1.26	0.93-1.69	1.83***	1.31-2.56	1.93***	1.37-2.72
	basic or low[4]	1.22	0.94-1.60	2.34***	1.69-3.24	2.46***	1.76-3.43
	low[5]	2.47***	1.68-3.65	3.88***	2.53-5.97	4.15***	2.68-6.42
3	general, basic or low[3]	1.35*	1.02-1.80	2.05***	1.48-2.84	2.05***	1.47-2.86
	basic or low[4]	1.51**	1.15-1.98	3.07***	2.19-4.32	3.12***	2.20-4.42
	low[5]	3.95***	2.64-5.91	6.46***	4.09-10.2	7.21***	4.56-11.4
4	general, basic or low[3]	1.43*	1.08-1.90	1.75**	1.25-2.45	1.68**	1.19-2.38
	basic or low[4]	1.51**	1.13-2.03	2.17***	1.49-3.17	2.17***	1.48-3.18
	low[5]	3.42***	1.97-5.92	3.90***	2.17-7.01	4.07***	2.26-7.34
5 - least deprived	general, basic or low[3]	2.05***	1.51-2.77	2.66***	1.86-3.80	2.84***	1.98-4.08
	basic or low[4]	1.82***	1.29-2.57	2.44***	1.57-3.78	2.60***	1.65-4.08
	low[5]	6.67***	3.58-12.4	6.98***	3.55-13.7	7.41***	3.72-14.7

Exponentiated coefficients; 95% confidence intervals in brackets

* p<0.05, ** p<0.01, *** p<0.001

[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)

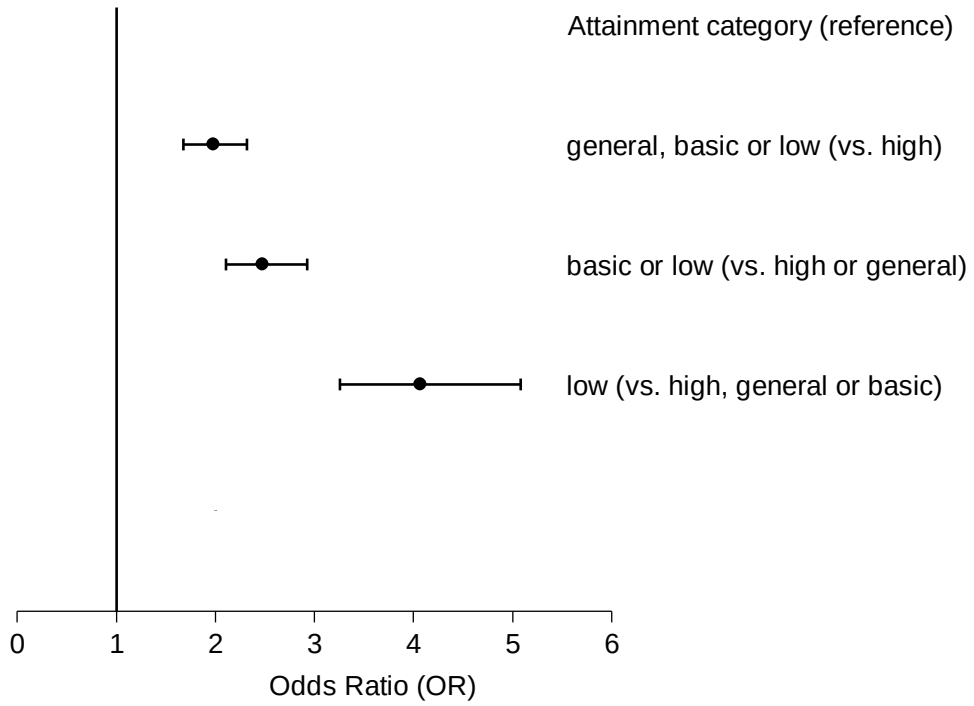
[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

[3]reference category of high attainment

[4]reference category of high or general attainment

[5]reference category of high, general or basic attainment

Figure 21 Forest plot of the fully adjusted multivariate association between epilepsy and poor academic attainment level



5.6.5 Association between attention deficit hyperactivity disorder and academic attainment

The ADHD analysis dataset contained 139,205 records. The assumption of parallel odds was violated on univariate analysis (p=0.023) and after adjustment for sociodemographic confounders. Therefore, the association between ADHD and academic attainment was described using different odds ratios within each category of attainment: OR1 - odds of general, basic or low attainment; OR2 - odds of basic or low attainment; OR3 - odds of low attainment. However, a final model adjusted for sociodemographic and maternity confounders met the assumption of parallel odds (p=0.106). Therefore OR1, OR2 and OR3 described above were the same and the association between ADHD and academic attainment was accurately described using one odds ratio (OR) across all categories of attainment. ADHD was significantly associated with academic attainment level on univariate analysis (OR1 8.72, 95% CI 7.16-10.60; OR2 6.66, 95% CI 5.85-7.58; OR3 6.18, 95% CI 5.44-7.03) and after adjusting for sociodemographic (OR1 5.06, 95% CI 4.04-6.34; OR2 4.72, 95% CI 3.98-5.60; OR3 3.60, 95% CI 3.12-4.16) and maternity (OR 3.64, 95% CI 3.26-4.07) confounders. Children with ADHD had 3.64 times greater odds of having poorer academic attainment compared to children without ADHD. There were significant univariate (p=0.003) and multivariate (p=0.013) interactions between ADHD

and gender whereby the association between ADHD and academic attainment was stronger for girls than for boys. On fully adjusted subgroup analyses, girls with ADHD had 5.3 times greater odds of poorer attainment compared to unaffected girls. The picture was more complex for boys because the parallel odds assumption did not hold between categories. Boys with ADHD had 4.01 times greater odds of having general, basic or low attainment rather than high attainment compared to boys without ADHD. Boys with ADHD similarly experienced 4.09 times greater odds of having specifically basic or low attainment and 2.94 times greater odds of having specifically low attainment compared to boys without ADHD. Whilst there was an initial interaction between ADHD and deprivation ($p=0.023$), this disappeared after adjustment for sociodemographic and maternity confounders ($p=0.106$). For unaffected children, boys were more likely to have poorer attainment than girls; therefore the relative impact of ADHD on attainment was greater for girls. Table 50 presents odds ratios and confidence intervals from univariate and multivariate analyses; overall and within sex and deprivation subgroups. The fully adjusted multivariate odds ratios and confidence intervals are summarised visually in Figure 22.

Table 50 Association between attention deficit hyperactivity disorder and poor academic attainment level by sex and area deprivation

		Univariate model		Multivariate[1]		Multivariate[2]	
		OR	95% CI	OR	95% CI	OR	95% CI
Overall	general, basic or low[3]	8.72***	7.16-10.6	5.06***	4.04-6.34	3.64***	3.26-4.07
	basic or low[4]	6.66***	5.85-7.58	4.72***	3.98-5.60	"	"
	low[5]	6.18***	5.44-7.03	3.60***	3.12-4.16	"	"
Boys	general, basic or low[3]	7.12***	5.73-8.84	4.62***	3.62-5.90	4.01***	3.13-5.15
	basic or low[4]	5.75***	4.98-6.63	4.63***	3.84-5.57	4.09***	3.39-4.94
	low[5]	5.26***	4.55-6.08	3.24***	2.76-3.80	2.94***	2.50-3.45
Girls	general, basic or low[3]	9.92***	6.20-15.8	6.11***	4.67-7.99	5.30***	4.04-6.96
	basic or low[4]	6.91***	5.09-9.38	"	"	"	"
	low[5]	9.58***	7.18-12.8	"	"	"	"
1 - most deprived	general, basic or low[3]	18.47***	8.25-41.39	10.46***	4.47-24.49	3.59***	2.93-4.41
	basic or low[4]	6.73***	4.90-9.24	4.67***	3.17-6.89	"	"
	low[5]	4.90***	3.97-6.06	3.48***	2.75-4.39	"	"
2	general, basic or low[3]	8.87***	5.59-14.0	3.63***	2.94-4.48	3.31***	2.68-4.09
	basic or low[4]	5.97***	4.57-7.79	"	"	"	"
	low[5]	4.64***	3.62-5.95	"	"	"	"
3	general, basic or low[3]	6.57***	5.21-8.29	4.68***	3.66-5.98	3.88***	3.03-4.97
	basic or low[4]	"	"	"	"	"	"
	low[5]	"	"	"	"	"	"
4	general, basic or low[3]	8.26***	6.26-10.9	5.93***	4.42-7.95	4.91***	3.64-6.64
	basic or low[4]	"	"	"	"	"	"
	low[5]	"	"	"	"	"	"
5 - least deprived	general, basic or low[3]	5.93***	4.37-8.05	3.20***	2.28-4.49	2.67***	1.89-3.78
	basic or low[4]	"	"	"	"	"	"
	low[5]	"	"	"	"	"	"

Exponentiated coefficients; 95% confidence intervals in brackets

* p<0.05, ** p<0.01, *** p<0.001

[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)

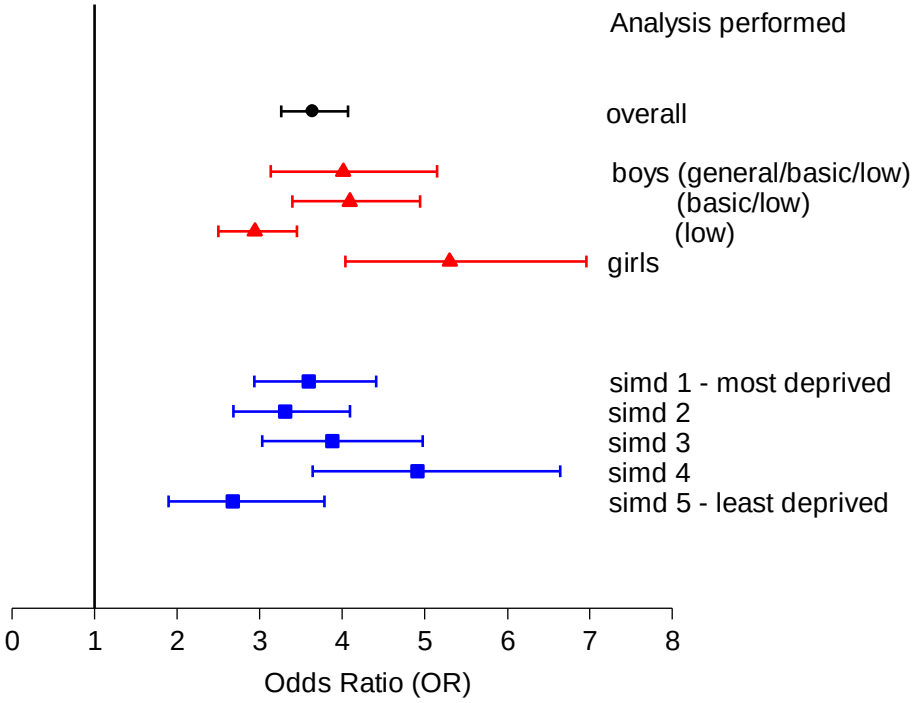
[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

[3]reference category of high attainment

[4]reference category of high or general attainment

[5]reference category of high, general or basic attainment

Figure 22 Forest plot of the fully adjusted multivariate association between attention deficit hyperactivity disorder and poor academic attainment level by sex and area deprivation



5.6.6 Association between depression and academic attainment

The depression analysis dataset contained 139,199 pupil census records after excluding 6 where ineligible anti-depressant drugs had been prescribed (section 4.3.2.2.5). The assumption of parallel odds was violated on univariate and multivariate analyses ($p<0.001$). Therefore, the association between depression and academic attainment was described using different odds ratios between each category of attainment: OR1 - odds of general, basic or low attainment; OR2 - odds of basic or low attainment; OR3 - odds of low attainment. Depression was significantly associated with academic attainment level on univariate analysis (OR1 1.07, 95% CI 0.99-1.17; OR2 0.94, 95% CI 0.87-1.03; OR3 1.69, 95% CI 1.45-1.97) and after adjusting for sociodemographic (OR1 1.95, 95% CI 1.77-2.16; OR2 2.10, 95% CI 1.88-2.34; OR3 3.28, 95% CI 2.76-3.89) and maternity (OR1 1.96, 95% CI 1.77-2.17; OR2 2.14, 95% CI 1.92-2.39; OR3 3.44, 95% CI 2.89-4.09) confounders. Children with depression had a 96% increased odds of having attainment classed as general, basic or low compared to children without depression. These same children had 2.14 times greater odds of having attainment classed specifically as basic or low compared to unaffected peers. Finally, compared to unaffected peers, children with depression had 3.44 times greater odds of specifically low attainment.

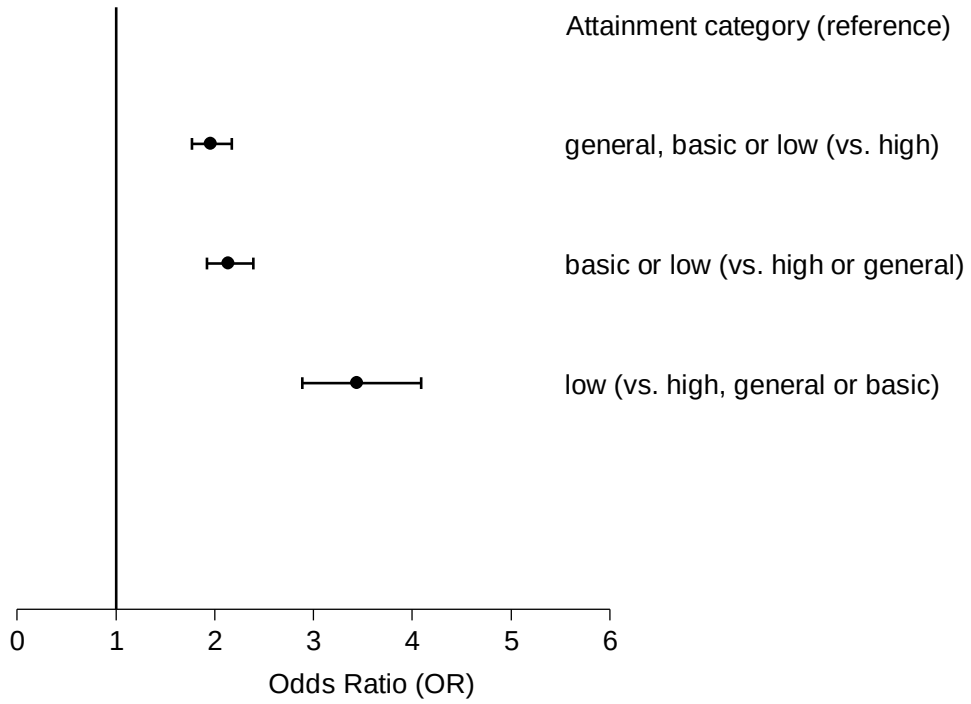
There was no significant univariate interaction between depression and gender (p=0.142). Whilst there was a significant univariate interaction between depression and deprivation (p<0.001) this disappeared after adjusting for sociodemographic and maternity confounders (p=0.090). Table 51 presents odds ratios and confidence intervals from univariate and multivariate analyses; overall and within deprivation subgroups. The fully adjusted multivariate odds ratios and confidence intervals are visually summarised in Figure 23.

Table 51 Association between depression and poor academic attainment level by area deprivation

		Univariate model		Multivariate[1]		Multivariate[2]	
		OR	95% CI	OR	95% CI	OR	95% CI
Overall	general, basic or low[3]	1.07	0.99-1.17	1.95***	1.77-2.16	1.96***	1.77-2.17
	basic or low[4]	0.94	0.87-1.03	2.10***	1.88-2.34	2.14***	1.92-2.39
	low[5]	1.69***	1.45-1.97	3.28***	2.76-3.89	3.44***	2.89-4.09
1 - most deprived	general, basic or low[3]	0.73**	0.58-0.91	1.37*	1.07-1.76	1.34*	1.04-1.73
	basic or low[4]	0.75**	0.62-0.91	1.80***	1.41-2.29	1.84***	1.44-2.35
	low[5]	1.52**	1.14-2.02	3.22***	2.34-4.44	3.38***	2.44-4.69
2	general, basic or low[3]	1.07	0.88-1.30	1.88***	1.51-2.35	1.90***	1.52-2.38
	basic or low[4]	0.97	0.81-1.16	1.97***	1.57-2.47	1.99***	1.58-2.51
	low[5]	2.07***	1.56-2.73	3.44***	2.52-4.68	3.62***	2.65-4.94
3	general, basic or low[3]	1.06	0.89-1.26	2.00***	1.68-2.39	2.00***	1.67-2.40
	basic or low[4]	0.91	0.76-1.10	"	"	"	"
	low[5]	1.34	0.90-2.00	"	"	"	"
4	general, basic or low[3]	1.29**	1.08-1.55	2.33***	1.93-2.82	2.39***	1.97-2.89
	basic or low[4]	1.16	0.95-1.42	"	"	"	"
	low[5]	1.96**	1.25-3.10	"	"	"	"
5 - least deprived	general, basic or low[3]	1.49***	1.24-1.80	2.40***	1.92-3.00	2.47***	1.98-3.10
	basic or low[4]	1.38**	1.10-1.73	2.37***	1.78-3.16	2.50***	1.87-3.33
	low[5]	4.09***	2.51-6.66	5.65***	3.35-9.55	6.11***	3.60-10.3

Exponentiated coefficients; 95% confidence intervals in brackets
* p<0.05, ** p<0.01, *** p<0.001
[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)
[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score
[3]reference category of high attainment
[4]reference category of high or general attainment
[5]reference category of high, general or basic attainment

Figure 23 Forest plot of the fully adjusted multivariate association between depression and poor academic attainment level



5.6.7 Association between each chronic condition and academic attainment after excluding children with special educational need and adjusting for absenteeism

Adjusting for number of days absent from school per census year attenuated the associations between asthma, epilepsy and depression and academic attainment. However this had no notable impact on the association between ADHD and academic attainment. Excluding children with SEN attenuated the associations between epilepsy, ADHD and depression and academic attainment. However this had no notable impact on the association between asthma and academic attainment. The association between diabetes and academic attainment was not further investigated by adjusting for absenteeism or excluding children with SEN because the fully adjusted multivariate model in 5.6.2 was not significant.

There was a significant association between asthma and academic attainment after adjusting for sociodemographic and maternity (OR 1.11, 95% CI 1.06-1.16) confounders whereby children with asthma had an 11% increased odds of poorer academic attainment. The increased odds reduced slightly to 8% but remained significant after excluding

children with SEN from the analysis (OR 1.08, 95% CI 1.03-1.14). Further, the association was no longer significant after adjusting for absenteeism (OR 0.96, 95% CI 0.91-1.01).

There was a significant association between ADHD and academic attainment after adjusting for sociodemographic and maternity (OR 3.64, 95% CI 3.26-4.07) confounders whereby children with ADHD had 3.6 times greater odds of poorer academic attainment. After excluding children with SEN, children with ADHD experienced 2.8 times greater odds of poorer attainment (OR 2.79, 95% CI 2.35-3.31) which increased after adjusting for absenteeism (OR 4.07, 95% CI 3.62-4.59). The association remained significant in both cases.

The associations between epilepsy and depression and academic attainment were described using different odds ratios between each category of attainment: OR1 - odds of general, basic or low attainment; OR2 - odds of basic or low attainment; OR3 - odds of low attainment. Epilepsy was significantly associated with academic attainment after adjusting for sociodemographic and maternity (OR1 1.98, 95% CI 1.68-2.32; OR2 2.48, 95% CI 2.11-2.93; OR3 4.07, 95% CI 3.26-5.08) confounders. Excluding children with SEN from the analysis attenuated the strength of this association (OR1 1.70, 95% CI 1.41-2.05; OR2 1.94, 95% CI 1.57-2.39; OR3 3.29, 95% CI 2.36-4.58). Adjusting for absenteeism also attenuated the association (OR1 1.23, 95% CI 1.03-1.46; OR2 1.62, 95% CI 1.35-1.93; OR3 3.07, 95% CI 2.39-3.95) but both remained significant.

There was a significant association between depression and academic attainment after adjusting for sociodemographic and maternity (OR1 1.96, 95% CI 1.77-2.17; OR2 2.14, 95% CI 1.92-2.39; OR3 3.44, 95% CI 2.89-4.09) confounders. Excluding children with SEN from the analysis attenuated the strength of this association (OR1 1.83, 95% CI 1.64-2.05; OR2 1.95, 95% CI 1.71-2.22; OR3 3.28, 95% CI 2.61-4.12) but it remained significant. Adjusting for absenteeism also attenuated the association (OR1 0.98, 95% CI 0.87-1.09; OR2 1.03, 95% CI 0.91-1.17; OR3 1.89, 95% CI 1.55-2.30); however only the lowest category of attainment remained significant. The results of each of the original analyses and of analyses excluding children with a SEN and adjusted for absenteeism are shown in Table 52 below.

Table 52 Association between chronic diseases and academic attainment in children after adjusting for absenteeism and excluding children with special educational needs

		multivariate model[1]		adjusted for absenteeism[2a]		excluding children with SEN[2b]	
		OR	95% CI	OR	95% CI	OR	95% CI
Asthma		1.11***	1.06-1.16	0.96	0.91-1.01	1.08**	1.03-1.14
Epilepsy	general, basic or low[3]	1.98***	1.68-2.32	1.23*	1.03-1.46	1.70***	1.41-2.05
	basic or low[4]	2.48***	2.11-2.93	1.62***	1.35-1.93	1.94***	1.57-2.39
	low[5]	4.07***	3.26-5.08	3.07***	2.39-3.95	3.29***	2.36-4.58
ADHD		3.64***	3.26-4.07	4.07***	3.62-4.59	2.79***	2.35-3.31
Depression	general, basic or low[3]	1.96***	1.77-2.17	0.98	0.87-1.09	1.83***	1.64-2.05
	basic or low[4]	2.14***	1.92-2.39	1.03	0.91-1.17	1.95***	1.71-2.22
	low[5]	3.44***	2.89-4.09	1.89***	1.55-2.30	3.28***	2.61-4.12

Exponentiated coefficients; 95% confidence intervals in brackets
** p<0.05, ** p<0.01, *** p<0.001*

[1]adjusted for sex, age, SIMD deprivation quintile, ethnic group, maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

[2a] mulivariate model repeated additionally adjusting for number of days absent from school per census year

[2b] mulivariate model repeated after excluding children with special educational needs

[3]reference category of high attainment

[4]reference category of high or general attainment

[5]reference category of high, general or basic attainment

5.7 Unemployment six months after leaving school

5.7.1 Characteristics of children by whether or not they were unemployed

The final leaver dataset contained 217,924 pupil records across census years 2009 through to 2013 with one record present per pupil. Each chronic condition with the exception of diabetes was significantly associated with unemployment (Table 53). Significantly more boys than girls were not in employment, education or training six months after leaving school. Children who were unemployed were significantly younger, more deprived and significantly more likely to be of white ethnicity. They had significantly lower sex-gestation-specific birthweight centiles, gestational ages and 5-minute Apgar scores and their mothers were significantly more likely to have been younger than 25 years of age, multiparous and to have smoked during pregnancy. Children who were unemployed were more likely to have been delivered by spontaneous vaginal delivery and less likely to have been delivered by assisted vaginal delivery or caesarean section (Table 54).

Table 53 Unemployment (NEET) by presence or absence of chronic conditions

		Not NEET N=195,205		NEET N=22,719		Total N=217,924		
		n	%	n	%	n	%	P value
Diabetes								
	No	194,019	99.4	22,569	99.3	216,588	99.4	0.336
	Yes	1,186	0.6	150	0.7	1,336	0.6	
	Missing	0		0		0		
Asthma								
	No	184,583	94.6	21,357	94.0	205,940	94.5	<0.001
	Yes	10,622	5.4	1,362	6.0	11,984	5.5	
	Missing	0		0		0		
Epilepsy								
	No	193,714	99.2	22,400	98.6	216,114	99.2	<0.001
	Yes	1,491	0.8	319	1.4	1,810	0.8	
	Missing	0		0		0		
ADHD								
	No	193,352	99.1	22,170	97.6	215,522	98.9	<0.001
	Yes	1,853	0.9	549	2.4	2,402	1.1	
	Missing	0		0		0		
Depression								
	No	192,364	98.5	22,166	97.6	214,530	98.4	<0.001
	Yes	2,841	1.5	553	2.4	3,394	1.6	
	Missing	0		0		0		

Table 54 Characteristics of children by whether or not they were unemployed (NEET)

		Not NEET N=195,205		NEET N=22,719		Total N=217,924		
		n	%	n	%	n	%	P value
Sociodemographic factors								
Sex								
	male	97,371	49.9	13,121	57.8	110,492	50.7	<0.001
	female	97,834	50.1	9,598	42.2	107,432	49.3	
	Missing	0		0		0		
Age	14	900	0.5	224	1.0	1,124	0.5	<0.001
	15	49,291	25.3	12,366	54.4	61,657	28.3	
	16	71,460	36.6	5,645	24.8	77,105	35.4	
	17	72,975	37.4	4,361	19.2	77,336	35.5	
	18	579	0.3	123	0.5	702	0.3	
	Missing	0		0		0		
Deprivation quintile								
	1 (most deprived)	39,986	20.5	8,149	35.9	48,135	22.1	<0.001
	2	38,340	19.7	5,587	24.6	43,927	20.2	
	3	38,639	19.8	4,047	17.8	42,686	19.6	
	4	39,231	20.1	2,853	12.6	42,084	19.3	
	5 (least deprived)	38,811	19.9	2,042	9.0	40,853	18.8	
	Missing	198		41		239		
Ethnic group								
	White	188,194	97.4	21,910	97.9	210,104	97.5	<0.001
	Asian	3,375	1.7	270	1.2	3,645	1.7	
	Black	98	0.1	9	0.0	107	0.0	
	Mixed	1,169	0.6	126	0.6	1,295	0.6	
	Other	364	0.2	70	0.3	434	0.2	
	Missing	2,005		334		2,339		
Maternity factors								
Maternal age (years)								
	<=24	52,380	26.8	8,915	39.2	61,295	28.1	<0.001
	25-29	69,310	35.5	6,963	30.6	76,273	35.0	
	30-34	53,723	27.5	4,840	21.3	58,563	26.9	
	>=35	19,791	10.1	2,000	8.8	21,791	10.0	
	Missing	1		1		2		
Maternal smoking								
	No	113,893	70.6	9,637	50.9	123,530	68.5	<0.001
	Yes	47,483	29.4	9,313	49.1	56,796	31.5	
	Missing	33,829		3,769		37,598		
Parity								
	0	88,699	45.4	8,394	37.0	97,093	44.6	<0.001
	1	69,216	35.5	7,732	34.0	76,948	35.3	
	>1	37,258	19.1	6,585	29.0	43,843	20.1	
	Missing	32		8		40		

		Not NEET N=195,205		NEET N=22,719		Total N=217,924		
		n	%	n	%	n	%	P value
Maternity factors								
Mode of delivery								
	SVD	137,162	70.3	16,938	74.6	154,100	70.7	<0.001
	Cephalic	2,979	1.5	395	1.7	3,374	1.5	
	Assisted VD	22,893	11.7	2,004	8.8	24,897	11.4	
	Breech Delivery	760	0.4	95	0.4	855	0.4	
	Elective Caesarean	11,672	6.0	1,222	5.4	12,894	5.9	
	Emergency Caesarean	19,697	10.1	2,059	9.1	21,756	10.0	
	Other Delivery	42	0.0	6	0.0	48	0.0	
	Missing	0		0		0		
Gestation (weeks)								
	<24	7	0.0	0	0.0	7	0.0	<0.001
	24-27	225	0.1	32	0.1	257	0.1	
	28-32	1,673	0.9	219	1.0	1,892	0.9	
	33-36	8,410	4.3	1,184	5.2	9,594	4.4	
	37	9,387	4.8	1,255	5.5	10,642	4.9	
	38	24,362	12.5	2,956	13.0	27,318	12.6	
	39	37,763	19.4	4,499	19.8	42,262	19.4	
	40	62,091	31.9	7,010	30.9	69,101	31.8	
	41	41,600	21.3	4,547	20.0	46,147	21.2	
	42	9,136	4.7	960	4.2	10,096	4.6	
	43	201	0.1	18	0.1	219	0.1	
	>43	54	0.0	4	0.0	58	0.0	
	Missing	296		35		331		
Sex-gestation-specific birthweight centile								
	1-3	8,177	4.2	1,235	5.4	9,412	4.3	<0.001
	4-10	17,921	9.2	2,461	10.9	20,382	9.4	
	11-20	23,801	12.2	3,048	13.4	26,849	12.3	
	21-80	114,995	59.0	12,969	57.2	127,964	58.8	
	81-90	16,110	8.3	1,581	7.0	17,691	8.1	
	91-97	9,968	5.1	992	4.4	10,960	5.0	
	98-100	3,859	2.0	390	1.7	4,249	2.0	
	Missing	374		43		417		
5 min Apgar								
	1-3	1,447	0.7	204	0.9	1,651	0.8	0.006
	4-6	1,963	1.0	244	1.1	2,207	1.0	
	7-10	191,748	98.3	22,265	98.0	214,013	98.2	
	Missing	47		6		53		

5.7.2 Association between diabetes and unemployment

The diabetes analysis dataset contained 217,805 pupil census records after excluding 119 where ineligible diabetes drugs had been prescribed (section 4.3.2.2.1). Diabetes was not significantly associated with unemployment on univariate analysis (OR 1.09, 95% CI 0.92-1.29) or after adjusting for sociodemographic (OR 1.13, 95% CI 0.94-1.34) and maternity (OR 1.18, 95% CI 0.99-1.41) confounders.

5.7.3 Association between asthma and unemployment

The asthma analysis dataset contained 197,430 pupil census records after excluding 20,494 where ineligible asthma drugs had been prescribed (section 4.3.2.2.2). Asthma was significantly associated with unemployment on univariate analysis (OR 1.11, 95% CI 1.05-1.18) however this association disappeared after adjusting for sociodemographic (OR 1.00, 95% CI 0.94-1.06) and maternity (OR 1.00, 95% CI 0.95-1.07) confounders.

5.7.4 Association between epilepsy and unemployment

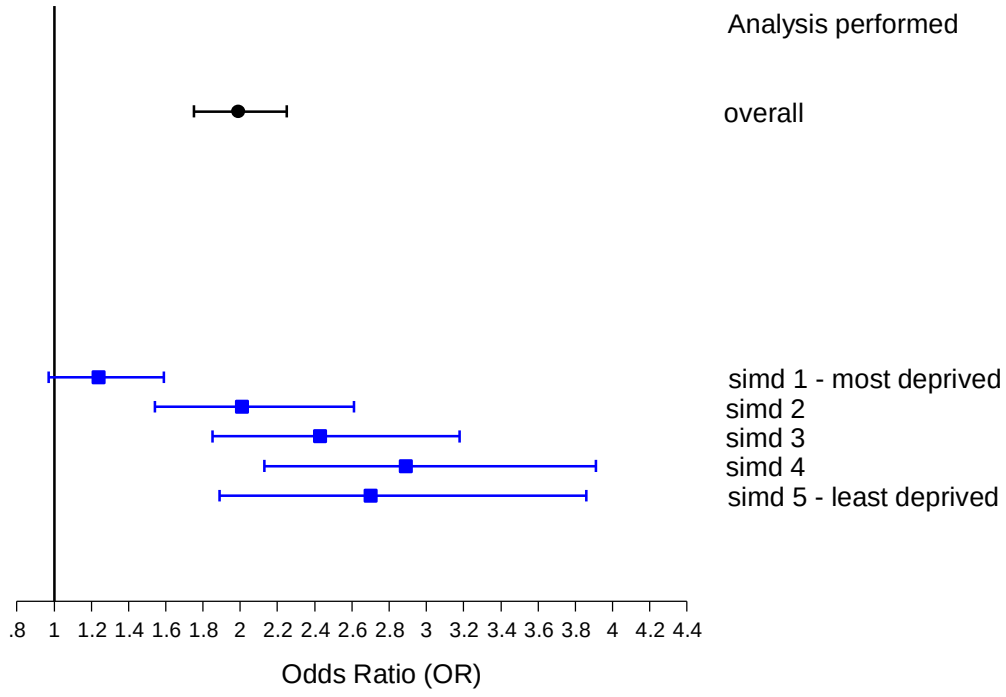
The epilepsy analysis dataset contained 217,924 records. Epilepsy was significantly associated with unemployment on univariate analysis (OR 1.85, 95% CI 1.64-2.09) and after adjusting for sociodemographic (OR 1.96, 95% CI 1.73-2.23) and maternity (OR 1.99, 95% CI 1.75-2.25) confounders. Pupils with epilepsy had a 99% increased odds of unemployment compared to pupils who did not have epilepsy. There was no significant univariate interaction between epilepsy and gender ($p=0.740$); however, there were significant univariate and multivariate interactions between epilepsy and deprivation (both $p<0.001$). On fully adjusted subgroup analyses the association between epilepsy and unemployment was stronger for less deprived children. The greater relative impact of epilepsy on unemployment was due to higher baseline unemployment in the most deprived children among those without epilepsy (8,068 [16.9%] vs. 2,005 [4.9%] in the most and least deprived quintiles respectively). Among children with epilepsy, unemployment was still higher in the most deprived quintile (81 [18.5%] vs. 37 [12.2%]). Table 55 presents odds ratios and confidence intervals from univariate and multivariate analyses; overall and within deprivation subgroups. The fully adjusted multivariate odds ratios and confidence intervals are summarised visually in Figure 24.

Table 55 Association between epilepsy and unemployment by area deprivation

	Univariate model		Multivariate[1]		Multivariate[2]	
	OR	95% CI	OR	95% CI	OR	95% CI
Overall	1.85***	1.64-2.09	1.96***	1.73-2.23	1.99***	1.75-2.25
1 - most deprived	1.13	0.88-1.43	1.22	0.95-1.57	1.24	0.97-1.59
2	1.72***	1.33-2.22	1.97***	1.51-2.57	2.01***	1.54-2.61
3	2.28***	1.76-2.95	2.45***	1.88-3.20	2.43***	1.85-3.18
4	2.90***	2.17-3.88	2.87***	2.13-3.87	2.89***	2.13-3.91
5 - least deprived	2.65***	1.88-3.75	2.64***	1.85-3.76	2.70***	1.89-3.86

Exponentiated coefficients; 95% confidence intervals in brackets
* p<0.05, ** p<0.01, *** p<0.001
[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)
[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

Figure 24 Forest plot of the fully adjusted multivariate association between epilepsy and unemployment by area deprivation



5.7.5 Association between attention deficit hyperactivity disorder and unemployment

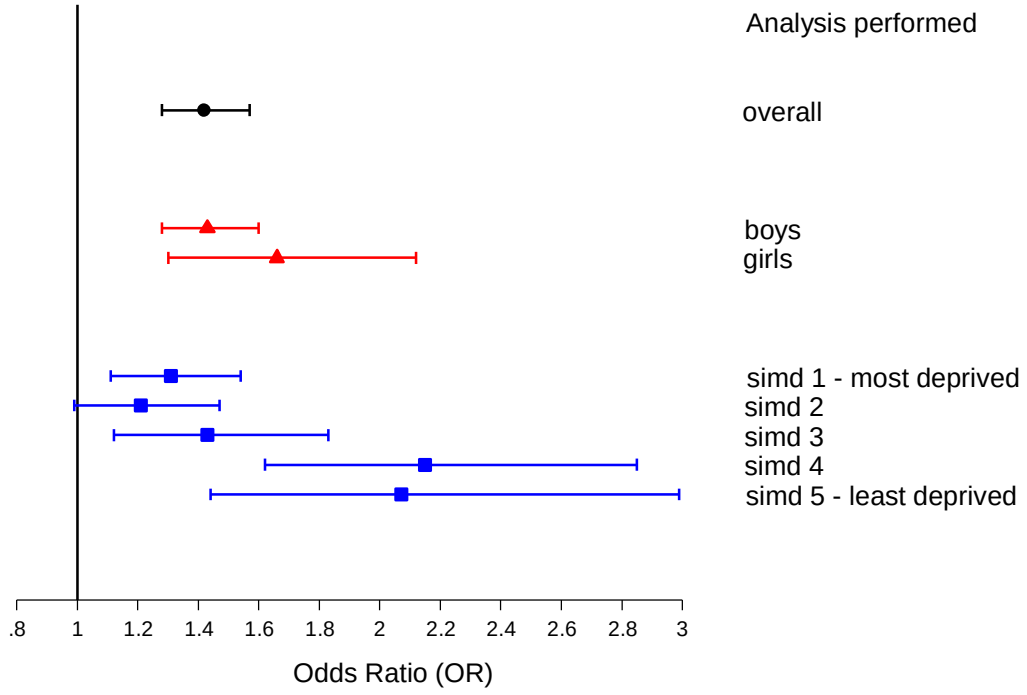
The ADHD analysis dataset contained 217,924 records. ADHD was significantly associated with unemployment on univariate analysis (OR 2.58, 95% CI 2.35-2.84) and after adjusting for sociodemographic (OR 1.52, 95% CI 1.37-1.68) and maternity (OR 1.42, 95% CI 1.28-2.57) confounders. Pupils with ADHD had 42% increased odds of unemployment compared to pupils who did not have ADHD. There were significant univariate and multivariate interactions between ADHD and gender ($p=0.007$ and $p=0.036$ respectively) and deprivation ($p<0.001$ and $p=0.001$ respectively). On fully adjusted subgroup analyses the association between ADHD and unemployment was stronger for girls and children in the least deprived 40% of the population. The relative impact of ADHD on unemployment was weaker for boys because, among unaffected children, they had a higher absolute risk of unemployment compared to girls (12,644 [11.7%] vs. 9,506 [8.9%]). Among children with ADHD, unemployment was still higher in boys (457 [22.7%] vs. 92 [23.6%]). The greater relative impact of ADHD on unemployment in less deprived children was also due to higher baseline unemployment among the most deprived children for those without ADHD (7,936 [16.7%] vs. 2,003 [4.9%] in the most and least deprived quintiles respectively). Among children with ADHD, unemployment was still higher in the most deprived quintile (213 [27.7%] vs. 39 [16.1%]). Table 56 presents odds ratios and confidence intervals from univariate and multivariate analyses; overall and within sex and deprivation subgroups. The fully adjusted multivariate odds ratios and confidence intervals are summarised visually in Figure 25.

Table 56 Association between attention deficit hyperactivity disorder and unemployment by sex and area deprivation

	Univariate model		Multivariate[1]		Multivariate[2]	
	OR	95% CI	OR	95% CI	OR	95% CI
Overall	2.58***	2.35-2.84	1.52***	1.37-1.68	1.42***	1.28-1.57
Boys	2.22***	2.00-2.47	1.54***	1.38-1.71	1.43***	1.28-1.60
Girls	3.17***	2.50-4.01	1.74***	1.36-2.22	1.66***	1.30-2.12
1 - most deprived	1.90***	1.62-2.23	1.37***	1.16-1.61	1.31**	1.11-1.54
2	1.99***	1.65-2.41	1.28*	1.05-1.55	1.21	0.99-1.47
3	2.47***	1.96-3.13	1.61***	1.27-2.05	1.43**	1.12-1.83
4	3.78***	2.88-4.95	2.40***	1.82-3.17	2.15***	1.62-2.85
5 - least deprived	3.70***	2.62-5.23	2.27***	1.59-3.23	2.07***	1.44-2.99

Exponentiated coefficients; 95% confidence intervals in brackets
* p<0.05, ** p<0.01, *** p<0.001
[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)
[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

Figure 25 Forest plot of the fully adjusted multivariate association between attention deficit hyperactivity disorder and unemployment by sex and area deprivation



5.7.6 Association between depression and unemployment

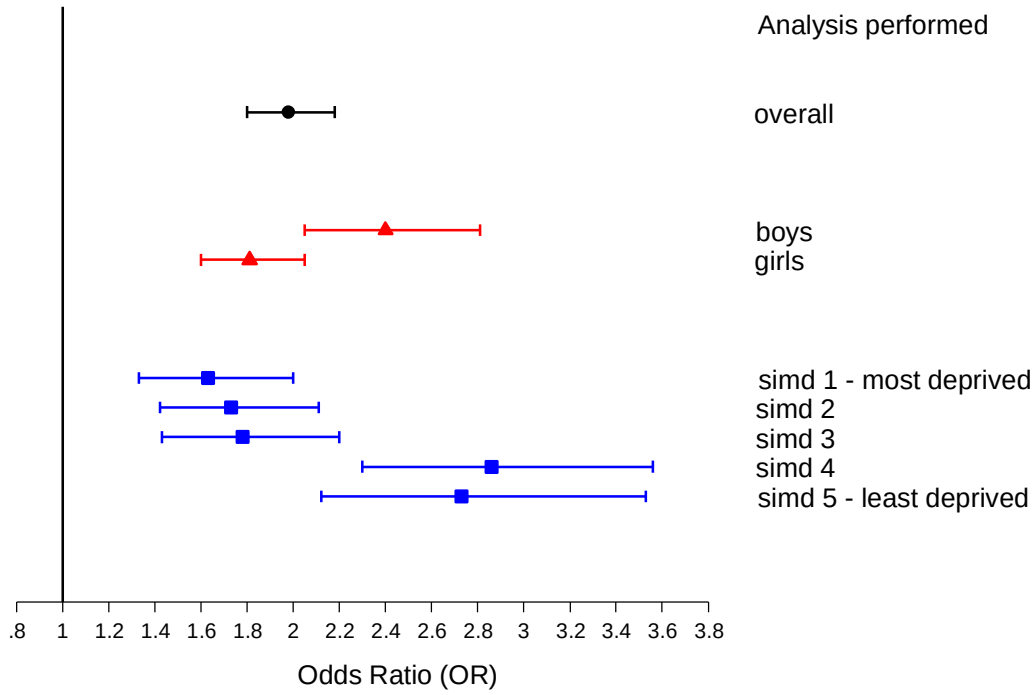
The depression analysis dataset contained 217,919 pupil census records after excluding 5 where ineligible anti-depressant drugs had been prescribed (section 4.3.2.2.5). Depression was significantly associated with unemployment on univariate analysis (OR 1.69, 95% CI 1.54-1.85) and after adjusting for sociodemographic (OR 1.96, 95% CI 1.78-2.16) and maternity (OR 1.98, 95% CI 1.80-2.18) confounders. Pupils with depression had 98% increased odds of unemployment compared to pupils who did not have depression. There were significant univariate and multivariate interactions between depression and gender (both $p=0.001$) and deprivation (both $p<0.001$). On fully adjusted subgroup analyses the association between depression and unemployment was stronger for boys than girls (OR 2.40 versus 1.81). Boys who were depressed experienced 2.4 fold odds of unemployment compared to boys who were not depressed. The corresponding increase in odds was 1.8 fold for depressed girls. Among unaffected children, boys had a higher absolute risk of unemployment compared to girls (12,898 [11.8%] vs. 9,267 [8.8%]). Among children with depression, unemployment was still higher in boys (223 [23.0%] vs. 330 [13.6%]).The association between depression and unemployment was stronger for less deprived children. Within the least deprived 20% of the population, depressed children had 2.73 times greater odds of unemployment compared to children who were not depressed. The corresponding increase in odds was 1.63 fold within the most deprived 20% of the population. The relative impact of depression on unemployment was greater for less deprived children which was due to higher baseline unemployment among the most deprived children for those unaffected by depression (8,019 [16.9%] vs. 1,969 [4.9%] in the most and least deprived quintiles respectively). Among depressed children, unemployment was still higher in the most deprived quintile (130 [21.3%] vs. 73 [11.4%]). Table 57 presents odds ratios and confidence intervals from univariate and multivariate analyses; overall and within sex and deprivation subgroups. The multivariate odds ratios and confidence intervals are summarised visually in Figure 26.

Table 57 Association between depression and unemployment by sex and area deprivation

	Univariate model		Multivariate[1]		Multivariate[2]	
	OR	95% CI	OR	95% CI	OR	95% CI
Overall	1.69***	1.54-1.85	1.96***	1.78-2.16	1.98***	1.80-2.18
Boys	2.24***	1.92-2.60	2.35***	2.01-2.75	2.40***	2.05-2.81
Girls	1.63***	1.45-1.83	1.80***	1.59-2.04	1.81***	1.60-2.05
1 - most deprived	1.33**	1.10-1.62	1.58***	1.29-1.93	1.63***	1.33-2.00
2	1.56***	1.29-1.89	1.71***	1.41-2.08	1.73***	1.42-2.11
3	1.63***	1.33-2.01	1.82***	1.47-2.25	1.78***	1.43-2.20
4	2.68***	2.17-3.31	2.88***	2.32-3.58	2.86***	2.30-3.56
5 - least deprived	2.50***	1.95-3.20	2.68***	2.08-3.46	2.73***	2.12-3.53

Exponentiated coefficients; 95% confidence intervals in brackets
* p<0.05, ** p<0.01, *** p<0.001
[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)
[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

Figure 26 Forest plot of the fully adjusted multivariate association between depression and unemployment by sex and area deprivation



5.7.7 Association between each chronic condition and unemployment after excluding children with special educational need and adjusting for absenteeism

Adjusting for the number of days absent from school per census year attenuated the associations between epilepsy and depression and unemployment but they remained significant. However this adjustment had no notable impact on the magnitude of the association between ADHD and unemployment. Excluding children with SEN attenuated the associations between ADHD and epilepsy and unemployment but they remained significant. However this had no notable impact on the magnitude of the association between depression and unemployment. The association between diabetes and unemployment was not further investigated by adjusting for absenteeism or excluding children with SEN because the fully adjusted multivariate model in 5.7.2 was not significant. This was also true for the association between asthma and unemployment (5.7.3).

There was a significant association between epilepsy and unemployment after adjusting for sociodemographic and maternity (OR 1.99, 95% CI 1.75-2.25) confounders whereby children with epilepsy had 2-fold odds of unemployment. However, after adjusting for absenteeism, children with epilepsy experienced a 70% increased odds of unemployment (OR 1.70, 95% CI 1.49-1.55). Furthermore, after excluding children with SEN, children with epilepsy experienced a 49% increased odds of unemployment (OR 1.49, 95% CI 1.24-1.80).

There was also a significant association between ADHD and unemployment after adjusting for sociodemographic and maternity (OR 1.42, 95% CI 1.28-1.57) confounders whereby children with ADHD experienced 42% increased odds of unemployment. However, after excluding children with SEN, children with ADHD experienced 30% increased odds of unemployment (OR 1.30, 95% CI 1.10-1.55). Adjusting for absenteeism did not notably impact the magnitude of the association between ADHD and unemployment and the increase in odds for children with ADHD remained around 40% (OR 1.44, 95% CI 1.29-1.60).

There was a significant association between depression and unemployment after adjusting for sociodemographic and maternity (OR 1.98, 95% CI 1.80-2.18) confounders whereby children with depression experienced 2-fold odds of unemployment. The increase in odds remained 2-fold after excluding children with SEN from the analysis (OR 2.02, 95% CI

1.80-2.26). However, after adjusting for absenteeism, children with depression experienced 42% increased odds of unemployment (OR 1.42, 95% CI 1.28-1.57). The results of each of the original analyses versus each of the analyses excluding children with a SEN and adjusted for absenteeism are shown in Table 58 below.

Table 58 Association between chronic diseases and unemployment in children after adjusting for absenteeism and excluding children with special educational needs

	multivariate model[1]		adjusted for absenteeism[2a]		excluding children with SEN[2b]	
	OR	95% CI	OR	95% CI	OR	95% CI
Epilepsy	1.99***	1.75-2.25	1.70***	1.49-1.95	1.49***	1.24-1.80
ADHD	1.42***	1.28-1.57	1.44***	1.29-1.60	1.30**	1.10-1.55
Depression	1.98***	1.80-2.18	1.42***	1.28-1.57	2.02***	1.80-2.26

Exponentiated coefficients; 95% confidence intervals in brackets
** p<0.05, ** p<0.01, *** p<0.001*

[1]adjusted for sex, age, SIMD deprivation quintile, ethnic group, maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

[2a] mulivariate model repeated additionally adjusting for number of days absent from school per census year

[2b] mulivariate model repeated after excluding children with special educational needs

5.8 Association between each chronic condition and each educational outcome after adjusting for other comorbid chronic conditions

Of the 766,244 children attending school between 2009 and 2013, 61,897 (8.1%) had one of the five chronic conditions investigated. However, a further 2,501 (0.33%) pupils had two comorbid conditions and 132 (0.02%) pupils had three comorbid conditions. Therefore, given that some children experienced more than one condition, the final fully adjusted models for all of the educational outcomes were run again including all conditions together. The first two columns in Table 59 present incidence rate ratios and 95% confidence intervals from generalised estimating equations with a negative binomial distribution and log link adjusted for sociodemographic and maternity confounders. These models investigated associations between each chronic condition and number of days absent and number of exclusions. The third and fourth columns in Table 59 present the same information after adjusting each condition-specific model for presence of any of the other chronic conditions identified in this thesis.

The first two columns in Table 60 present odds ratios and 95% confidence intervals from models, adjusted for sociodemographic and maternity confounders, investigating associations between each chronic condition and SEN, unemployment and academic attainment. SEN was investigated using generalised estimating equations with a binomial distribution and logit link. Unemployment and attainment were investigated using binomial and ordinal logistic regression respectively. The third and fourth columns in Table 60 present the same information after adjusting each condition-specific model for presence of any of the other chronic conditions identified in this thesis.

Adjusting for the presence of other chronic conditions had little impact on the effect size of the associations between any of the chronic conditions and number of days absent. This was also true for the associations between both asthma and ADHD and number of exclusions. However, the association between depression and number of exclusions was attenuated but remained significant after adjustment for the presence of other chronic conditions. Depressed children had a 65% increased risk of exclusion compared to children who were not depressed; however they experienced a 48% increased risk after further adjustment for the presence of other conditions.

Adjusting for other conditions did not notably attenuate the associations between diabetes and SEN, asthma and SEN, asthma and attainment or ADHD and unemployment. However, the association between ADHD and attainment was attenuated but remained significant. The most notable impact of adjusting for other conditions was observed for epilepsy and depression. Associations of these conditions with SEN, unemployment and academic attainment were all attenuated but remained significant after further adjustment for presence of other chronic conditions.

Table 59 Association between each chronic condition and absences and exclusions after adjusting for presence of other comorbid chronic conditions

	multivariate model[1]		adjusted for comorbid conditions[2]	
	IRR	95% CI	IRR	95% CI
Absence				
Diabetes	1.34***	1.30-1.39	1.33***	1.28-1.37
Asthma	1.25***	1.24-1.26	1.24***	1.23-1.26
Epilepsy	1.50***	1.45-1.55	1.43***	1.38-1.48
ADHD	1.18***	1.16-1.21	1.16***	1.14-1.19
Depression	1.95***	1.90-2.00	1.90***	1.85-1.95
Exclusion				
Asthma	0.91**	0.85-0.97	0.89***	0.84-0.95
ADHD	5.82***	5.47-6.18	5.79***	5.45-6.16
Depression	1.65***	1.45-1.88	1.48***	1.29-1.69

Exponentiated coefficients; 95% confidence intervals in brackets
** p<0.05, ** p<0.01, *** p<0.001*

[1]adjusted for sex, age, SIMD deprivation quintile, ethnic group, maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

[2] also adjusted for diabetes, asthma, epilepsy, ADHD and depression (as appropriate)

Table 60 Association between each chronic condition and special educational need, unemployment (NEET) and academic attainment after adjusting for presence of other comorbid chronic conditions

		multivariate model[1]		adjusted for comorbid conditions[2]	
		OR	95% CI	OR	95% CI
SEN					
Diabetes		2.45***	2.26-2.66	2.46***	2.26-2.68
Asthma		1.28***	1.25-1.32	1.27***	1.23-1.30
Epilepsy		10.11***	9.51-10.75	9.60***	9.02-10.23
ADHD		8.77***	8.40-9.15	8.62***	8.26-9.00
Depression		2.24***	2.10-2.39	1.77***	1.65-1.90
NEET					
Epilepsy		1.99***	1.75-2.25	1.82***	1.60-2.07
ADHD		1.42***	1.28-1.57	1.39***	1.25-1.53
Depression		1.98***	1.80-2.18	1.88***	1.71-2.08
Attainment					
Asthma		1.11***	1.06-1.16	1.09***	1.04-1.14
Epilepsy	general, basic or low[3]	1.98***	1.68-2.32	1.73***	1.47-2.04
	basic or low[4]	2.48***	2.11-2.93	2.16***	1.83-2.55
	low[5]	4.07***	3.26-5.08	3.43***	2.74-4.29
ADHD		3.64***	3.26-4.07	3.50***	3.14-3.91
Depression	general, basic or low[3]	1.96***	1.77-2.17	1.86***	1.68-2.05
	basic or low[4]	2.14***	1.92-2.39	1.99***	1.78-2.23
	low[5]	3.44***	2.89-4.09	2.98***	2.50-3.56

Exponentiated coefficients; 95% confidence intervals in brackets
** p<0.05, ** p<0.01, *** p<0.001*

[1]adjusted for sex, age, SIMD deprivation quintile, ethnic group, maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

[2] also adjusted for diabetes, asthma, epilepsy, ADHD and depression (as appropriate)

[3]reference category of high attainment

[4]reference category of high or general attainment

[5]reference category of high, general or basic attainment

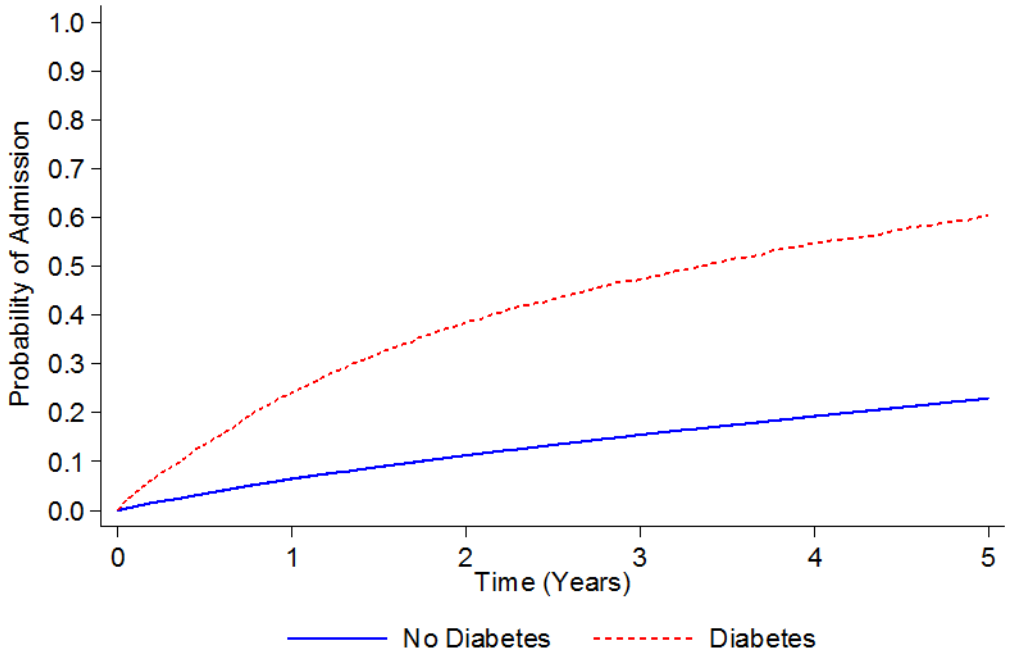
5.9 All-cause and cause-specific hospital admissions

5.9.1 Association between diabetes and hospital admission

5.9.1.1 All-cause hospital admission

197 pupils who received ineligible diabetes drugs (section 4.3.2.2.1) were excluded. Therefore 766,047 pupils were followed up to assess the impact of diabetes on risk of all-cause hospital admission. The follow up period ranged from 1 year to 5 years and the mean follow up time was 4.34 years. 157,294 pupils (20.5%) were admitted to hospital at least once over the follow up period and there were 305,580 admissions in total. The probability of being admitted to hospital for any cause was significantly greater for pupils with diabetes compared to those without diabetes at all points during follow up. The probability of being admitted to hospital within the 5-year follow-up period was around 60% for children with diabetes compared to 20% for children without diabetes (Figure 27).

Figure 27 Kaplan-Meier survival curve: Diabetes and all-cause hospital admission



Diabetes was significantly associated with increased risk of all-cause hospital admission on univariate Cox regression analysis (HR 3.84, 95% CI 3.67-4.03) and following adjustment for sociodemographic (HR 3.89, 95% CI 3.71-4.07) and maternity (HR 3.97, 95% CI 3.79-4.16) confounders. However, the model did not satisfy the assumption of proportional hazards ($p<0.001$) and Poisson piecewise regression was used to more accurately model the variation in the hazard ratio over the follow up period. Table 61 presents univariate and

multivariate incidence rate ratios within each specific year of follow up. Figure 28 presents the fully adjusted incidence rate ratios from the Poisson piecewise model (black solid circles) alongside the static multivariate hazard ratio derived from the Cox proportional hazards model (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines). The Cox model demonstrated that, over the full follow up period, diabetic children had 3.9 times greater risk of hospital admission compared to non-diabetic pupils. However, the Poisson piecewise model demonstrated that the elevated risk of hospital admission experienced by diabetic pupils compared to non-diabetic pupils actually decreased in magnitude with increasing follow-up time. Diabetic children experienced a 4.4 fold risk of admission within the first year of follow-up; however their increase in risk was 2.9 fold across the final year of follow-up. Therefore, as children moved through the follow-up period, the relative impact of diabetes on risk of admission became weaker. A possible explanation for this observation is that it may take some time to stabilise children on insulin and for them, and their parents, to become confident in managing the condition. However, thereafter most diabetic children can avoid future symptoms such as hypoglycaemic episodes.

Table 61 Diabetes and risk of all-cause hospital admission modelled as time from diagnosis until first admission

	Univariate model		Multivariate[1]		Multivariate[2]	
	IRR	95% CI	IRR	95% CI	IRR	95% CI
0 - 1 years	4.31***	4.02-4.62	4.34***	4.05-4.66	4.43***	4.13-4.75
1 - 2 years	4.11***	3.74-4.51	4.16***	3.79-4.57	4.23***	3.85-4.65
2 - 3 years	3.32***	2.93-3.76	3.37***	2.97-3.82	3.47***	3.06-3.93
3 - 4 years	3.29***	2.85-3.82	3.35***	2.90-3.88	3.41***	2.94-3.95
4 - 5 years	2.80***	2.34-3.36	2.86***	2.39-3.43	2.92***	2.44-3.50

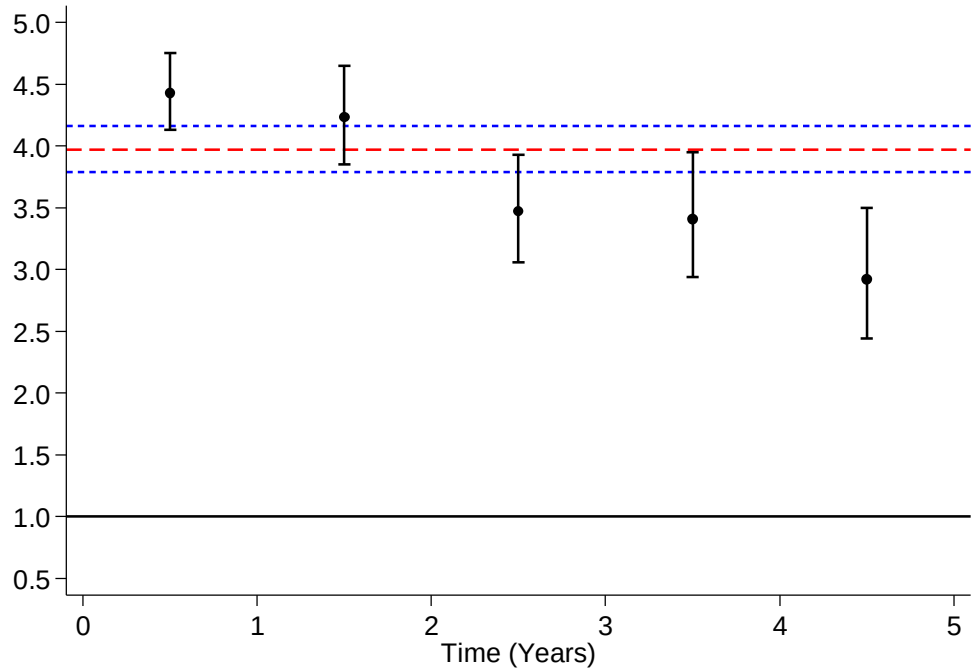
Exponentiated coefficients; 95% confidence intervals in brackets

* p<0.05, ** p<0.01, *** p<0.001

[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)

[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

Figure 28 Diabetes and risk of all-cause hospital admission modelled as time from diagnosis until first admission



The risk of hospital admission associated with diabetes also varied depending on pupil age at time of admission. Table 62 presents univariate and multivariate incidence rate ratios from a Poisson piecewise model for different age categories at admission. Figure 29 presents the fully adjusted incidence rate ratios from the Poisson piecewise model (black solid circles) alongside the static multivariate hazard ratio derived from the Cox proportional hazards model (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines). Diabetic children of all ages experienced an increased risk of hospital admission compared to non-diabetic children; however, the association between diabetes and risk of hospital admission was stronger for children aged between 9 and 14 years of age. This could be due to reported differing patterns in treatment adherence whereby pupils in their early teens are harder to manage and more likely to for example miss insulin injections [272].

Table 62 Diabetes and risk of all-cause hospital admission modelled as age at first admission

	Univariate model		Multivariate[1]		Multivariate[2]	
	IRR	95% CI	IRR	95% CI	IRR	95% CI
4 - 6 years	4.35***	3.71-5.11	4.45***	3.80-5.23	4.42***	3.75-5.22
7 - 8 years	4.00***	3.37-4.74	4.04***	3.41-4.79	4.12***	3.46-4.90
9 - 10 years	5.01***	4.36-5.75	5.05***	4.40-5.80	5.20***	4.53-5.97
11 - 12 years	5.13***	4.56-5.77	5.17***	4.60-5.82	5.25***	4.67-5.91
13 - 14 years	4.99***	4.51-5.53	5.03***	4.55-5.57	5.14***	4.65-5.70
15 - 16 years	3.68***	3.30-4.09	3.70***	3.32-4.13	3.79***	3.40-4.22
17 - 23 years	2.77***	2.48-3.10	2.83***	2.53-3.16	2.89***	2.59-3.24

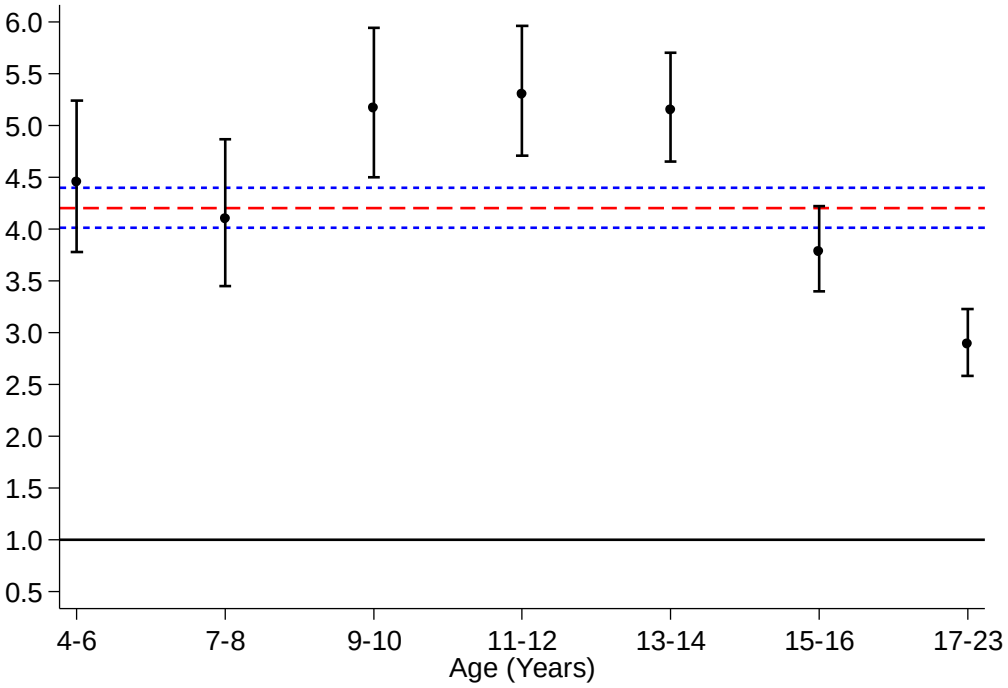
Exponentiated coefficients; 95% confidence intervals in brackets

* p<0.05, ** p<0.01, *** p<0.001

[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)

[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

Figure 29 Diabetes and risk of all-cause hospital admission modelled as age at first admission



There was a significant interaction between diabetes and gender (p<0.001) whereby the association between diabetes and risk of admission was consistently stronger for girls than for boys during all points of follow up. This was despite similar absolute rates of hospital admission among unaffected children (76,410 [20.4%] vs. 79,090 [20.4%] among girls and boys respectively). Among children with diabetes, admissions were higher among girls than boys (943 [58.6%] vs. 851 [49.5%]). The overall patterns observed within the main analyses remained within each group, whereby the association between diabetes and risk of hospital admission became weaker throughout the follow up period and became stronger between the ages of 9 and 14. Figure 30 and Figure 31 show the fully adjusted gender specific incidence rate ratios from the Poisson piecewise models alongside the static multivariate gender specific hazard ratios derived from the corresponding proportional hazards models. Confidence intervals are only presented for the Poisson piecewise models.

Figure 30 Diabetes and gender specific risk of all-cause hospital admission modelled as time from diagnosis until first admission (blue=boys; red=girls)

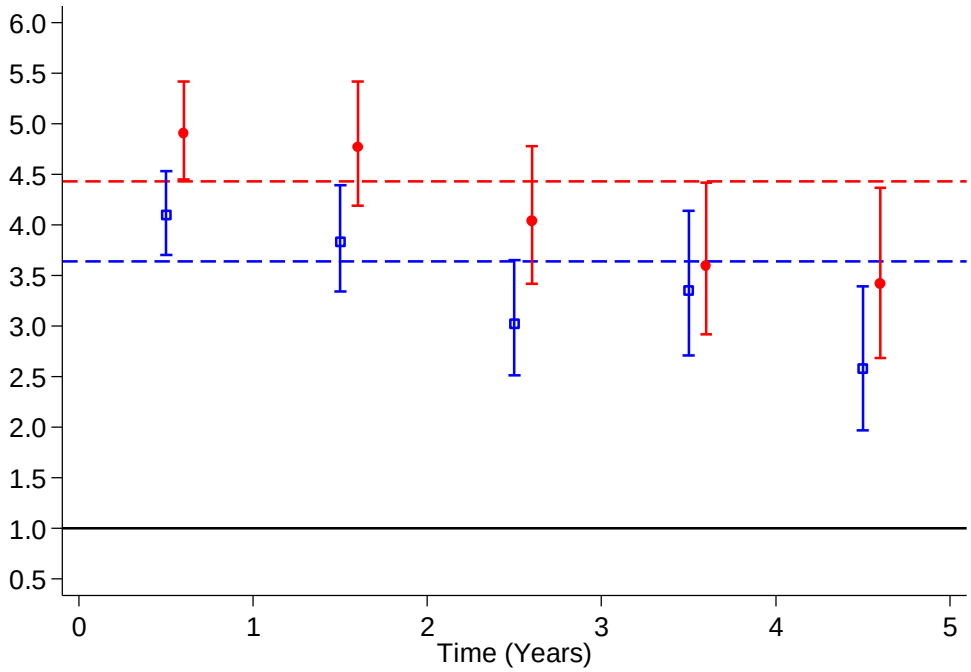
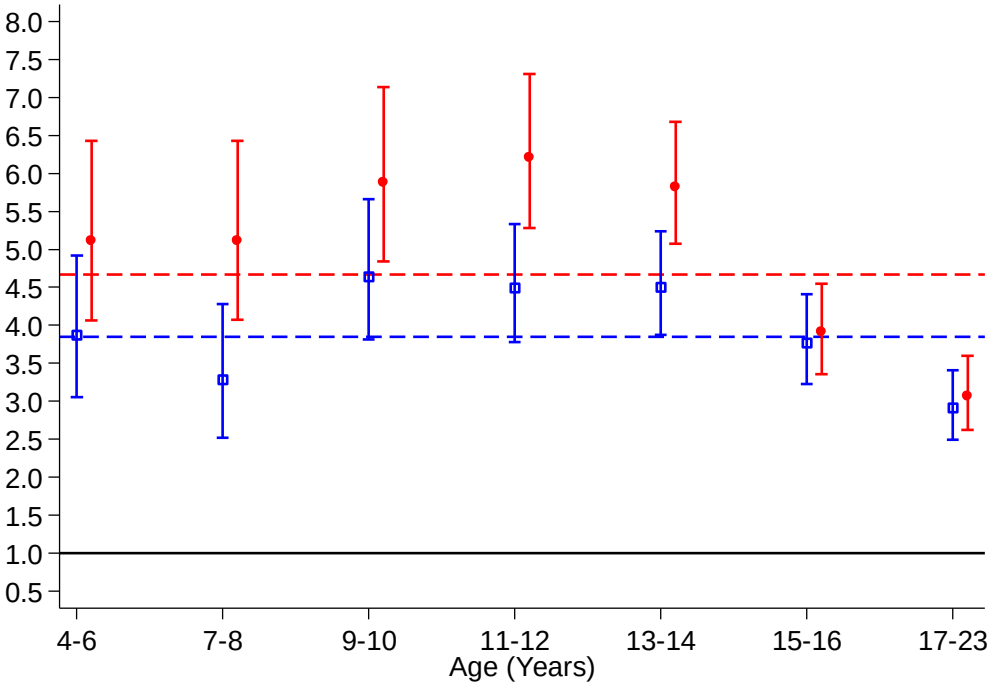


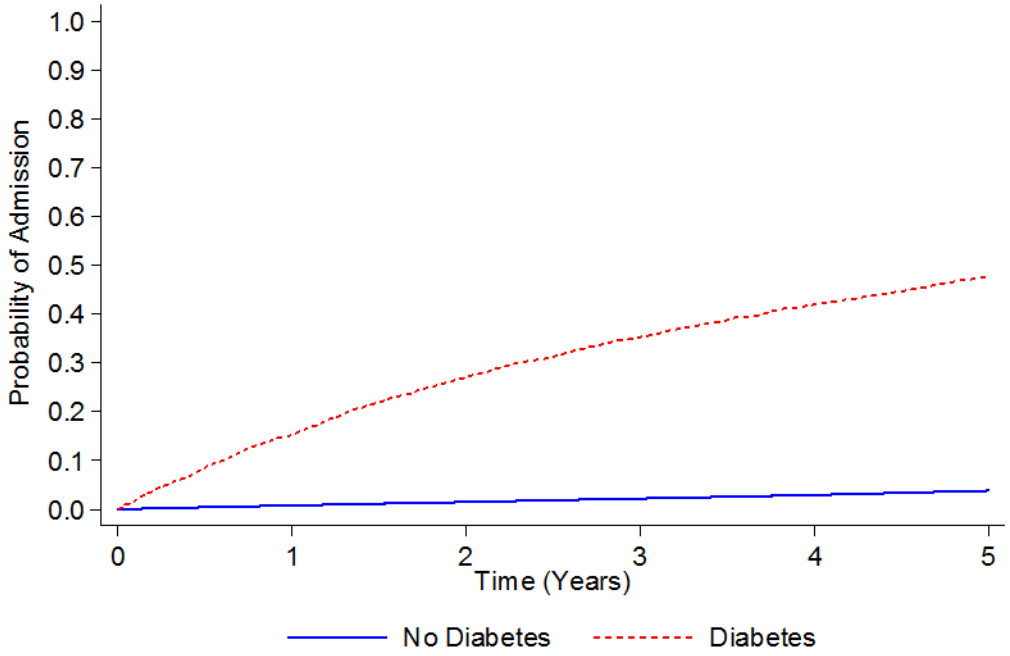
Figure 31 Diabetes and gender specific risk of all-cause hospital admission modelled by age at first admission (blue=boys; red=girls)



5.9.1.2 Hospital admission related to endocrine, nutritional or metabolic diseases or abnormal clinical or lab findings not elsewhere classified

197 pupils who received ineligible diabetes drugs (section 4.3.2.2.1) were excluded. Therefore 766,047 pupils were followed up to assess the impact of diabetes on risk of hospital admission related to endocrine, nutritional or metabolic diseases (ICD10 Chapter E00-E99) or abnormal clinical or laboratory findings not elsewhere classified (ICD10 Chapter R00-R99) as these were two of the main causes of admission for children with diabetes. Admissions with these ICD10 codes were likely to relate to stabilisation on insulin or acute complications of diabetes such as hypoglycaemia or diabetic ketoacidosis (DKA). The follow up period ranged from 1 year to 5 years and the mean follow up time was 4.34 years. 26,631 pupils (3.5%) were admitted to hospital at least once during the follow up period for one of these causes and there were 42,461 admissions in total. The probability of having an admission with one of the ICD10 codes within 5 years was around 50% for children with diabetes compared to less than 5% for children without diabetes (Figure 32).

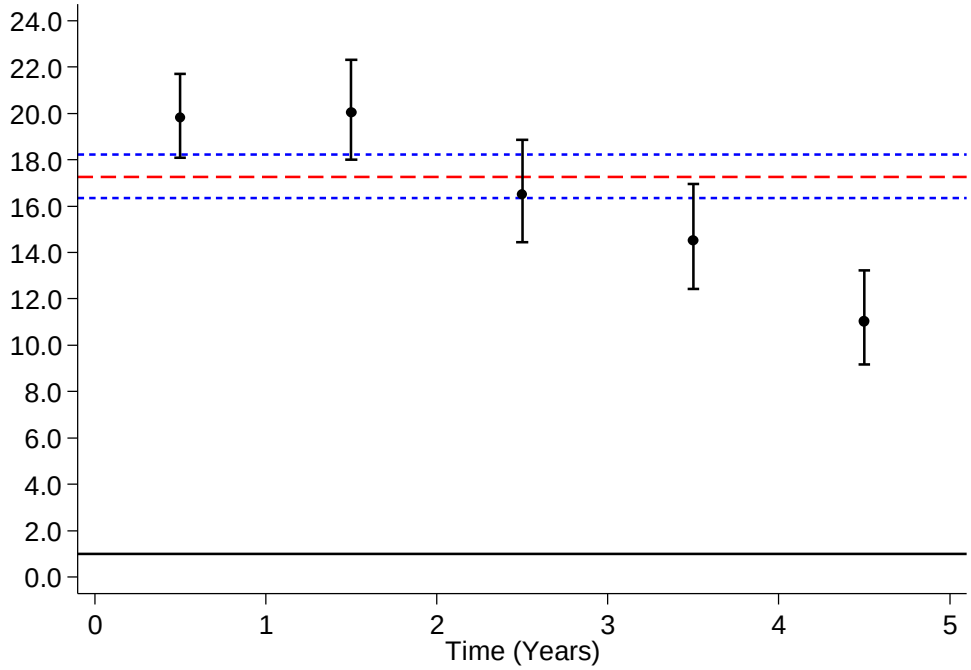
Figure 32 Kaplan-Meier survival curve: Diabetes and hospital admission related to endocrine, nutritional or metabolic diseases or abnormal clinical or lab findings not elsewhere classified



Diabetes was significantly associated with increased risk of hospital admission related to endocrine, nutritional or metabolic diseases or abnormal clinical or lab findings not elsewhere classified on univariate Cox regression analysis (HR 16.75, 95% CI 15.87-17.68) and following adjustment for sociodemographic (HR 17.06, 95% CI 16.15-18.01) and maternity confounders (HR 17.32, 95% CI 16.40-18.29). However, the model did not satisfy the assumption of proportional hazards ($p<0.001$) and Poisson piecewise regression was used to more accurately model variation in the hazard ratio over the follow up period. Figure 33 presents the fully adjusted incidence rate ratios from the Poisson piecewise model (black solid circles) alongside the static multivariate hazard ratio derived from the Cox proportional hazards model (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines). The Cox model demonstrated that, over the full follow up period, diabetic children experienced a 17-fold risk of hospital admission compared to non-diabetic pupils. However, the Poisson piecewise model demonstrated that the elevated risk of hospital admission experienced by diabetic pupils compared to non-diabetic pupils actually decreased in magnitude with increasing follow-up time. Diabetic children experienced a 20-fold risk of admission within the first year of follow-up; however their increase in risk fell to 11-fold across the final year of follow-up. Therefore,

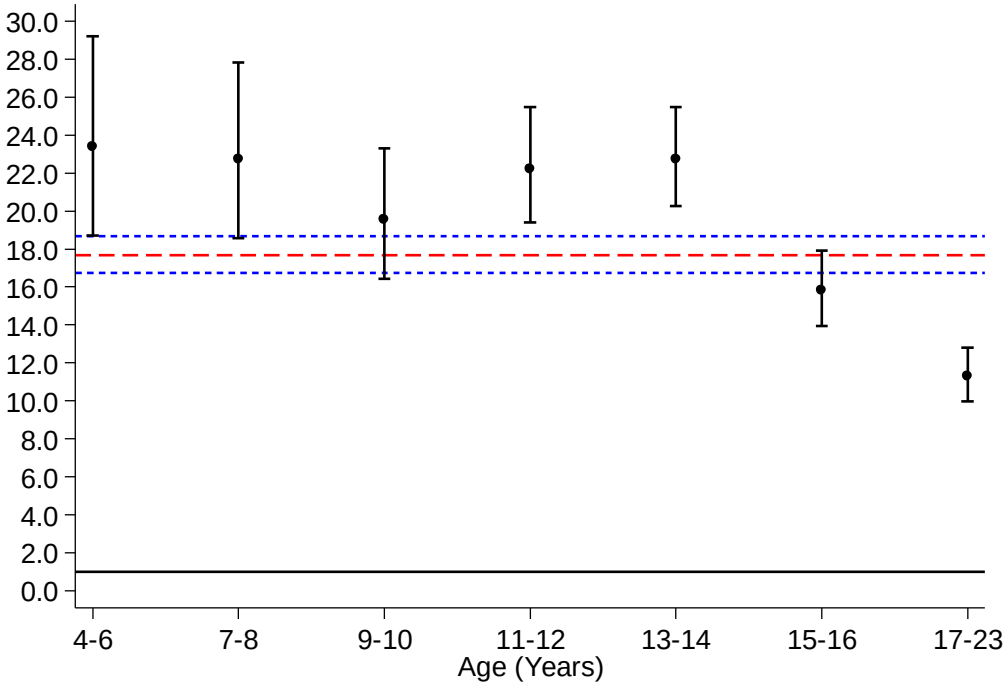
as children moved through the follow-up period, the relative impact of diabetes on risk of hospital admission became weaker. Again, this may potentially be explained because it takes time to stabilise children on insulin and for them, and their parents, to become confident in managing their condition.

Figure 33 Diabetes and risk of hospital admission related to endocrine, nutritional or metabolic diseases or abnormal clinical or lab findings not elsewhere classified modelled as time from diagnosis until first admission



The risk of admission associated with diabetes also varied depending on pupil age at time of admission. Figure 34 presents the fully adjusted incidence rate ratios from the Poisson piecewise model (black solid circles) alongside the static multivariate hazard ratio derived from the Cox proportional hazards model (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines). Diabetic children of all ages experienced an increased risk of hospital admission compared to non-diabetic children; however, similar to all-cause hospital admissions, the association was again weaker for children in their late teens suggesting that they may be better able to manage their condition than younger adolescents. There was no significant univariate interaction between diabetes and gender (p=0.272).

Figure 34 Diabetes and risk of hospital admission related to endocrine, nutritional or metabolic diseases or abnormal clinical or lab findings not elsewhere classified modelled by age at first admission

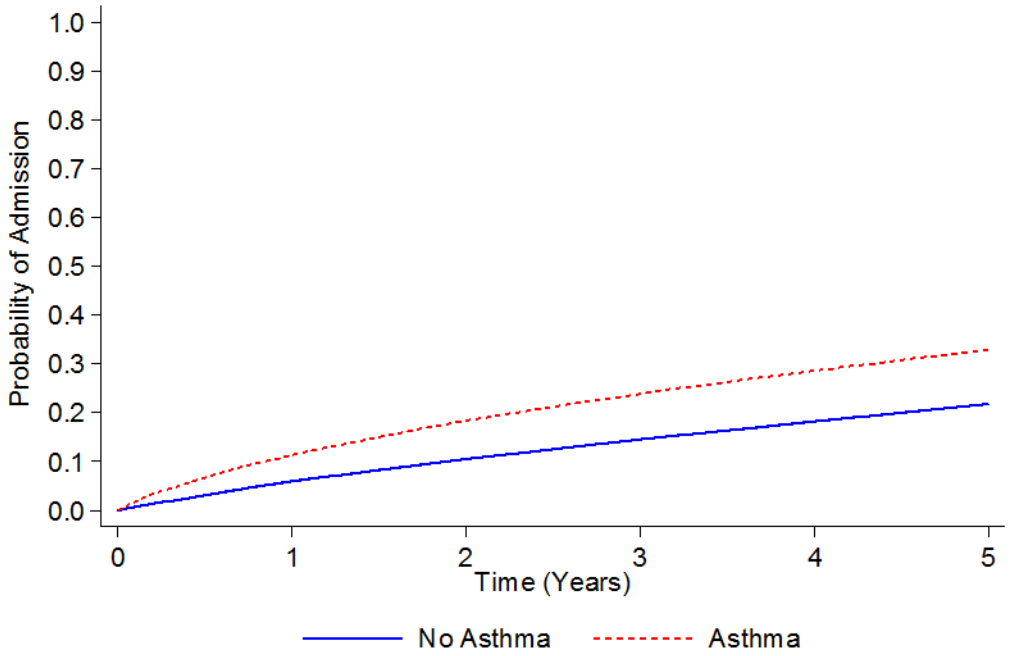


5.9.2 Association between asthma and hospital admission

5.9.2.1 All-cause hospital admission

82,528 pupils who received ineligible asthma drugs (section 4.3.2.2.2) were excluded. Therefore 683,716 pupils were followed up to assess the impact of asthma on risk of all-cause admission. The follow up period ranged from 1 year to 5 years and the mean follow up time was 4.29 years. 135,482 pupils (19.8%) were admitted to hospital at least once during the follow up period and there were 260,705 admissions in total. The probability of having an all-cause admission was significantly higher for pupils with asthma compared to those without asthma at all points over follow up. The probability of being admitted to hospital for any cause within 5 years was around 30% for children with asthma compared to 20% for children without asthma (Figure 35).

Figure 35 Kaplan-Meier survival curve: Asthma and all-cause hospital admission



Asthma was significantly associated with increased risk of hospital admission on univariate Cox regression analysis (HR 1.74, 95% CI 1.71-1.77) and following adjustment for sociodemographic (HR 1.71, 95% CI 1.68-1.74) and maternity (HR 1.69, 95% CI 1.66-1.72) confounders. However, the model did not satisfy the assumption of proportional hazards ($p<0.001$) and Poisson piecewise regression was used to more accurately model the variation in the hazard ratio over the follow up period. Table 63 presents the univariate and multivariate incidence rate ratios within each specific year of follow up. Figure 36 presents the fully adjusted incidence rate ratios from the Poisson piecewise model (black solid circles) alongside the static multivariate hazard ratio derived from the Cox proportional hazards model (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines). The Cox model demonstrated that, over the full follow up period, asthmatic children had a 70% increased risk of hospital admission compared to non-asthmatic pupils. However, the Poisson piecewise model demonstrated that the elevated risk of hospital admission experienced by asthmatic pupils compared to non-asthmatic pupils actually decreased in magnitude with increasing follow-up time. Asthmatic children experienced a 98% increased risk of admission within the first year of follow-up which reduced to 38% increased risk across the final year of follow-up. Therefore, as children moved through the follow-up period, the relative impact of asthma on risk of all-cause admission became weaker. A possible explanation may be that

asthmatic pupils become stabilised on their medication and are more able to manage their condition with time resulting in less risk of admission to hospital.

Table 63 Asthma and risk of all-cause hospital admission modelled as time from diagnosis until first admission

	Univariate model		Multivariate[1]		Multivariate[2]	
	IRR	95% CI	IRR	95% CI	IRR	95% CI
0 - 1 years	2.04***	1.98-2.10	2.01***	1.95-2.06	1.98***	1.93-2.04
1 - 2 years	1.73***	1.67-1.80	1.71***	1.64-1.77	1.69***	1.63-1.76
2 - 3 years	1.53***	1.47-1.60	1.51***	1.44-1.58	1.49***	1.42-1.56
3 - 4 years	1.51***	1.43-1.59	1.48***	1.41-1.56	1.47***	1.39-1.55
4 - 5 years	1.41***	1.33-1.50	1.39***	1.30-1.47	1.38***	1.30-1.47

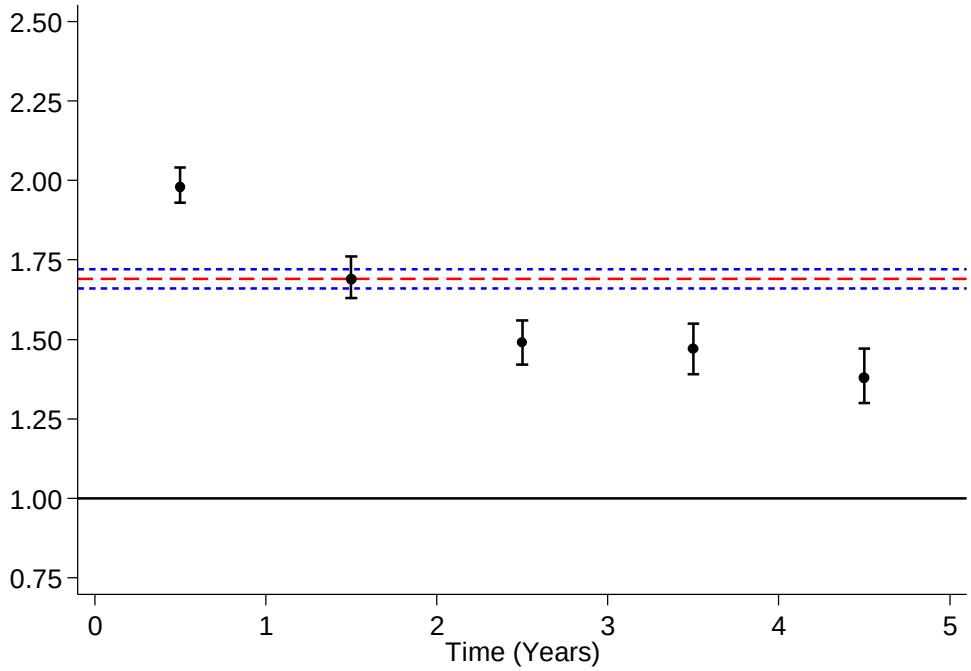
Exponentiated coefficients; 95% confidence intervals in brackets

* p<0.05, ** p<0.01, *** p<0.001

[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)

[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

Figure 36 Asthma and risk of all-cause hospital admission modelled as time from diagnosis until first admission



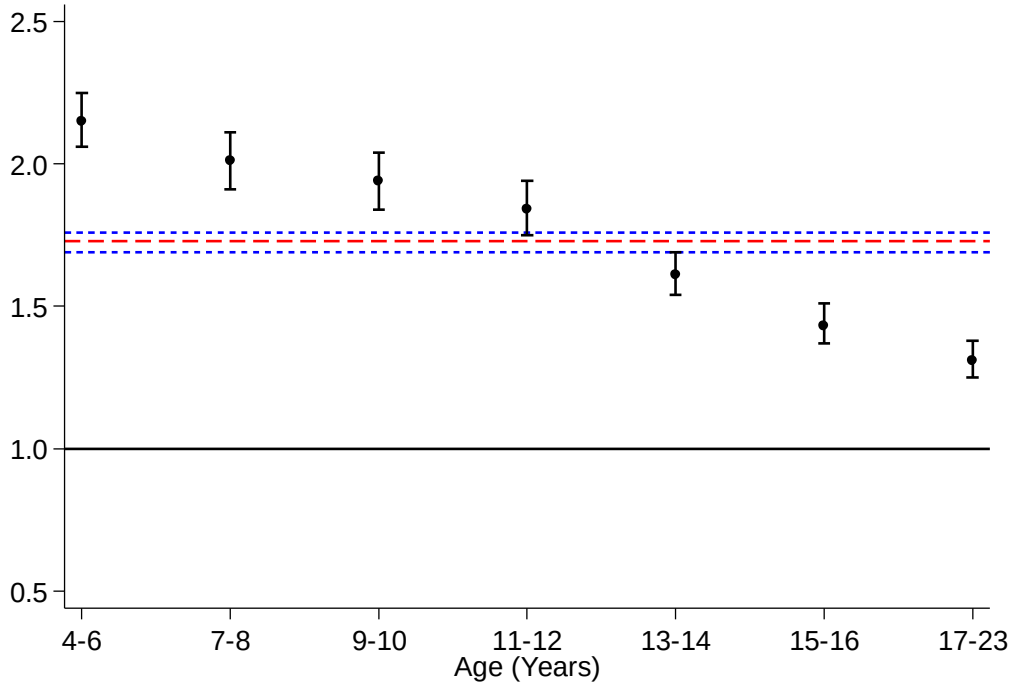
The risk of all-cause hospital admission associated with asthma also varied depending on pupil age at time of admission. Table 64 presents univariate and multivariate incidence rate ratios from a Poisson piecewise model within different age categories at admission. Figure 37 presents the fully adjusted incidence rate ratios from the Poisson piecewise model (black solid circles) alongside the static multivariate hazard ratio derived from the Cox proportional hazards model (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines). Asthmatic children of all ages experienced an increased risk of all-cause hospital admission compared to non-asthmatic children; however, there was a distinct trend whereby the association between asthma and risk of hospital admission was stronger in younger children. The magnitude of the excess risk decreased from 2.15 fold in children aged 4-6 years to 1.31 fold in children aged 17-23 years. This observation may potentially be due to older asthmatic children having a greater understanding of their condition and its limitations and risks compared to younger children who as a result are impacted more severely in terms of risk of admission. There was no significant interaction between asthma and gender (p=0.405).

Table 64 Asthma and risk of all-cause hospital admission modelled as age at first admission

	Univariate model		Multivariate[1]		Multivariate[2]	
	IRR	95% CI	IRR	95% CI	IRR	95% CI
4 - 6 years	2.25***	2.16-2.35	2.20***	2.11-2.30	2.15***	2.06-2.25
7 - 8 years	2.07***	1.98-2.18	2.03***	1.94-2.13	2.01***	1.91-2.11
9 - 10 years	2.00***	1.90-2.10	1.96***	1.86-2.06	1.94***	1.84-2.04
11 - 12 years	1.88***	1.79-1.98	1.85***	1.76-1.94	1.84***	1.75-1.94
13 - 14 years	1.64***	1.57-1.73	1.62***	1.54-1.70	1.61***	1.54-1.69
15 - 16 years	1.46***	1.39-1.54	1.44***	1.37-1.51	1.43***	1.37-1.51
17 - 23 years	1.34***	1.27-1.40	1.32***	1.25-1.38	1.31***	1.25-1.38

Exponentiated coefficients; 95% confidence intervals in brackets
* p<0.05, ** p<0.01, *** p<0.001
[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)
[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

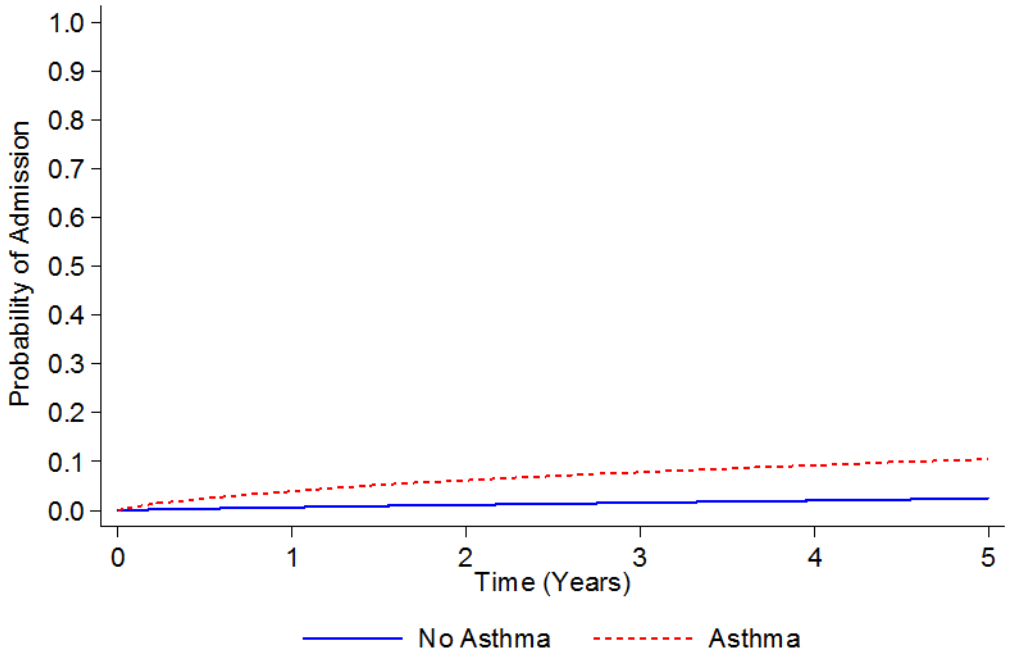
Figure 37 Asthma and risk of all-cause hospital admission modelled as age at first admission



5.9.2.2 Hospital admission related to the respiratory system

82,528 pupils who received ineligible asthma drugs (section 4.3.2.2.2) were excluded. Therefore 683,716 pupils were followed up to assess the impact of asthma on risk of respiratory related hospital admission as this was the main cause of admission among children with asthma. The follow up period ranged from 1 year to 5 years and the mean follow up time was 4.29 years. 17,721 pupils (2.6%) were admitted to hospital at least once during the follow up period and there were 25,733 admissions in total. The probability of having a respiratory related admission was significantly higher for pupils with asthma compared to those without asthma at all points during follow up. The probability of being admitted to hospital for a respiratory admission within 5 years was around 10% for children with asthma compared to less than 5% for children without asthma (Figure 38).

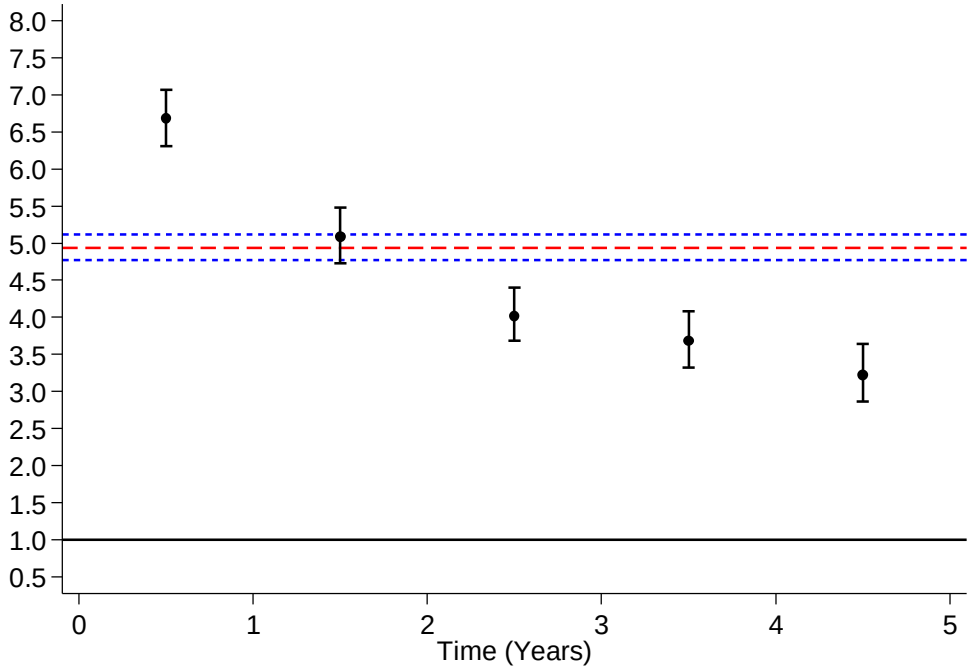
Figure 38 Kaplan-Meier survival curve: Asthma and respiratory related hospital admission



Asthma was significantly associated with increased risk of a respiratory related hospital admission on univariate Cox regression analysis (HR 4.99, 95% CI 4.82-5.17) and following adjustment for sociodemographic (HR 5.00, 95% CI 4.83-5.18) and maternity (HR 4.94, 95% CI 4.77-5.12) confounders. However, the model did not satisfy the assumption of proportional hazards ($p<0.001$) and Poisson piecewise regression was used to more accurately model the hazard ratios over the follow up period. Figure 39 presents the fully adjusted incidence rate ratios from the Poisson piecewise model (black solid circles) alongside the static multivariate hazard ratio derived from the Cox proportional hazards model (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines). The Cox model demonstrated that, over the full follow up period, asthmatic children experienced an almost 5-fold risk of respiratory hospital admission compared to non-asthmatic pupils. However, the Poisson piecewise model demonstrated that the elevated risk of respiratory hospital admission experienced by asthmatic pupils compared to non-asthmatic pupils actually decreased in magnitude with increasing follow-up time. Asthmatic children experienced a 6.7-fold risk of respiratory admission within the first year of follow-up which reduced to 3.2-fold across the final year of follow-up. Therefore, as children moved through the follow-up period, the relative impact of asthma on risk of respiratory admission became weaker. This pattern was similar

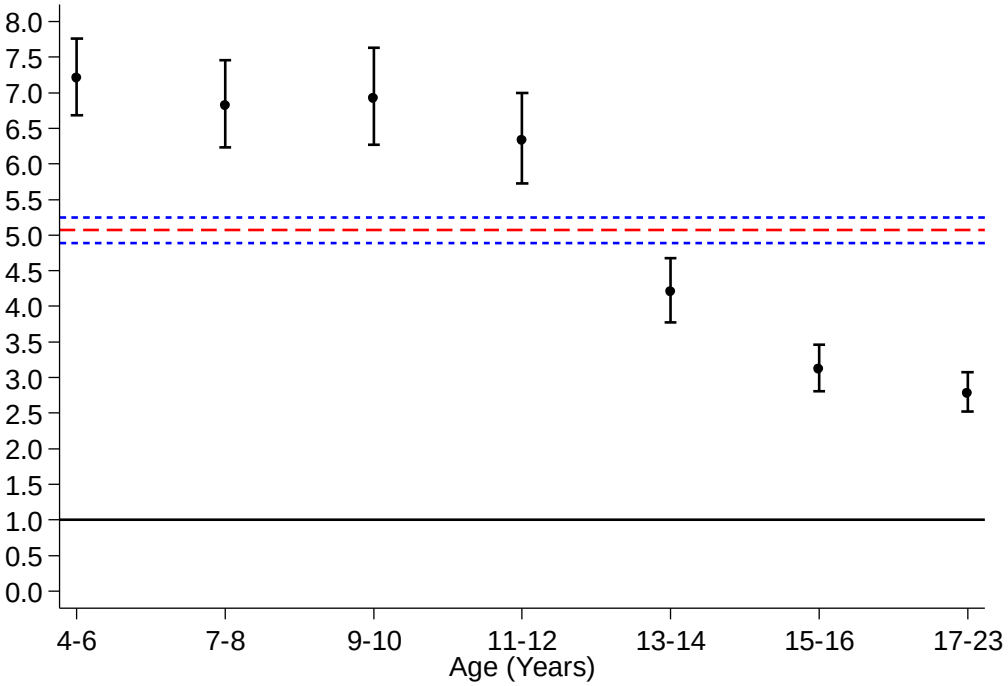
to that observed for asthma and risk of all-cause hospital admission and may again be attributed to children becoming stabilised on medication over time.

Figure 39 Asthma and risk of respiratory related hospital admission modelled as time from diagnosis until first admission



The risk of respiratory admission associated with asthma also varied depending on pupil age at time of admission. Figure 40 presents the fully adjusted incidence rate ratios from the Poisson piecewise model (black solid circles) alongside the static multivariate hazard ratio derived from the Cox proportional hazards model (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines). Asthmatic children of all ages experienced an increased risk of respiratory hospital admission compared to non-asthmatic children; however, there was a trend whereby the association between asthma and risk of respiratory hospital admission became weaker for older children. The magnitude of the excess risk decreased from 7.24 fold in children aged 4-6 years to 2.75 fold in children aged 17-23 years. This pattern was similar to that observed for asthma and risk of all-cause hospital admission.

Figure 40 Asthma and risk of respiratory related hospital admission modelled as age at first admission



There was a significant interaction between asthma and gender ($p<0.001$) whereby the association between asthma and risk of respiratory admission was stronger for boys than girls during all points of follow up. Girls however had higher rates of respiratory admission among unaffected children (7,852 [2.5%] vs. 5,737 [1.8%] among girls and boys respectively). Among children with asthma, admissions were still higher among girls than boys (1,831 [9.6%] vs. 2,301 [8.6%]). Figure 41 and Figure 42 show the fully adjusted gender specific incidence rate ratios from the Poisson piecewise models alongside the static multivariate gender specific hazard ratios derived from the corresponding proportional hazards models. Confidence intervals are only presented for the Poisson piecewise models. The same patterns observed within the main analyses were observed within both gender groups namely that the excess risk of respiratory admission associated with asthma decreased over follow up time and with increasing age.

Figure 41 Asthma and gender specific risk of respiratory related hospital admission modelled as time from diagnosis until first admission (blue=boys; red=girls)

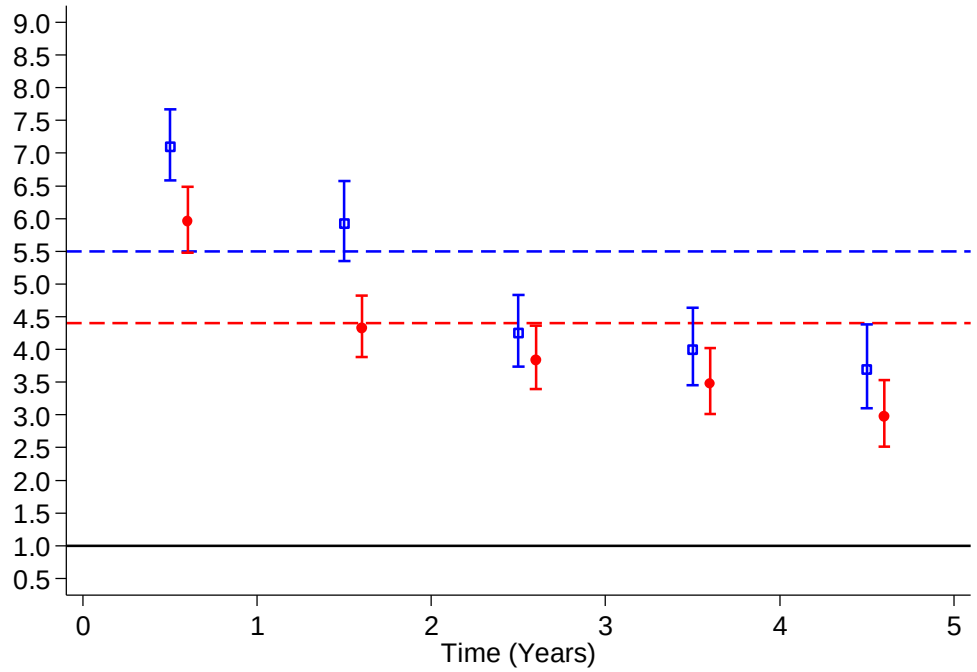
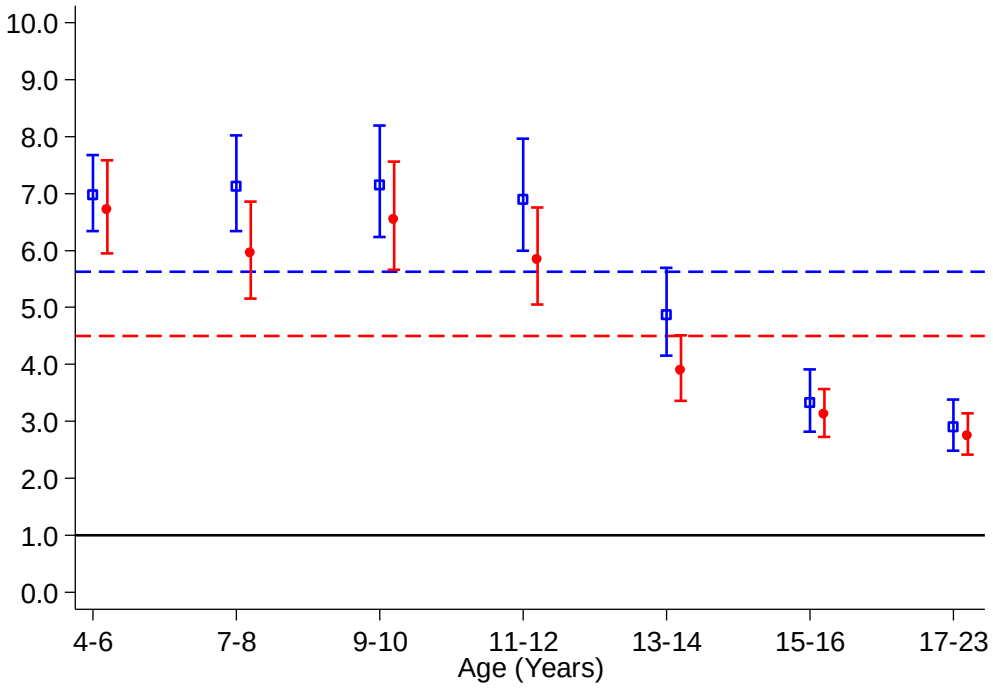


Figure 42 Asthma and gender specific risk of respiratory related hospital admission modelled by age at first admission (blue=boys; red=girls)

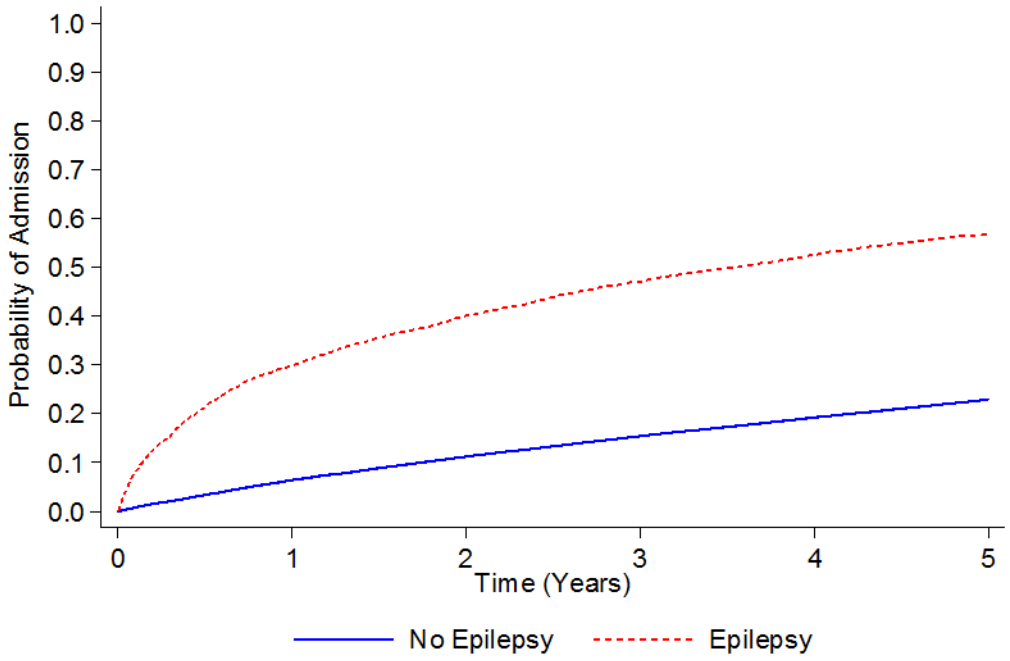


5.9.3 Association between epilepsy and hospital admission

5.9.3.1 All-cause hospital admission

766,244 pupils were followed up to assess the impact of epilepsy on risk of all-cause admission. The follow up period ranged from 1 year to 5 years and the mean follow up time was 4.33 years. 157,350 pupils (20.5%) were admitted to hospital at least once during the follow up period and there were 303,718 admissions in total. The probability of having an all-cause admission was significantly higher for pupils with epilepsy compared to those without epilepsy at all points during follow up. The probability of being admitted to hospital for any cause within 5 years was around 60% for children with epilepsy compared to 20% for children without epilepsy (Figure 43).

Figure 43 Kaplan-Meier survival curve: Epilepsy and all-cause hospital admission



Epilepsy was significantly associated with increased risk of all-cause hospital admission on univariate Cox regression analysis (HR 3.86, 95% CI 3.71-4.01) and following adjustment for sociodemographic (HR 3.79, 95% CI 3.65-3.94) and maternity (HR 3.72, 95% CI 3.58-3.97) confounders. However, the model did not satisfy the assumption of proportional hazards ($p<0.001$) and Poisson piecewise regression was used to more accurately model the variation in the hazard ratio over the follow up period. Table 65 presents the univariate and multivariate incidence rate ratios within each specific year of follow up. Figure 44 presents the fully adjusted incidence rate ratios from the Poisson piecewise model (black

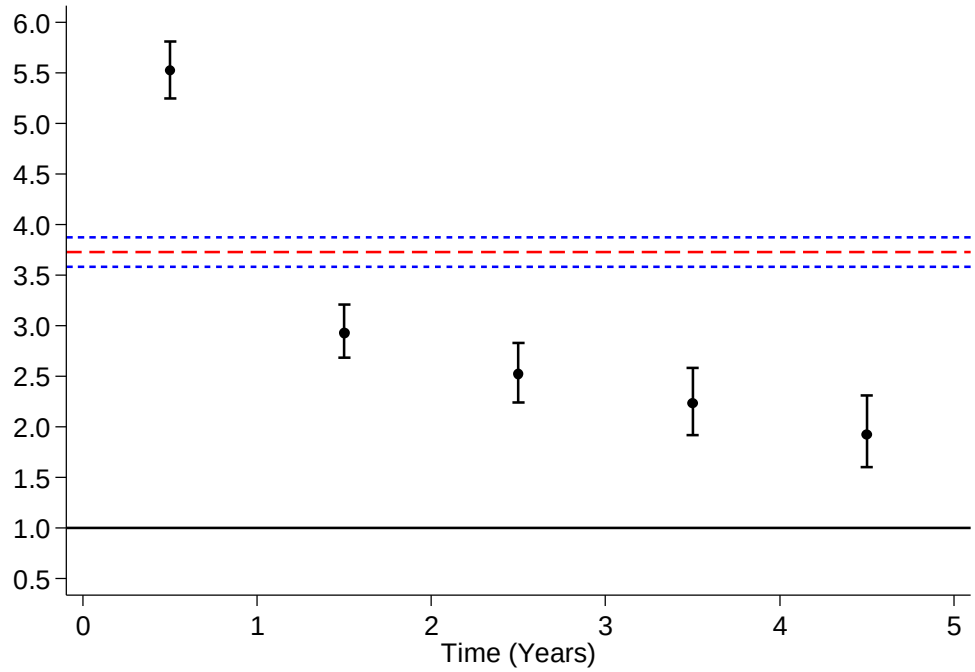
solid circles) alongside the static multivariate hazard ratio derived from the Cox proportional hazards model (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines). The Cox model demonstrated that, over the full follow up period, children with epilepsy had 3.7 times greater risk of hospital admission compared to non-epileptic pupils. However, the Poisson piecewise model demonstrated that the elevated risk of all-cause hospital admission experienced by epileptic pupils compared to non-epileptic pupils actually decreased in magnitude with increasing follow-up time. Epileptic children experienced a 5.5 fold risk of all-cause admission within the first year of follow-up. However, the increase in risk sharply decreased to 2.9 fold in the second year of follow-up and further decreased in later years. Therefore, the relative impact of epilepsy on risk of all-cause admission was particularly strong within the first year of follow up and consistently became weaker thereafter. This suggests that children with epilepsy are at highest risk of admission close to time of diagnosis. The relative increased risk of hospitalisation may diminish as children become stabilised on their medication and experience fewer seizures [273].

Table 65 Epilepsy and risk of all-cause hospital admission modelled as time from diagnosis until first admission

	Univariate model		Multivariate[1]		Multivariate[2]	
	IRR	95% CI	IRR	95% CI	IRR	95% CI
0 - 1 years	5.72***	5.44-6.01	5.63***	5.35-5.92	5.52***	5.25-5.81
1 - 2 years	3.01***	2.75-3.30	2.97***	2.71-3.25	2.93***	2.68-3.21
2 - 3 years	2.59***	2.30-2.91	2.54***	2.26-2.86	2.52***	2.24-2.83
3 - 4 years	2.34***	2.02-2.70	2.28***	1.97-2.64	2.23***	1.92-2.58
4 - 5 years	2.00***	1.67-2.41	1.96***	1.63-2.36	1.92***	1.60-2.31

Exponentiated coefficients; 95% confidence intervals in brackets
* p<0.05, ** p<0.01, *** p<0.001
[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)
[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

Figure 44 Epilepsy and risk of all-cause hospital admission modelled as time from diagnosis until first admission



The risk of all-cause admission associated with epilepsy also varied depending on pupil age at time of admission. Table 66 presents univariate and multivariate incidence rate ratios from the Poisson piecewise model within different age categories at time of admission. Figure 45 presents the fully adjusted incidence rate ratios from the Poisson piecewise model (black solid circles) alongside the static multivariate hazard ratio derived from the Cox proportional hazards model (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines). Epileptic children of all ages experienced an increased risk of all-cause hospital admission compared to non-epileptic children; however, there was a distinct trend whereby the association between epilepsy and risk of all-cause hospital admission became weaker for older children. The magnitude of the excess risk decreased from 7-fold in children aged 4-6 years to 2-fold in children aged 17-23 years. Epilepsy in younger children can be more severe [271] and this may explain why epilepsy has a greater impact on risk of hospital admission in younger pupils. There was no significant interaction between epilepsy and gender (p=0.349).

Table 66 Epilepsy and risk of all-cause hospital admission modelled as age at first admission

	Univariate model		Multivariate[1]		Multivariate[2]	
	IRR	95% CI	IRR	95% CI	IRR	95% CI
4 - 6 years	7.54***	6.85-8.30	7.33***	6.66-8.07	7.11***	6.45-7.83
7 - 8 years	5.47***	4.85-6.17	5.33***	4.72-6.01	5.28***	4.67-5.97
9 - 10 years	5.54***	4.94-6.21	5.39***	4.81-6.05	5.26***	4.68-5.91
11 - 12 years	5.04***	4.53-5.61	4.93***	4.43-5.49	4.86***	4.36-5.41
13 - 14 years	4.42***	4.01-4.87	4.36***	3.95-4.80	4.31***	3.91-4.75
15 - 16 years	3.72***	3.41-4.07	3.68***	3.37-4.02	3.62***	3.32-3.96
17 - 23 years	2.10***	1.91-2.32	2.09***	1.89-2.30	2.05***	1.85-2.26

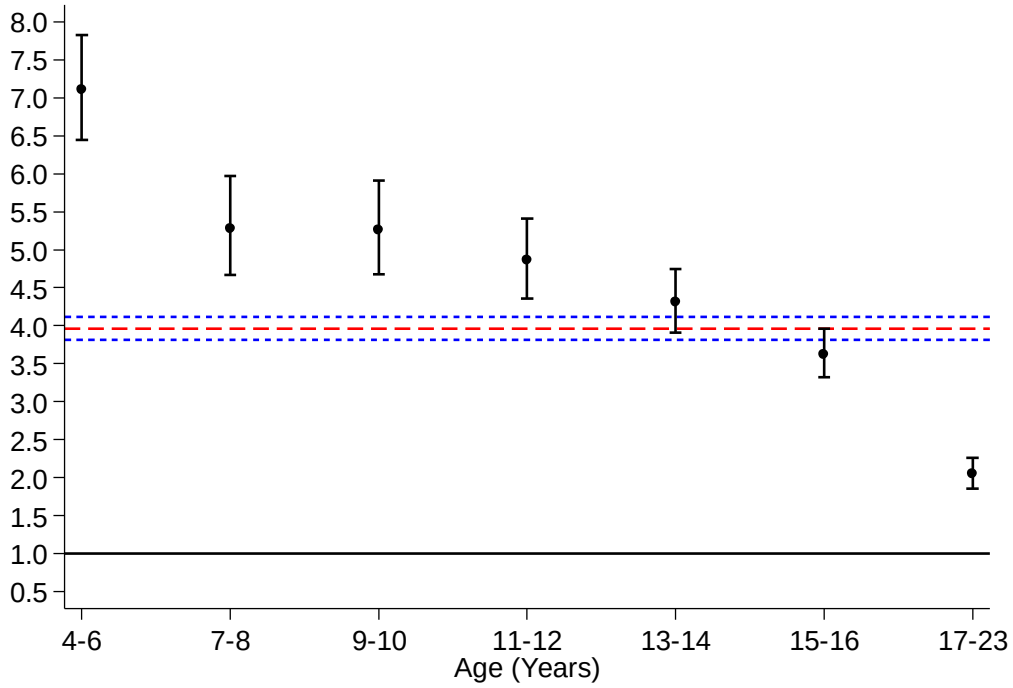
Exponentiated coefficients; 95% confidence intervals in brackets

* p<0.05, ** p<0.01, *** p<0.001

[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)

[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

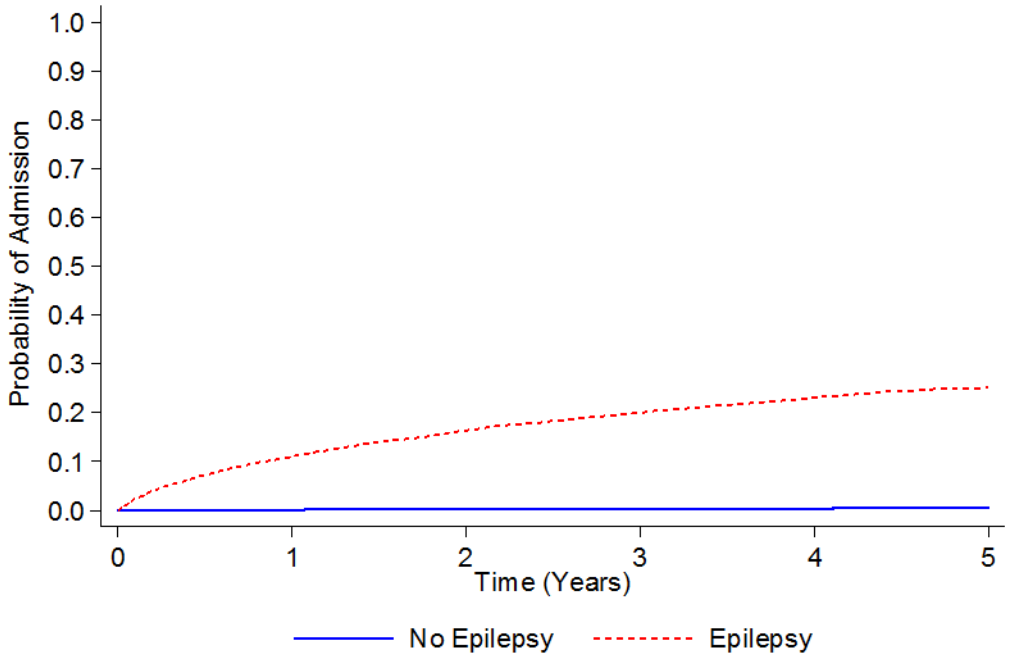
Figure 45 Epilepsy and risk of all-cause hospital admission modelled as age at first admission



5.9.3.2 Hospital admission related to the nervous system

766,244 pupils were followed up to assess the impact of epilepsy on risk of hospital admission related to the nervous system as this was the main cause of admission among children with epilepsy. The follow up period ranged from 1 year to 5 years and the mean follow up time was 4.33 years. 4,399 pupils (0.6%) were admitted to hospital at least once during the follow up period and there were 8,962 admissions in total. The probability of having an admission related to the nervous system was significantly higher for pupils with epilepsy compared to those without epilepsy at all points during follow up. The probability of being admitted to hospital for a nervous system admission within 5 years was around 20% for children with epilepsy compared to less than 5% for children without epilepsy (Figure 46).

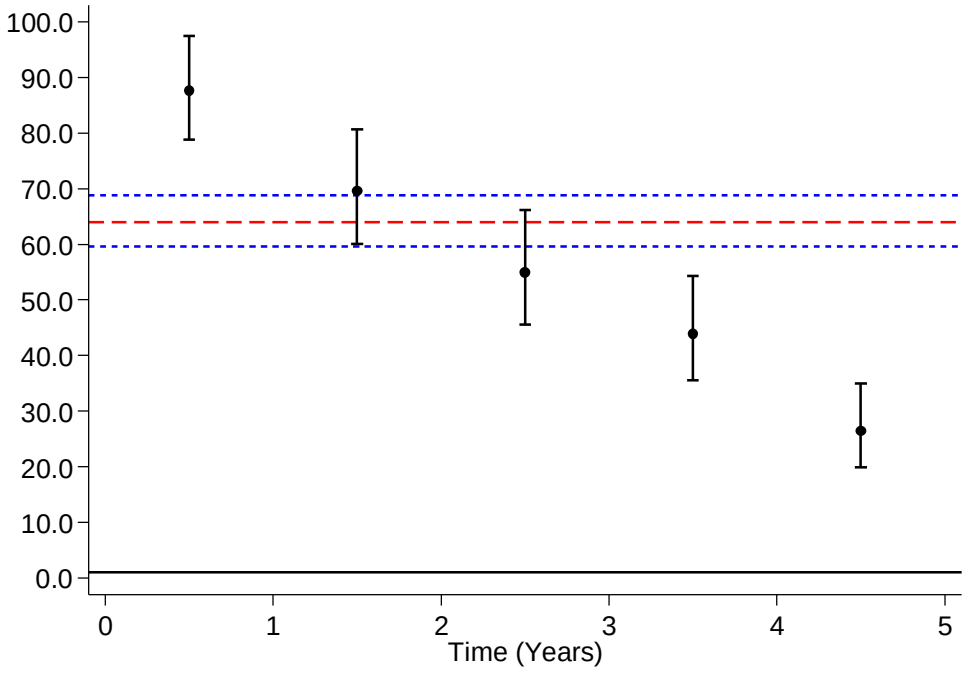
Figure 46 Kaplan-Meier survival curve: Epilepsy and hospital admission related to the nervous system



Epilepsy was significantly associated with increased risk of hospital admission related to the nervous system on univariate Cox regression analysis (HR 71.73, 95% CI 66.92-76.87) and following adjustment for sociodemographic (HR 70.49, 95% CI 65.76-75.56) and maternity (HR 63.20, 95% CI 58.85-67.88) confounders. However, the model did not satisfy the assumption of proportional hazards ($p<0.001$) and Poisson piecewise regression was used to more accurately model the hazard ratios over the follow up period. Figure 47

presents the fully adjusted incidence rate ratios from the Poisson piecewise model (black solid circles) alongside the static multivariate hazard ratio derived from the Cox proportional hazards model (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines). The Cox model demonstrated that, over the full follow up period, children with epilepsy had a 63-fold risk of hospital admission compared to non-epileptic pupils. However, the Poisson piecewise model demonstrated that the elevated risk of hospital admission related to the nervous system experienced by epileptic pupils compared to non-epileptic pupils actually decreased in magnitude with increasing follow-up time. Epileptic children experienced an 86-fold risk of admission in the first year of follow-up. However, the excess risk decreased with increasing follow up time to 26-fold in the 5th year of follow-up. This pattern was similar to that observed for epilepsy and risk of all-cause hospital admission and may again be explained by children becoming stabilised on their medication and experiencing fewer seizures [273].

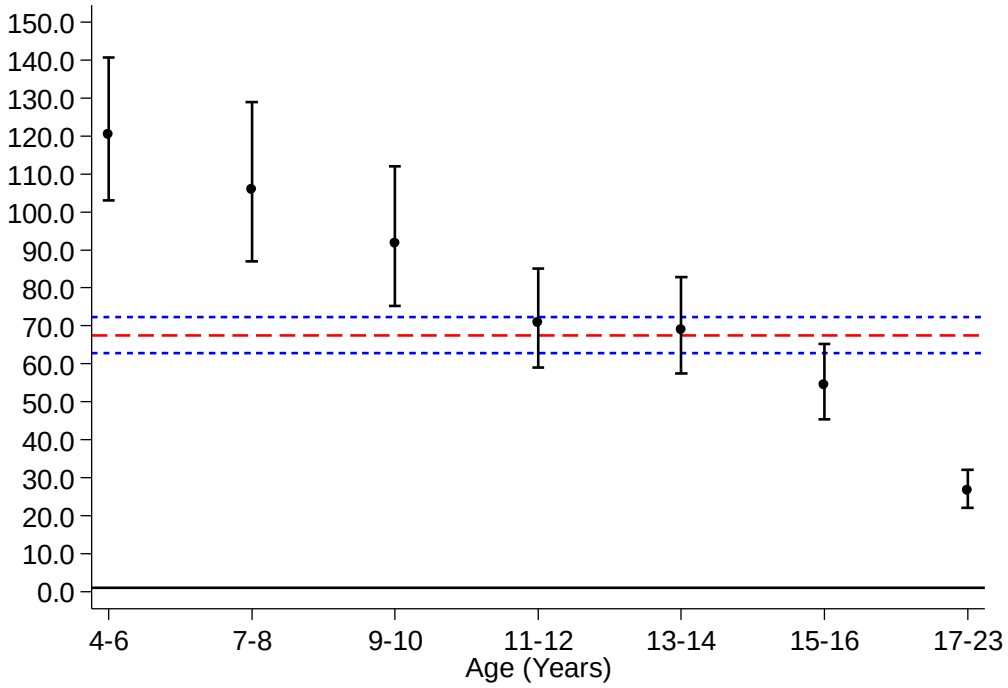
Figure 47 Epilepsy and risk of hospital admission related to the nervous system modelled as time from diagnosis until first admission



The association between epilepsy and risk of hospital admission related to the nervous system also varied depending on pupil age at time of admission. Figure 48 presents the fully adjusted incidence rate ratios from the Poisson piecewise model (black solid circles) alongside the static multivariate hazard ratio derived from the Cox proportional hazards

model (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines). Epileptic children of all ages experienced an increased risk of hospital admission related to the nervous system compared to non-epileptic children; however, there was a distinct trend whereby the association between epilepsy and risk of hospital admission related to the nervous system became weaker for older children. The magnitude of the excess risk decreased from 118-fold in children aged 4-6 years to 26-fold in children aged 17-23 years. This pattern was similar to that observed for epilepsy and risk of all-cause hospital admission.

Figure 48 Epilepsy and risk of hospital admission related to the nervous system modelled as age at first admission



There was a significant interaction between epilepsy and gender ($p<0.001$) whereby the association between epilepsy and risk of hospital admission related to the nervous system was stronger for boys than for girls during all points of follow up. This was despite similar absolute rates of hospital admission for this cause among unaffected children (1,668 [0.5%] vs. 1,590 [0.4%] among girls and boys respectively). Among children with epilepsy, admissions related to the nervous system were higher among boys than girls (615 [23.5%] vs. 526 [19.5%]). Figure 49 and Figure 50 show the fully adjusted gender specific incidence rate ratios from the Poisson piecewise models alongside the static multivariate gender specific hazard ratios derived from the corresponding proportional hazards models.

Confidence intervals are only presented for the Poisson piecewise models. The same patterns observed within the main analyses were observed within both gender groups namely that the excess risk of admission related to the nervous system for children with epilepsy decreased over follow up time and with increasing age.

Figure 49 Epilepsy and gender specific risk of hospital admission related to the nervous system modelled as time from diagnosis until first admission (blue=boys; red=girls)

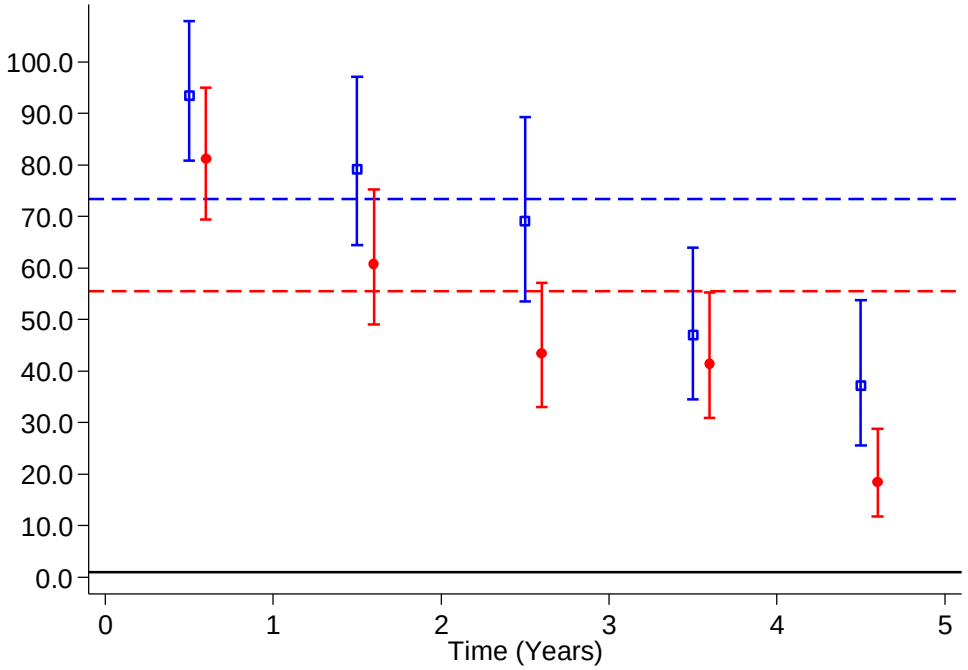
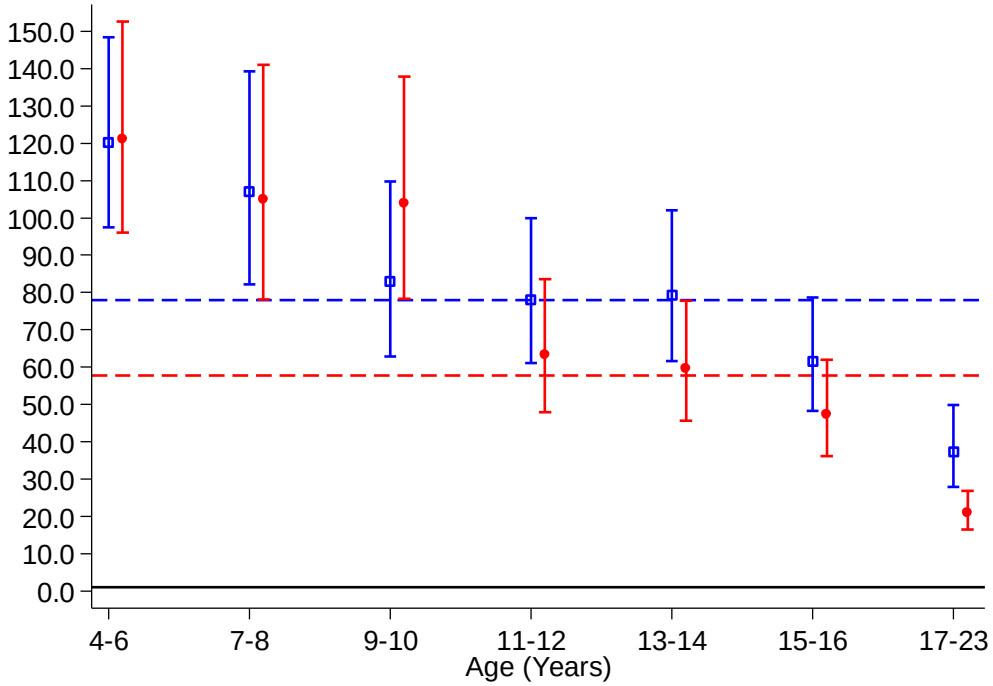


Figure 50 Epilepsy and gender specific risk of hospital admission related to the nervous system modelled by age at first admission (blue=boys; red=girls)

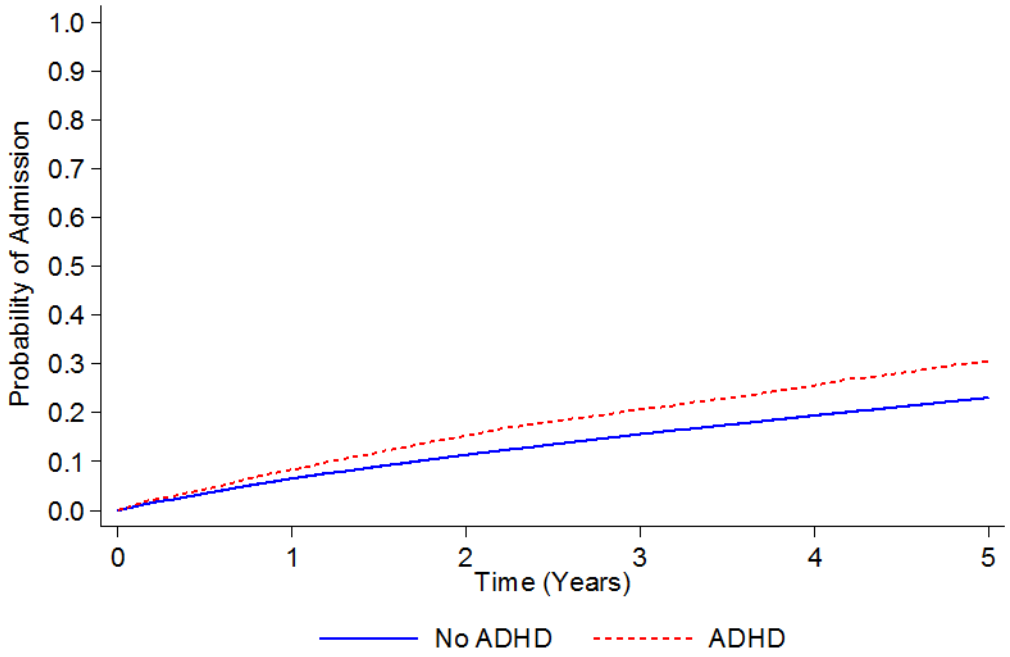


5.9.4 Association between attention deficit hyperactivity disorder and hospital admission

5.9.4.1 All-cause hospital admission

766,244 pupils were followed up to assess the impact of ADHD on risk of all-cause admission. The follow up period ranged from 1 year to 5 years and the mean follow up time was 4.33 years. 157,475 pupils (20.6%) were admitted to hospital at least once during the follow up period and there were 306,416 admissions in total. The probability of having an admission was significantly higher for pupils with ADHD compared to those without ADHD at all points during follow up. The probability of being admitted to hospital after 5 years was around 30% for children with ADHD compared to 20% for children without ADHD (Figure 51).

Figure 51 Kaplan-Meier survival curve: attention deficit hyperactivity disorder and all-cause hospital admission



ADHD was significantly associated with increased risk of hospital admission on univariate Cox regression analysis (HR 1.47, 95% CI 1.40-1.54) and following adjustment for sociodemographic (HR 1.40, 95% CI 1.34-1.47) and maternity (HR 1.33, 95% CI 1.27-1.39) confounders. The model satisfied the assumption of proportional hazards (p=0.983). Therefore this association was adequately described by a constant hazard ratio across the full follow up period. Figure 52 and Figure 53 present the fully adjusted incidence rate ratios from the Poisson piecewise models (black solid circles) alongside the static multivariate hazard ratios derived from the Cox proportional hazards models (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines). The Cox models demonstrated that, over the full follow up period, children with ADHD experienced 33% increased risk of hospital admission within 5 years compared to children without ADHD. The risk of admission associated with ADHD varied a little over the follow up period and for different ages but remained similar to the constant line in both cases therefore it could clearly be modelled via a constant hazard ratio.

Figure 52 Attention deficit hyperactivity disorder and risk of all-cause hospital admission modelled as time from diagnosis until first admission

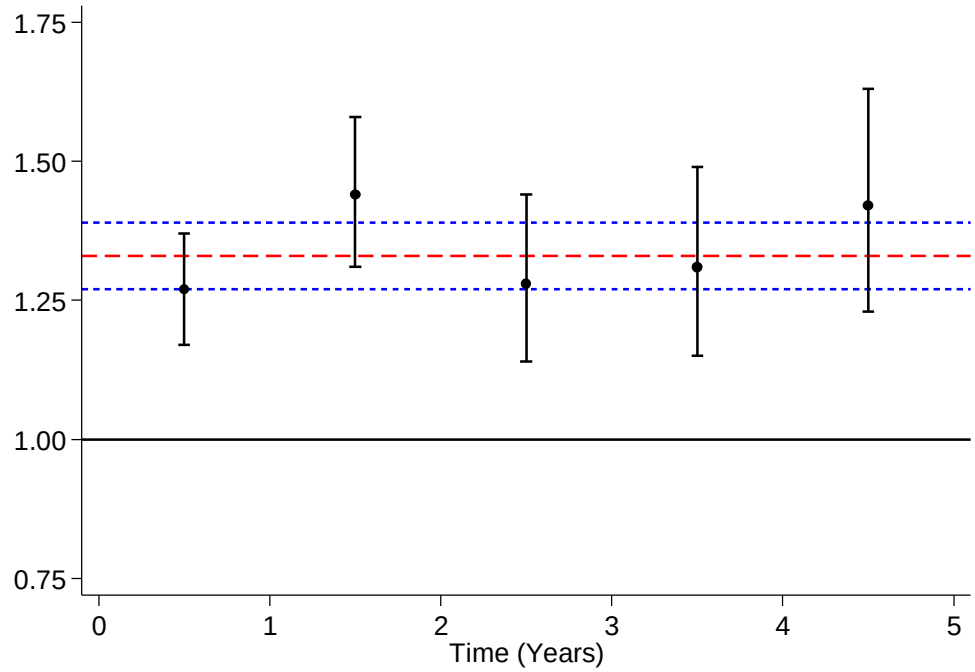
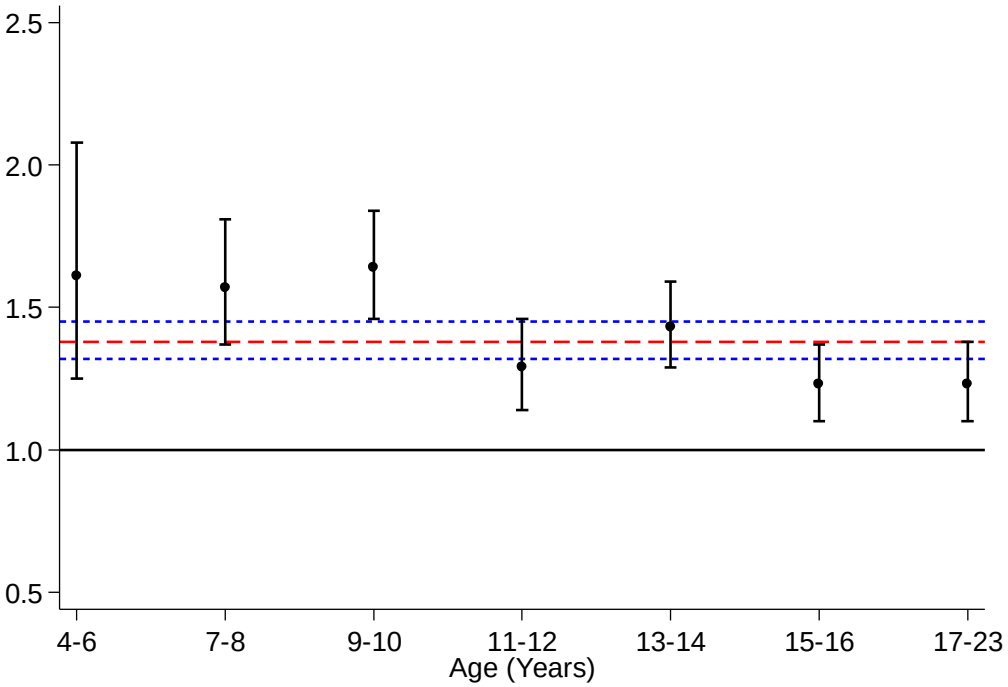


Figure 53 Attention deficit hyperactivity disorder and risk of all-cause hospital admission modelled as age at first admission

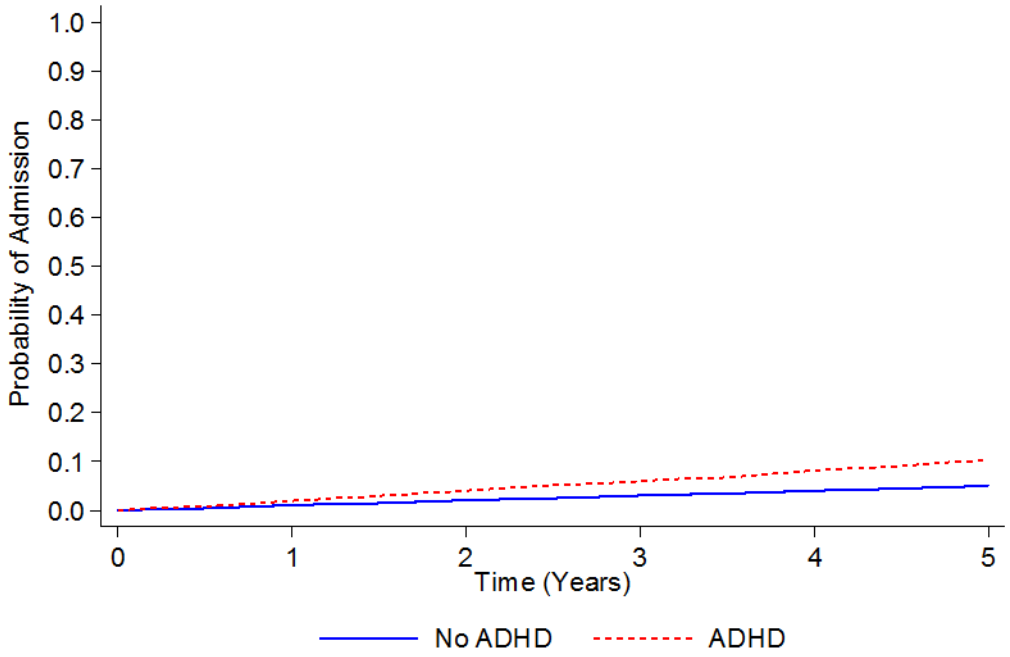


There was a significant interaction between ADHD and gender ($p<0.001$). The association between ADHD and risk of admission was stronger for girls than boys on univariate analysis (HR 1.87, 95% CI 1.68-2.08 versus HR 1.42, 95% CI 1.35-1.50) and after adjustment for sociodemographic (HR 1.77, 95% CI 1.59-1.97 versus HR 1.36, 95% CI 1.30-1.44) and maternity (HR 1.64, 95% CI 1.47-1.82 versus HR 1.31, 95% CI 1.24-1.38) confounders. Girls with ADHD experienced 64% increased risk of hospital admission compared to girls without ADHD whereas boys with ADHD experienced 31% increased risk of admission compared to boys without ADHD. This was despite similar absolute rates of hospital admission among unaffected children (77,232 [20.6%] vs. 78,426 [20.4%] among girls and boys respectively). Among children with ADHD, admissions were higher among girls than boys (338 [30%] vs. 1,479 [23.5%]).

5.9.4.2 Hospital admission related to injury, poisoning and other consequences of external causes

766,244 pupils were followed up to assess the impact of ADHD on risk of injury admission as this was the main cause of admission among children with ADHD. The probability of having an injury admission was significantly higher for pupils with ADHD compared to those without ADHD at all points during follow up. The follow up period ranged from 1 year to 5 years and the mean follow up time was 4.33 years. 33,630 pupils (4.4%) were admitted to hospital at least once during the follow up period and there were 41,765 admissions in total. The probability of being admitted to hospital after 5 years was around 10% for children with ADHD compared to less than 5% for children without ADHD (Figure 54).

Figure 54 Kaplan-Meier survival curve: Attention deficit hyperactivity disorder and hospital admission related to injury, poisoning or other external causes



ADHD was significantly associated with increased risk of hospital admission related to injury, poisoning or other external causes on univariate Cox regression analysis (HR 2.05, 95% CI 1.89-2.23) and following adjustment for sociodemographic (HR 1.70, 95% CI 1.57-1.85) and maternity (HR 1.59, 95% CI 1.47-1.73) confounders. The association between ADHD and risk of admission, specifically modelled using time from study as the time scale and adjusting for age at study entry, satisfied the proportional hazards assumption (p=0.456). However, modelling the association using age at admission as the time scale violated the assumption (p=0.037). This meant that, whilst the association between ADHD and risk of injury admission was adequately described using a constant hazard ratio over the full follow up period, it did differ in strength for different ages at admission. Children with ADHD experienced 59% increased risk of injury-related hospital admission within 5 years compared to children without ADHD. Figure 55 and Figure 56 present the fully adjusted incidence rate ratios from the Poisson piecewise models (black solid circles) alongside the static multivariate hazard ratio derived from the Cox proportional hazards models (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines). The risk of injury admission associated with ADHD varied a little over the follow up period but remained similar to the constant line therefore it could be modelled via a constant hazard ratio (Figure 55). However, the

risk of injury admission associated with ADHD increased for children older than 13 years of age (Figure 56). This may be partly explained by a combination of specific increased self-harm in children and adolescents with ADHD [274] and general increased self-harm in adolescence, particularly in girls [275,276].

Figure 55 Attention deficit hyperactivity disorder and risk of hospital admission related to injury, poisoning or other external causes modelled as time from diagnosis until first admission

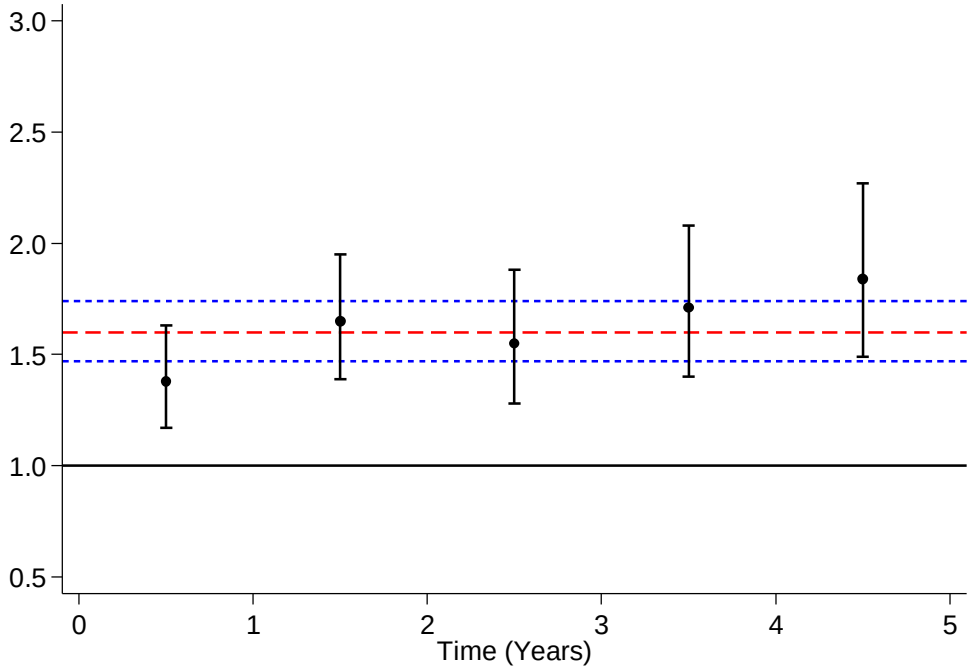
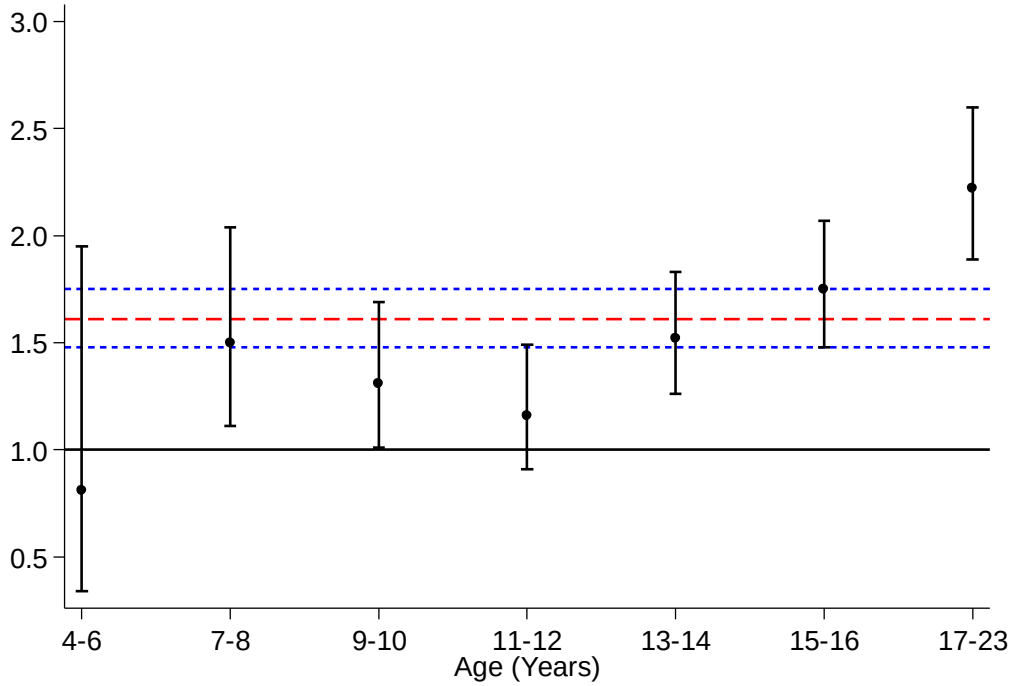
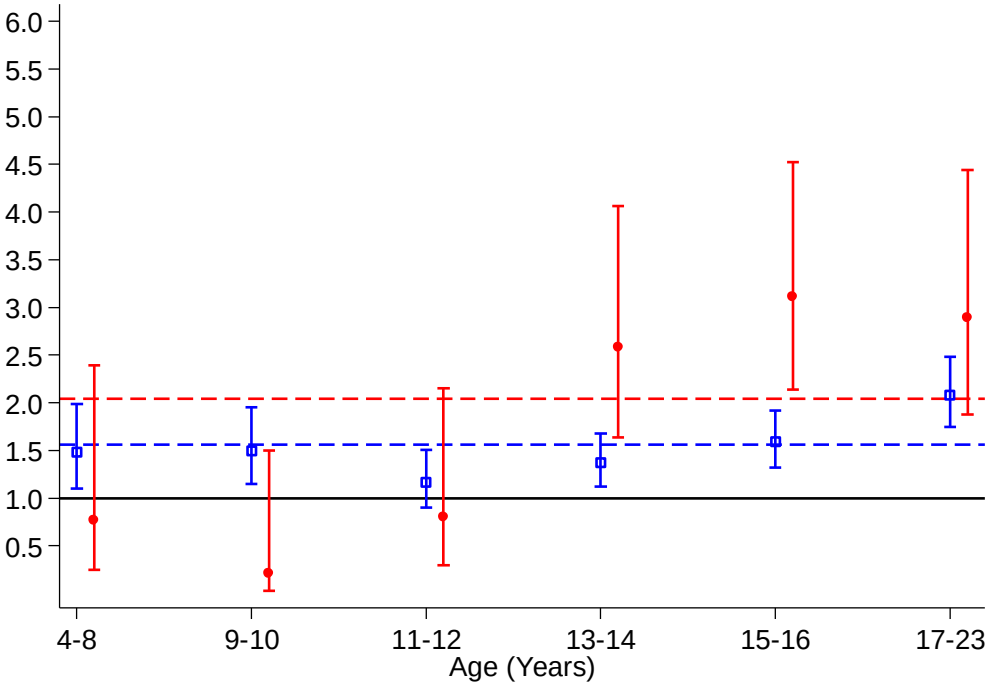


Figure 56 Attention deficit hyperactivity disorder and risk of hospital admission related to injury, poisoning or other external causes modelled as age at first admission



There was a significant interaction between ADHD and gender (p<0.001). The association between ADHD and risk of injury admission was stronger for girls than boys on univariate analysis (HR 2.39, 95% CI 1.91-2.99 versus HR 1.70, 95% CI 1.55-1.85) and after adjustment for sociodemographic (HR 2.24, 95% CI 1.79-2.80 versus HR 1.62, 95% CI 1.48-1.77) and maternity (HR 2.06, 95% CI 1.65-2.58 versus HR 1.53, 95% CI 1.40-1.67) confounders. The risk of injury admission for girls with ADHD was twice as high compared to girls without ADHD. The risk for boys with ADHD versus boys without ADHD was only 50% higher. Figure 57 presents fully adjusted gender specific incidence rate ratios from Poisson piecewise models using age at diagnosis alongside static multivariate gender specific hazard ratios from corresponding proportional hazards models. Confidence intervals are only presented for the Poisson piecewise models. The overall patterns observed within the main analyses remained within each group, whereby the association between ADHD and risk of injury admission increased for children older than 13 years of age. However, it was noticeable that the relative risk of injury admission was greater in girls which may be explained by their increased self-harm in adolescence [275].

Figure 57 Attention deficit hyperactivity disorder and gender specific risk of hospital admission related to injury, poisoning or other external causes modelled by age at first admission (blue=boys; red=girls)

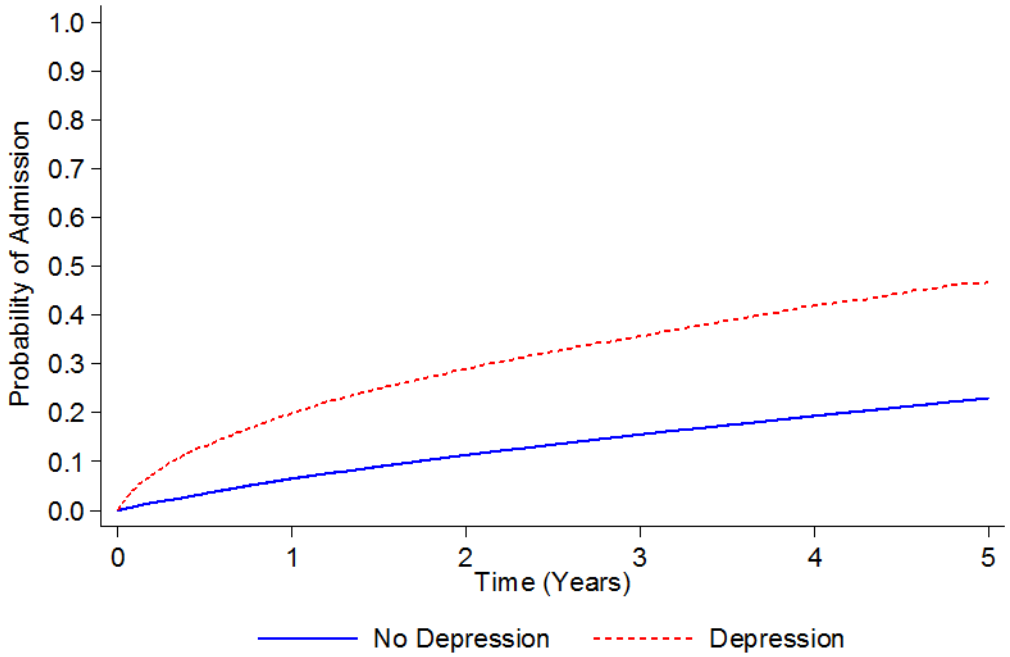


5.9.5 Association between depression and hospital admission

5.9.5.1 All-cause hospital admission

7 pupils who received ineligible anti-depressant drugs (section 4.3.2.2.5) were excluded. Therefore 766,237 pupils were followed up to assess the impact of depression on risk of all-cause admission. The follow up period ranged from 1 year to 5 years and the mean follow up time was 4.33 years. 157,291 pupils (20.5%) were admitted to hospital at least once during the follow up period and there were 304,507 admissions in total. The probability of having an admission was significantly higher for pupils with depression compared to those without depression at all points during follow up. The probability of being admitted to hospital after 5 years was around 40% for children with depression compared to 20% for children without depression (Figure 58).

Figure 58 Kaplan-Meier survival curve: Depression and all-cause hospital admission



Depression was significantly associated with increased risk of all-cause hospital admission on univariate Cox regression analysis (HR 2.57, 95% CI 2.45-2.69) and following adjustment for sociodemographic (HR 2.57, 95% CI 2.45-2.69) and maternity (HR 2.56, 95% CI 2.44-2.68) confounders. However, the model did not satisfy the assumption of proportional hazards ($p<0.001$) and Poisson piecewise regression was used to more accurately model the hazard ratios over the follow up period. Table 67 presents the univariate and multivariate incidence rate ratios within each specific year of follow up. Figure 59 presents the fully adjusted incidence rate ratios from the Poisson piecewise model (black solid circles) alongside the static multivariate hazard ratio derived from the Cox proportional hazards model (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines). The Cox model demonstrated that, over the full follow up period, depressed children had 2.56 times greater risk of hospital admission compared to pupils who were not depressed. However, the Poisson piecewise model demonstrated that the elevated risk of hospital admission experienced by depressed pupils actually decreased in magnitude with increasing follow-up time. Depressed children experienced 3.2 times greater risk of admission within the first year of follow-up. However, the excess risk sharply decreased to 2-fold across the second year of follow-up and decreased again in later years. Therefore, the relative impact of depression on risk of admission was particularly strong within the first year of follow up and consistently

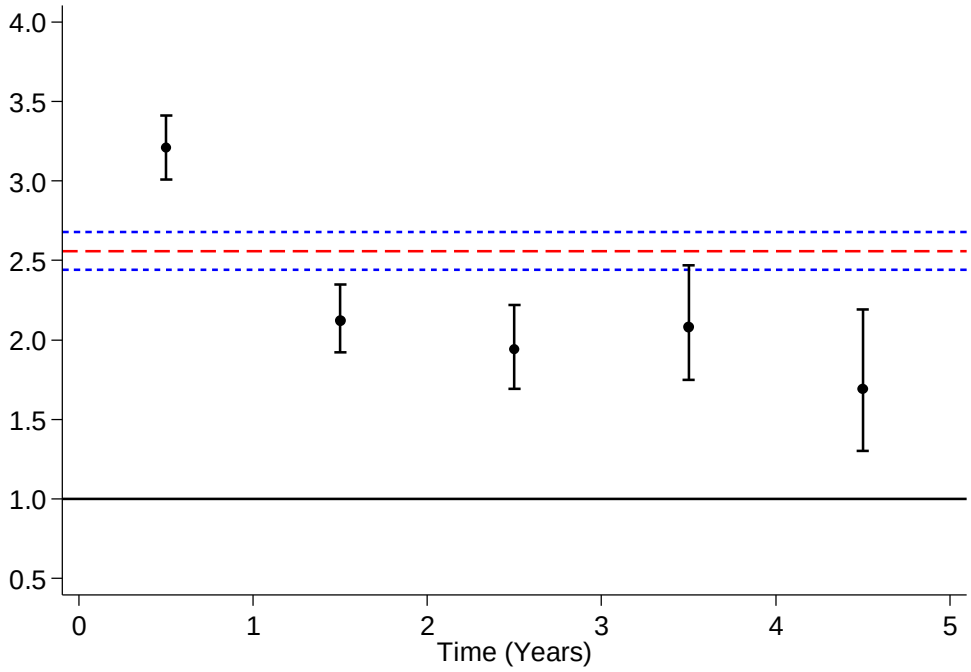
became weaker throughout the follow up period. This suggests that children with depression are at highest risk of admission close to time of diagnosis. A potential explanation may be that the excess risk of hospitalisation decreases through time as depressed children become stabilised on their medication.

Table 67 Depression and risk of all-cause hospital admission modelled as time from diagnosis until first admission

	Univariate model		Multivariate[1]		Multivariate[2]	
	IRR	95% CI	IRR	95% CI	IRR	95% CI
0 - 1 years	3.21***	3.02-3.42	3.21***	3.02-3.41	3.21***	3.01-3.41
1 - 2 years	2.13***	1.92-2.36	2.12***	1.92-2.35	2.12***	1.92-2.35
2 - 3 years	1.94***	1.69-2.22	1.94***	1.70-2.23	1.94***	1.69-2.22
3 - 4 years	2.08***	1.75-2.47	2.09***	1.76-2.49	2.08***	1.75-2.47
4 - 5 years	1.70***	1.31-2.20	1.71***	1.32-2.21	1.69***	1.30-2.19

Exponentiated coefficients; 95% confidence intervals in brackets
* p<0.05, ** p<0.01, *** p<0.001
[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)
[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

Figure 59 Depression and risk of all-cause hospital admission modelled as time from diagnosis until first admission



The risk of admission associated with depression also varied depending on pupil age at time of admission. Table 68 presents univariate and multivariate incidence rate ratios from the Poisson piecewise model for different age categories at admission. Figure 60 presents the fully adjusted incidence rate ratios from the Poisson piecewise model (black solid circles) alongside the static multivariate hazard ratio derived from the Cox proportional hazards model (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines). Depressed children of all ages experienced an increased risk of all-cause hospital admission compared to children without depression; however, the association between depression and risk of hospital admission was stronger for children aged between 13 and 14 years of age. This could be due to different patterns of behaviour such as increased self-harm in depressed teenage children putting them at higher risk of admission [277]. Given children appear to be most at risk of admission close to time of diagnosis, this observation may also partially reflect the higher onset of depression in teenage girls reported elsewhere [278-281].

Table 68 Depression and risk of all-cause hospital admission modelled as age at first admission

	Univariate model		Multivariate[1]		Multivariate[2]	
	IRR	95% CI	IRR	95% CI	IRR	95% CI
4 - 6 years	3.33***	1.97-5.62	3.34***	1.98-5.64	3.46***	2.05-5.85
7 - 8 years	2.03**	1.28-3.22	2.01**	1.27-3.19	2.09**	1.32-3.32
9 - 10 years	2.85***	2.12-3.83	2.83***	2.10-3.80	2.79***	2.07-3.76
11 - 12 years	3.36***	2.70-4.18	3.31***	2.66-4.12	3.24***	2.60-4.05
13 - 14 years	4.45***	3.94-5.02	4.45***	3.94-5.02	4.39***	3.88-4.96
15 - 16 years	3.76***	3.49-4.06	3.76***	3.49-4.06	3.73***	3.46-4.02
17 - 23 years	2.17***	2.01-2.34	2.19***	2.03-2.36	2.18***	2.02-2.35

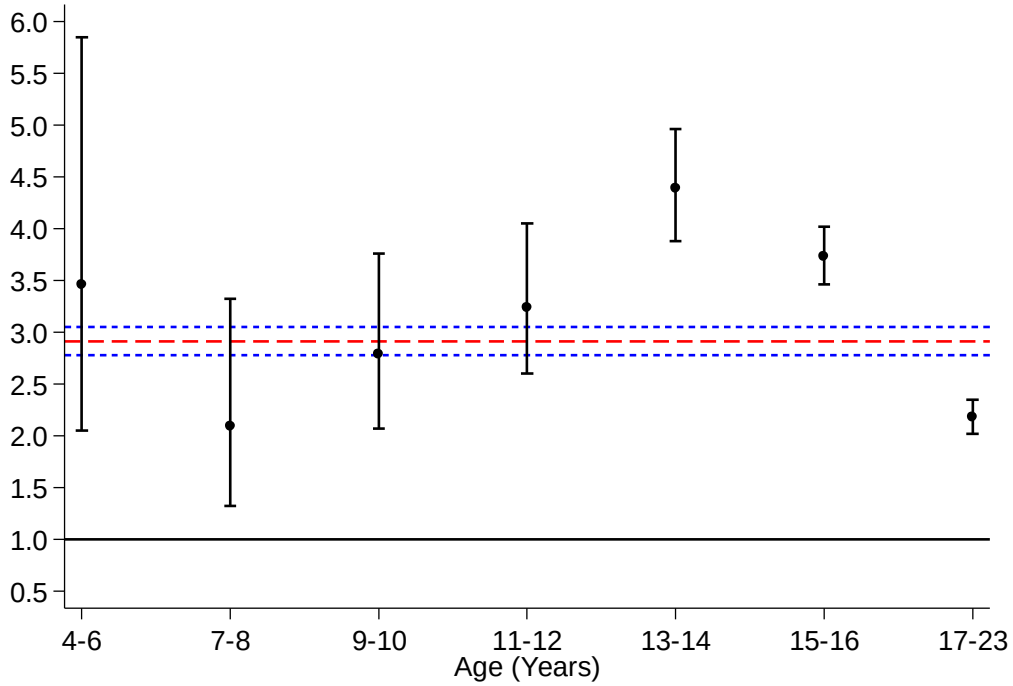
Exponentiated coefficients; 95% confidence intervals in brackets

* p<0.05, ** p<0.01, *** p<0.001

[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)

[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

Figure 60 Depression and risk of all-cause hospital admission modelled by age at first admission



There was a significant interaction between depression and gender ($p<0.001$) whereby the association between depression and risk of admission was stronger for girls than for boys. This was despite similar absolute rates of hospital admission among unaffected children (75,884 [20.4%] vs. 79,579 [20.5%] among girls and boys respectively). Among depressed children, admissions were higher among girls than boys (1,327 [37%] vs. 501 [28.6%]). The overall pattern remained within each group, whereby the association between depression and risk of hospital admission became weaker throughout the follow up period and was stronger for children in their early teens and decreased thereafter. The association between depression and risk of admission was visibly stronger for girls between the ages of 11 and 14 years. This may be explained by a combination of specific increased self-harm in depressed adolescents [277] and additional excess self-harm in adolescent girls [275,276]. Figure 61 and Figure 62 present fully adjusted gender specific incidence rate ratios from the Poisson piecewise models alongside static multivariate gender specific hazard ratios derived from the corresponding proportional hazards models. Confidence intervals are only presented for the Poisson piecewise models.

Figure 61 Depression and gender specific risk of all-cause hospital admission modelled as time from diagnosis until first admission (blue=boys; red=girls)

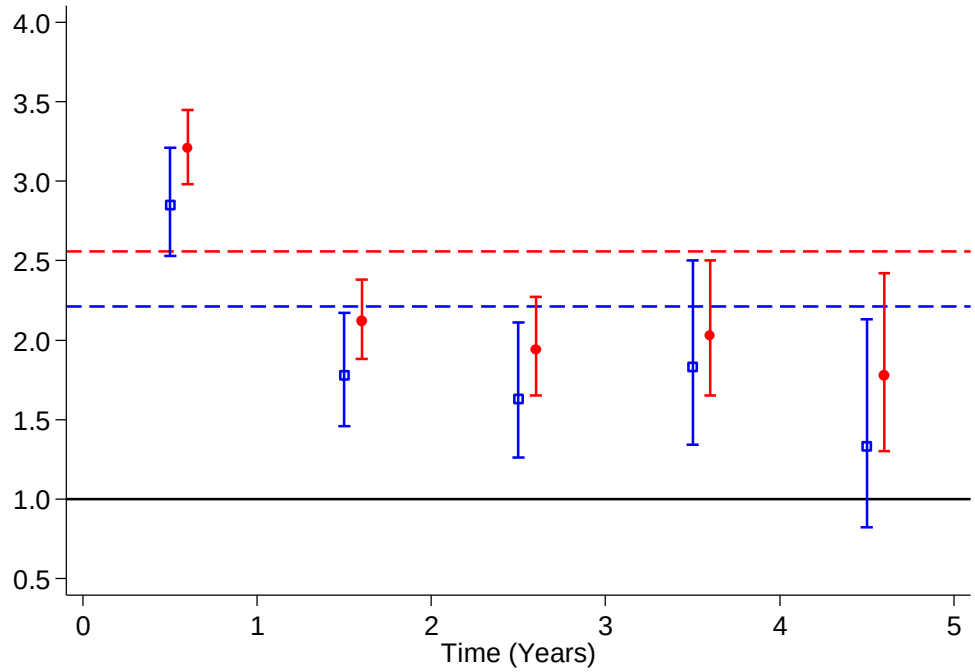
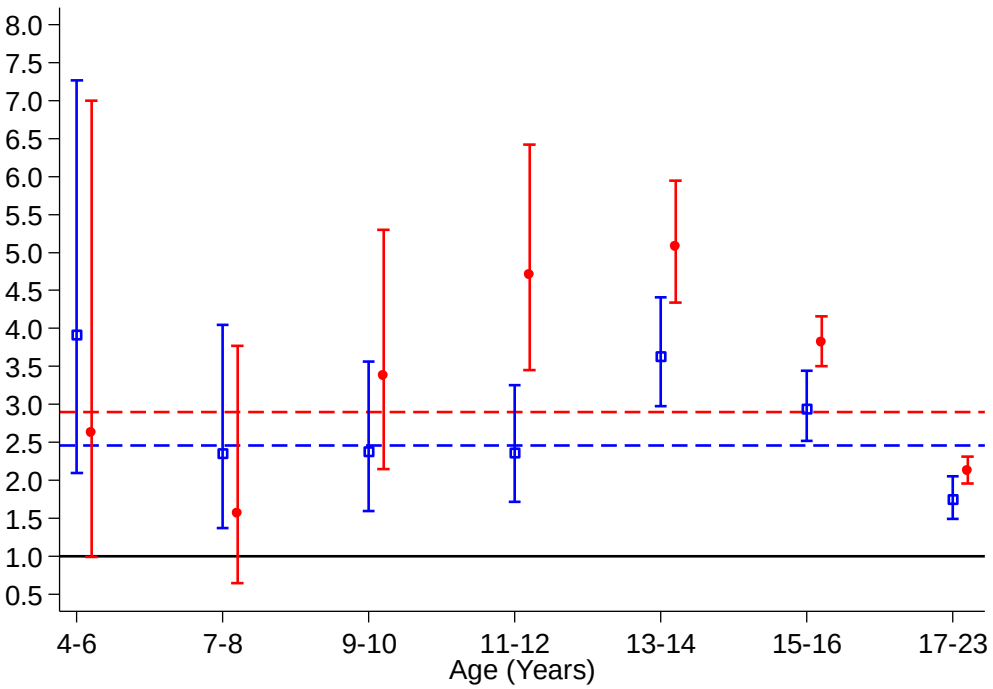


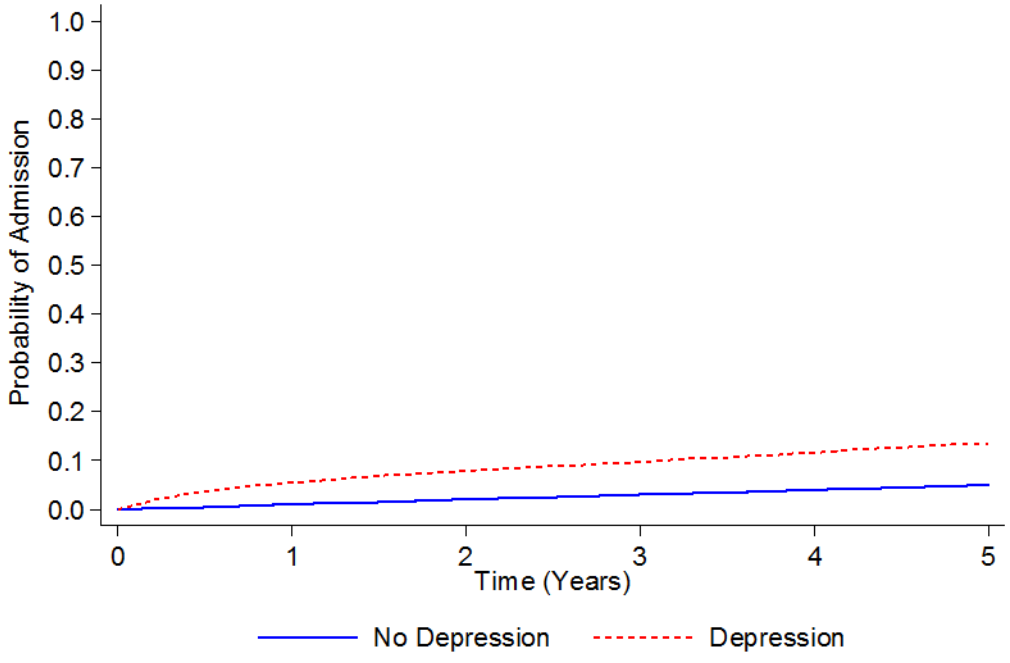
Figure 62 Depression and gender specific risk of all-cause hospital admission modelled by age at first admission (blue=boys; red=girls)



5.9.5.2 Hospital admission related to injury, poisoning and other consequences of external causes

7 pupils who received ineligible anti-depressant drugs (section 4.3.2.2.5) were excluded. Therefore 766,237 pupils were followed up to assess the impact of depression on risk of hospital admission related to injury, poisoning or other external causes as this was one of the main causes of admission among children with depression. The follow up period ranged from 1 year to 5 years and the mean follow up time was 4.33 years. 33,567 pupils (4.4%) were admitted to hospital at least once during the follow up period and there were 41,560 admissions in total. The probability of having an admission was significantly higher for pupils with depression compared to those without depression at all points during follow up. The probability of being admitted to hospital after 5 years was around 10% for children with depression compared to less than 5% for children without depression (Figure 63).

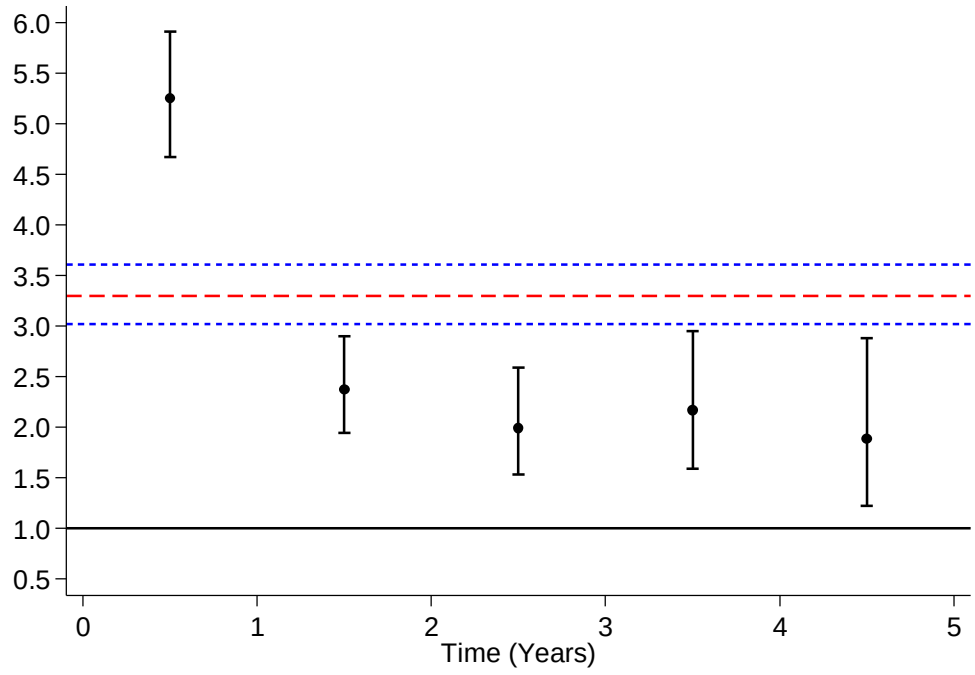
Figure 63 Kaplan-Meier survival curve: Depression and hospital admission related to injury, poisoning or other external causes



Depression was significantly associated with increased risk of hospital admission related to injury, poisoning or other external causes on univariate Cox regression analysis (HR 3.10, 95% CI 2.84-3.39) and following adjustment for sociodemographic (HR 3.31, 95% CI 3.03-3.62) and maternity (HR 3.32, 95% CI 3.04-3.63) confounders. However, the model

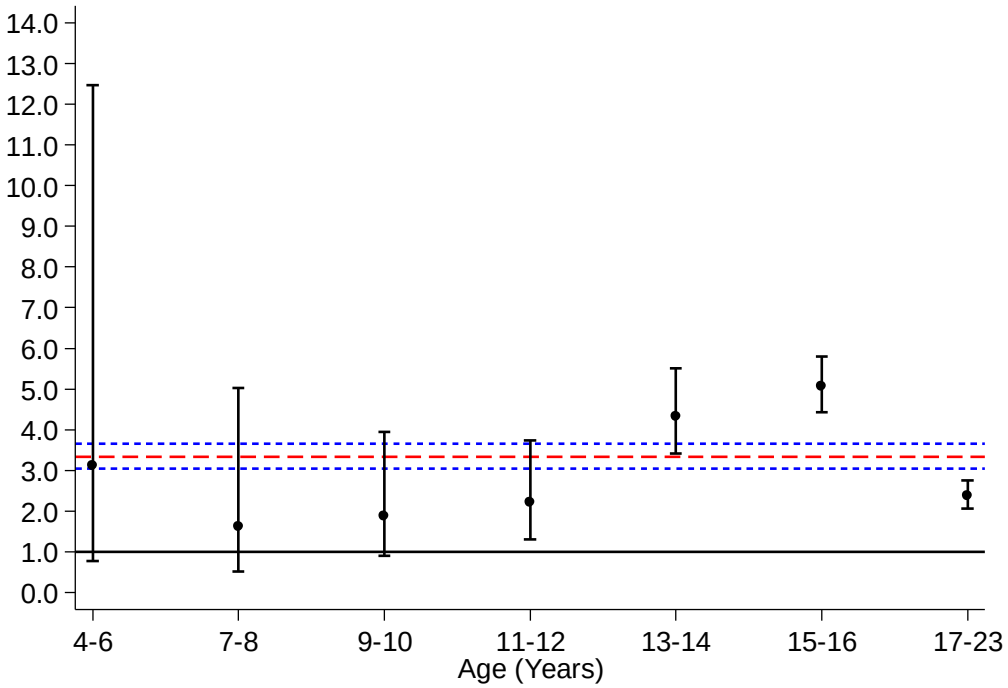
did not satisfy the assumption of proportional hazards ($p<0.001$) and Poisson piecewise regression was used to more accurately model the hazard ratios over the follow up period. Figure 64 presents the fully adjusted incidence rate ratios from the Poisson piecewise model (black solid circles) alongside the static multivariate hazard ratio derived from the Cox proportional hazards model (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines). The Cox model demonstrated that, over the full follow up period, depressed children had 3.3 times greater risk of injury-related hospital admission compared to pupils who were not depressed. However, the Poisson piecewise model demonstrated that the elevated risk of hospital admission experienced by depressed pupils actually decreased in magnitude with increasing follow-up time. Depressed children experienced a 5.3-fold risk of admission within the first year of follow-up. However, the excess risk sharply decreased to under 2.5-fold from the second year of follow-up. Therefore, the relative impact of depression on risk of hospital admission related to injury, poisoning or other external causes was particularly strong within the first year of follow up. This pattern was similar to that observed for depression and risk of all-cause hospital admission and may be explained by depressed children becoming stabilised on their medication with time.

Figure 64 Depression and risk of hospital admission related to injury, poisoning or other external causes modelled as time from diagnosis until first admission



The risk of admission associated with depression also varied depending on pupil age at time of admission. Figure 65 presents the fully adjusted incidence rate ratios from the Poisson piecewise model (black solid circles) alongside the static multivariate hazard ratio derived from the Cox proportional hazards model (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines). Depressed children of all ages experienced an increased risk of hospital admission compared to children without depression; however, the association between depression and risk of hospital admission related to injury, poisoning or other external causes was stronger for children aged between 13 years and 16 years. This pattern was also similar to that observed for depression and risk of all-cause hospital admission. This could be due to different patterns of behaviour such as increased self-harm in depressed teenage children putting them at higher risk of admission [277].

Figure 65 Depression and risk of hospital admission related to injury, poisoning or other external causes modelled by age at first admission



There was a significant interaction between depression and gender ($p<0.001$) whereby the association between depression and risk of hospital admission related to injury, poisoning or other external causes was stronger for girls than for boys (Figure 66 and Figure 67). There was a visible gender difference particularly within the first year of follow up and for children between the ages of 13 and 16 years whereby the strength of the association

between depression and risk of admission was greater for girls than boys. This suggests that girls are particularly at risk of admission at time of diagnosis and in teenage years. Again these findings may be partly explained by a combination of specific increased self-harm in depressed adolescents [277] and additional excess self-harm in adolescent girls [275,276] and may reflect reported higher onset of depression in teenage girls [278-281].

Figure 66 Depression and gender specific risk of hospital admission related to injury, poisoning or other external causes modelled as time from diagnosis until first admission (blue=boys; red=girls)

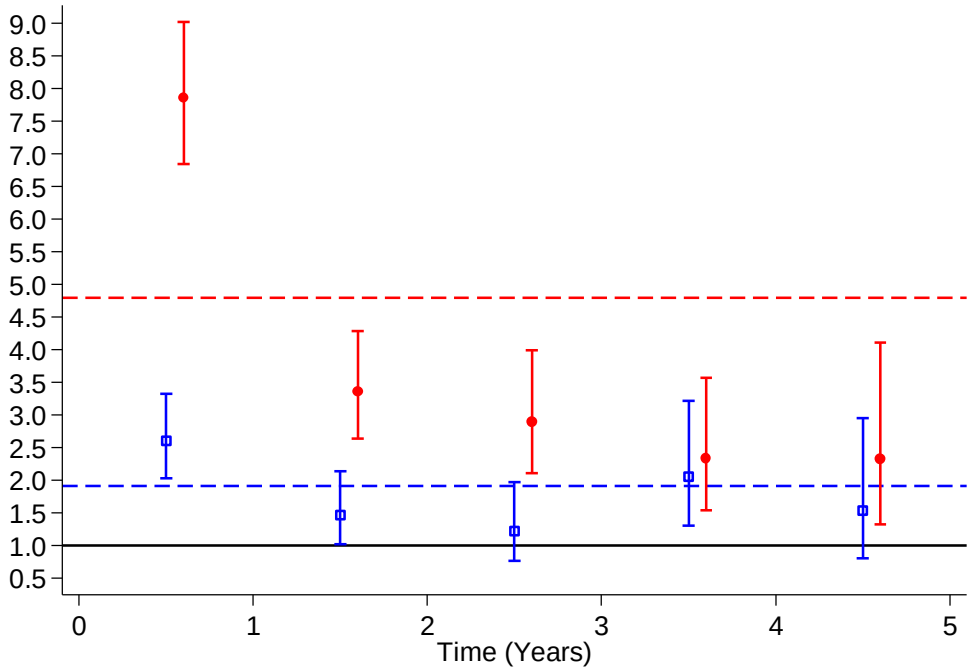
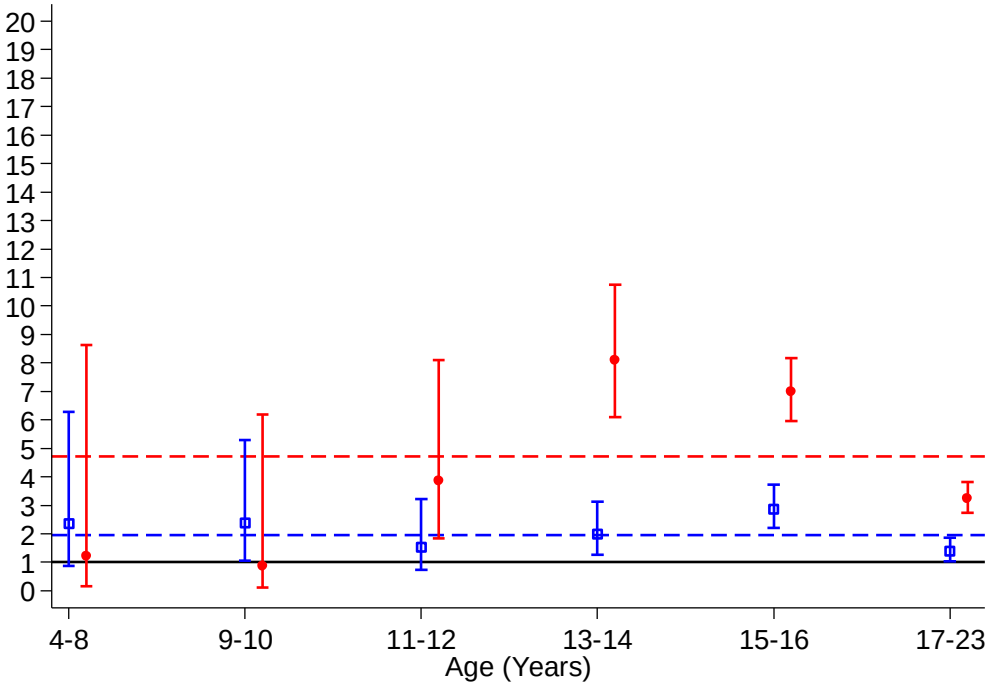


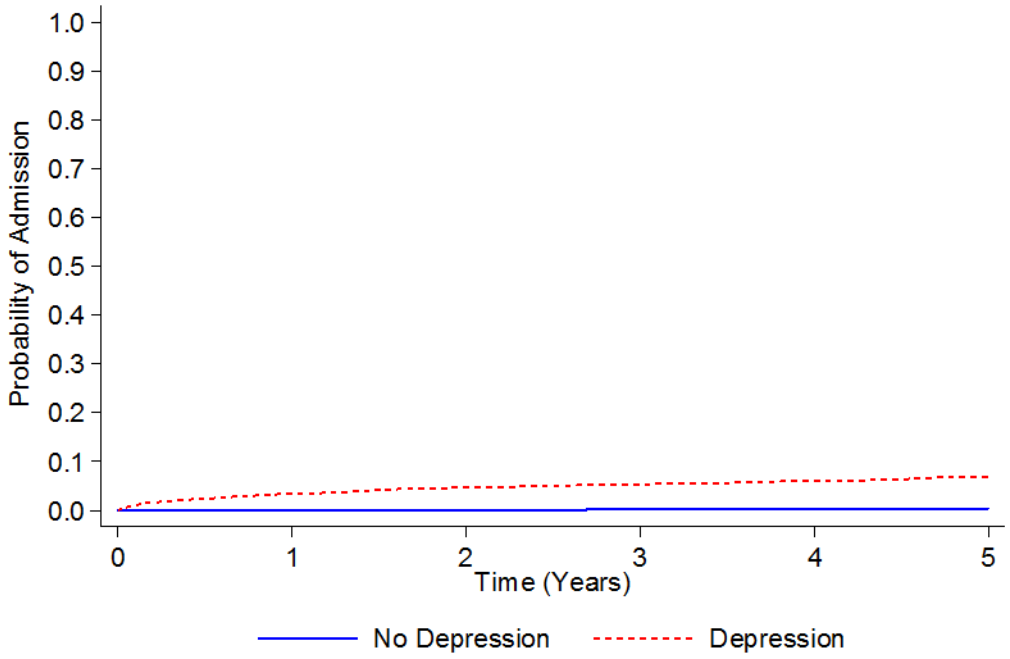
Figure 67 Depression and gender specific risk of hospital admission related to injury, poisoning or other external causes modelled by age at first admission (blue=boys; red=girls)



5.9.5.3 Hospital admission related to mental and behavioural disorders

7 pupils who received ineligible anti-depressant drugs (section 4.3.2.2.5) were excluded. Therefore 766,237 pupils were followed up to assess the impact of depression on risk of hospital admission related to mental and behavioural disorders as this was one of the main causes of admission among children with depression. The follow up period ranged from 1 year to 5 years and the mean follow up time was 4.33 years. 2,540 pupils (0.3%) were admitted to hospital at least once during the follow up period and there were 6,324 admissions in total. The probability of having an admission was significantly higher for pupils with depression compared to those without depression at all points during follow up. The probability of being admitted to hospital after 5 years was around 10% for children with depression compared to less than 5% for children without depression (Figure 68).

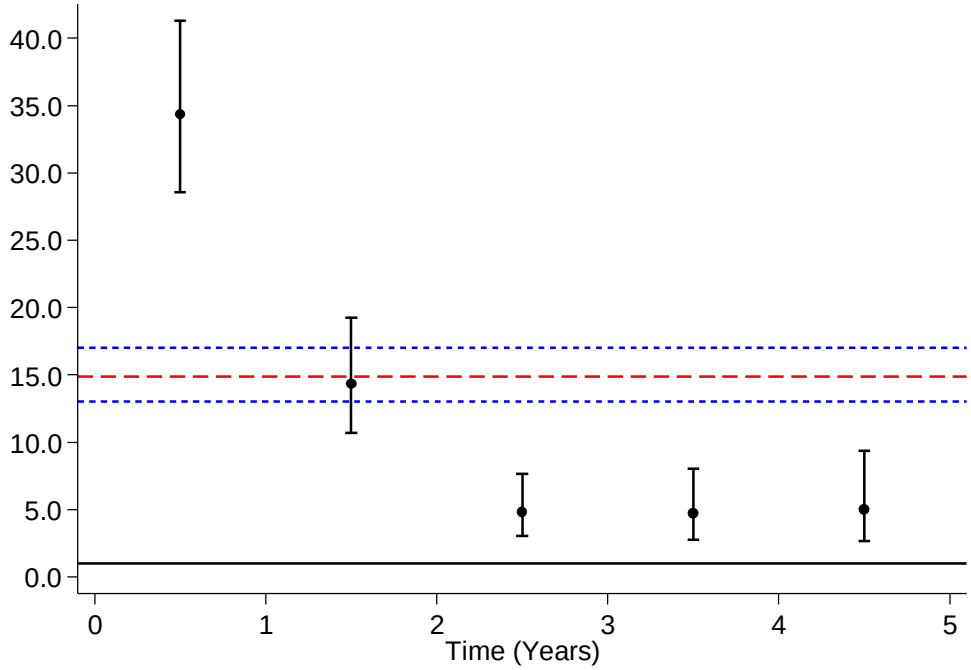
Figure 68 Kaplan-Meier survival curve: Depression and hospital admission related to mental and behavioural disorders



Depression was significantly associated with increased risk of hospital admission related to mental and behavioural disorders on univariate Cox regression analysis (HR 15.37, 95% CI 13.49-17.51) and following adjustment for sociodemographic (HR 15.12, 95% CI 13.25-17.24) and maternity (HR 14.89, 95% CI 13.05-17.00) confounders. The model did not satisfy the assumption of proportional hazards ($p<0.001$) therefore Poisson piecewise regression was used. Figure 69 presents the fully adjusted incidence rate ratios from the Poisson piecewise model (black solid circles) alongside the static multivariate hazard ratio derived from the Cox proportional hazards model (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines). The Cox model demonstrated that, over the full follow up period, depressed children had an almost 15-fold risk of mental health related hospital admission compared to pupils who were not depressed. However, the Poisson piecewise model demonstrated that the elevated risk of hospital admission experienced by depressed pupils actually decreased in magnitude with increasing follow-up time. Depressed children experienced a 34-fold risk of admission within the first year of follow-up. However, the excess risk sharply decreased to 15-fold in the second year of follow-up and decreased again thereafter. Therefore, the relative impact of depression on risk of hospital admission related to mental and behavioural disorders was

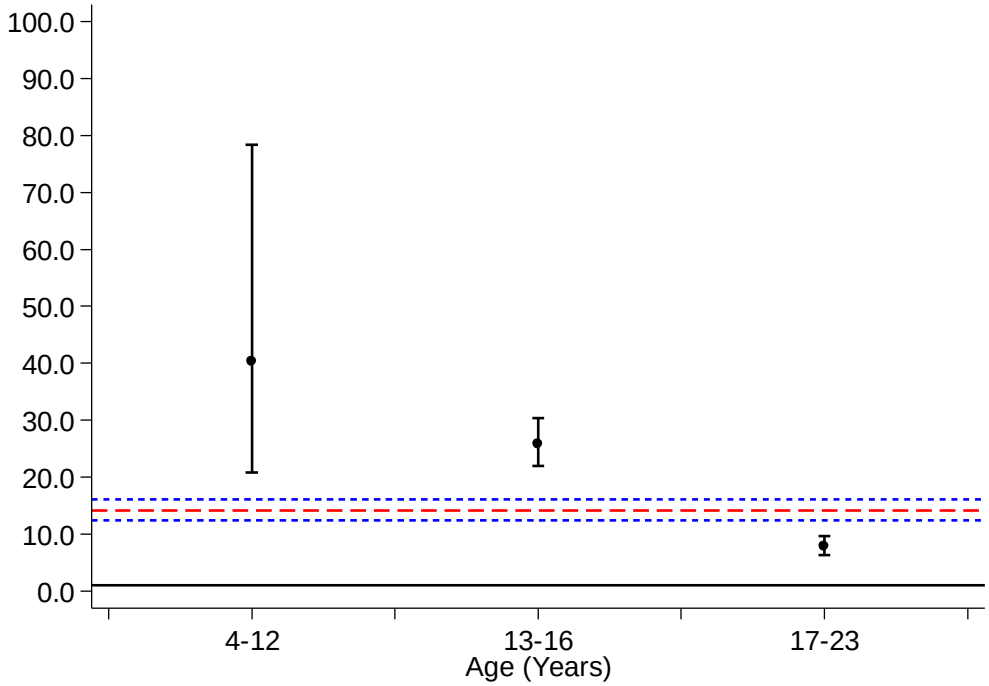
particularly strong within the first year of follow up. This may again be explained by depressed children becoming stabilised on their medication with time.

Figure 69 Depression and risk of hospital admission related to mental and behavioural disorders modelled as time from diagnosis until first admission



The association between depression and risk of hospital admission related to mental and behavioural disorders also varied depending on pupil age at time of admission. Figure 70 presents the fully adjusted incidence rate ratios from the Poisson piecewise model (black solid circles) alongside the static multivariate hazard ratio derived from the Cox proportional hazards model (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines). An elevated risk of admission was observed across all ages however the magnitude of the excess risk was smallest in children older than 17 years of age. It is not immediately clear why this pattern exists however it was notable that the confidence interval was very wide for the category representing children between 4 and 12 years of age due to small numbers of admissions. There was no significant interaction between depression and gender (p=0.146).

Figure 70 Depression and risk of hospital admission related to mental and behavioural disorders modelled by age at first admission



5.9.6 Total number and lengths of stay of all-cause hospital admissions

There was a significant association between epilepsy and total number of all-cause hospital admissions in the 5 year follow up period on univariate analysis (IRR 12.56, 95% CI 7.51-21.00) and following adjustment for sociodemographic (IRR 17.40, 95% CI 8.65-35.01) and maternity (IRR 14.72, 95% CI 7.34-29.51) confounders. There was also a significant association between depression and total number of all-cause hospital admissions on univariate analysis (IRR 2.99, 95% CI 1.56-5.73) and following adjustment for sociodemographic (IRR 3.24, 95% CI 2.28-4.59) and maternity (IRR 3.31, 95% CI 2.29-4.78) confounders. There was a significant association between ADHD and total number of admissions on univariate analysis (IRR 0.73, 95% CI 0.55-0.96); however, this association disappeared after adjustment for sociodemographic (IRR 1.18, 95% CI 0.97-1.45) and maternity (IRR 1.19, 95% CI 0.98-1.44) confounders. There was no significant association between either diabetes or asthma and total number of admissions (Table 69).

Table 69 Association between chronic disease and total number of admissions

	Univariate model		Multivariate[1]		Multivariate[2]	
	IRR	95% CI	IRR	95% CI	IRR	95% CI
Diabetes	1.22	0.70-2.14	1.36	0.99-1.88	1.37	0.97-1.93
Asthma	0.97	0.82-1.14	1.16	0.99-1.36	1.13	0.98-1.31
Epilepsy	12.56***	7.51-21.00	17.40***	8.65-35.01	14.72***	7.34-29.51
ADHD	0.73*	0.55-0.96	1.18	0.97-1.45	1.19	0.98-1.44
Depression	2.99***	1.56-5.73	3.24***	2.28-4.59	3.31***	2.29-4.78

Exponentiated coefficients; 95% confidence intervals in brackets
* p<0.05, ** p<0.01, *** p<0.001

[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)

[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

There was a significant association between epilepsy and total length of stay in hospital for any cause within the 5 year follow up period on univariate analysis (IRR 6.30, 95% CI 4.51-8.80) and following adjustment for sociodemographic (IRR 6.99, 95% CI 4.66-10.47) and maternity (IRR 7.55, 95% CI 5.24-10.88) confounders. There was also a significant association between depression and total length of stay in hospital for any cause on univariate analysis (IRR 5.53, 95% CI 3.66-8.34) and following adjustment for sociodemographic (IRR 6.54, 95% CI 4.37-9.80) and maternity (IRR 7.11, 95% CI 4.78-10.55) confounders. There was no significant association between diabetes, asthma or ADHD and total length of stay (Table 70). These findings suggest that children with epilepsy or depression may have particularly complex healthcare needs compared to children with the other chronic conditions investigated and place a greater burden on the health service

Table 70 Association between chronic disease and total length of stay

	Univariate model		Multivariate[1]		Multivariate[2]	
	IRR	95% CI	IRR	95% CI	IRR	95% CI
Diabetes	1.41	0.75-2.64	1.52	0.84-2.72	1.51	0.89-2.56
Asthma	0.8	0.63-1.02	0.89	0.71-1.12	0.97	0.78-1.19
Epilepsy	6.30***	4.51-8.80	6.99***	4.66-10.47	7.55***	5.24-10.88
ADHD	0.89	0.56-1.43	1.21	0.79-1.85	1.39	0.85-2.27
Depression	5.53***	3.66-8.34	6.54***	4.37-9.80	7.11***	4.78-10.55

Exponentiated coefficients; 95% confidence intervals in brackets

* p<0.05, ** p<0.01, *** p<0.001

[1]adjusted for sex, age, SIMD deprivation quintile and ethnic group (as appropriate)

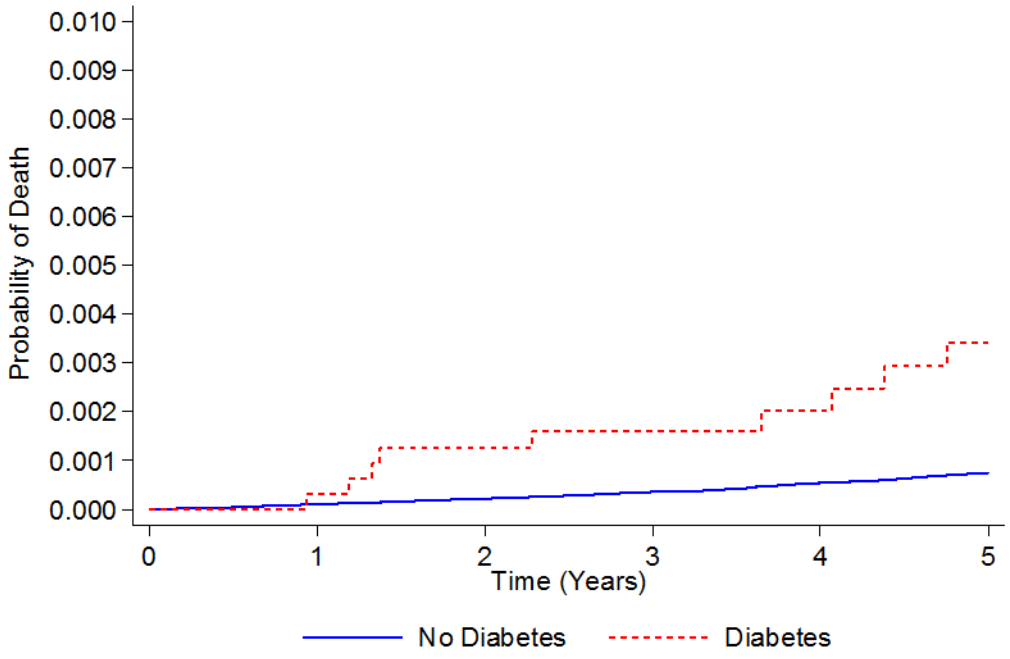
[2]also adjusted for maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

5.10 All-cause mortality

5.10.1 Association between diabetes and mortality

197 pupils who received diabetes drugs which did not meet the diabetes definition outlined in section 4.3.2.2.1 were omitted. Therefore 766,047 pupils were followed up to assess the impact of diabetes on risk of all-cause mortality. The follow up period ranged from 1 year to 5 years and the mean follow up time was 4.34 years. 490 pupils (0.06%) died during the follow up period. Whilst the absolute probability of dying over follow up was very low, the probability was significantly higher for pupils with diabetes compared to those without diabetes (Figure 71).

Figure 71 Kaplan-Meier survival curve: Diabetes and all-cause mortality

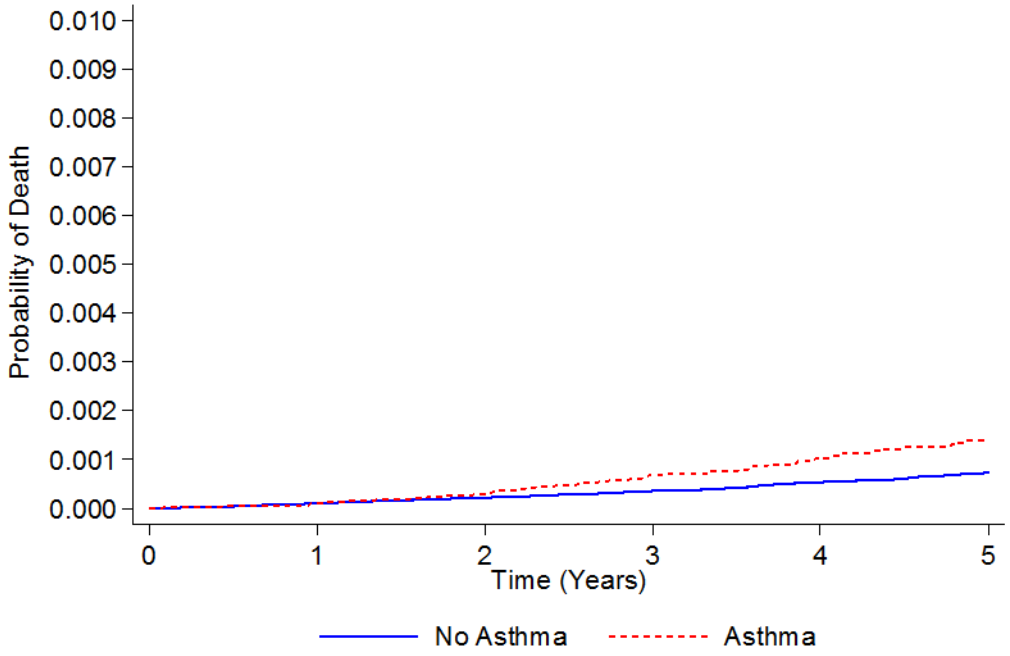


Diabetes was significantly associated with increased risk of mortality on univariate Cox regression analysis (HR 3.73, 95% CI 1.93-7.22) and following adjustment for sociodemographic (HR 3.71, 95% CI 1.92-7.19) and maternity (HR 3.84, 95% CI 1.98-7.43) confounders. The model satisfied the proportional hazards assumption ($p<0.001$); therefore, the association was adequately described using a constant hazard ratio over the full follow up period. Children with diabetes experienced a 3.8 fold risk of death over the five year follow up period compared to children without diabetes.

5.10.2 Association between asthma and mortality

82,528 pupils who received asthma drugs which did not meet the asthma definition outlined in section 4.3.2.2.2 were omitted. Therefore 683,716 pupils were followed up to assess the impact of asthma on risk of all-cause mortality. The follow up period ranged from 1 year to 5 years and the mean follow up time was 4.29 years. 437 pupils (0.06%) died during the follow up period. Whilst the absolute probability of dying over follow up was very low, the probability was significantly higher for pupils with asthma compared to those without asthma (Figure 72).

Figure 72 Kaplan-Meier survival curve: Asthma and all-cause mortality

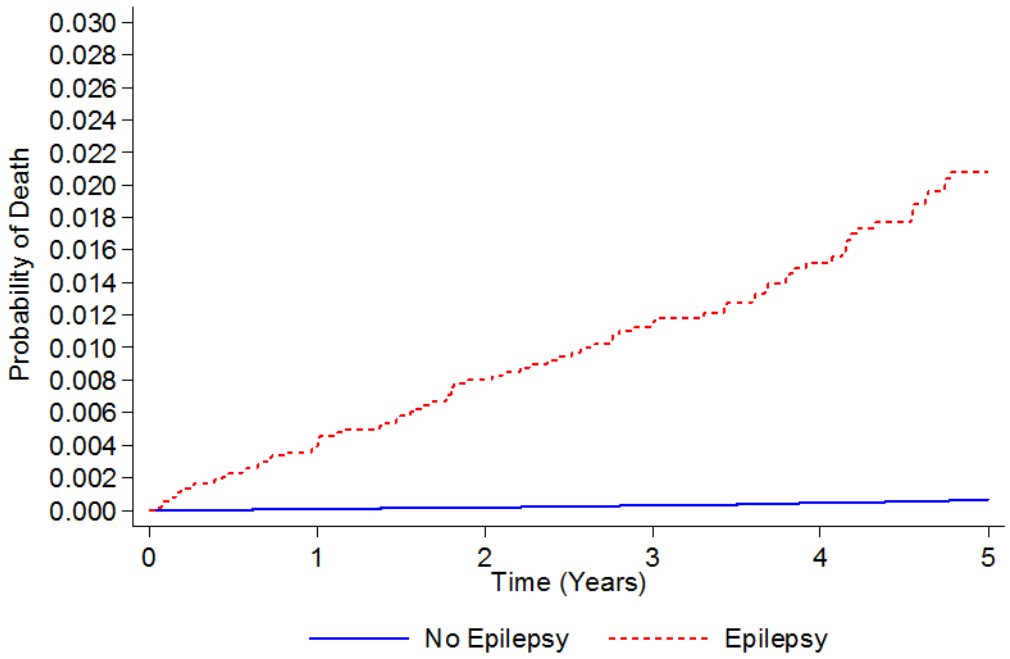


Asthma was significantly associated with increased risk of mortality on univariate Cox regression analysis (HR 1.92, 95% CI 1.42-2.60) and following adjustment for sociodemographic (HR 1.83, 95% CI 1.35-2.48) and maternity (HR 1.77, 95% CI 1.30-2.40) confounders. The model satisfied the proportional hazards assumption ($p<0.001$); therefore, the association was adequately described using a constant hazard ratio across the full follow up period. Children with asthma experienced a 77% increased risk of death in the five year follow up period compared to children without asthma.

5.10.3 Association between epilepsy and mortality

766,244 pupils were followed up to assess the impact of epilepsy on risk of all-cause mortality. The follow up period ranged from 1 year to 5 years and the mean follow up time was 4.33 years. 491 pupils (0.06%) died during the follow up period. Whilst the absolute probability of dying over follow up was very low, the probability was significantly higher for pupils with epilepsy compared to those without epilepsy (Figure 73).

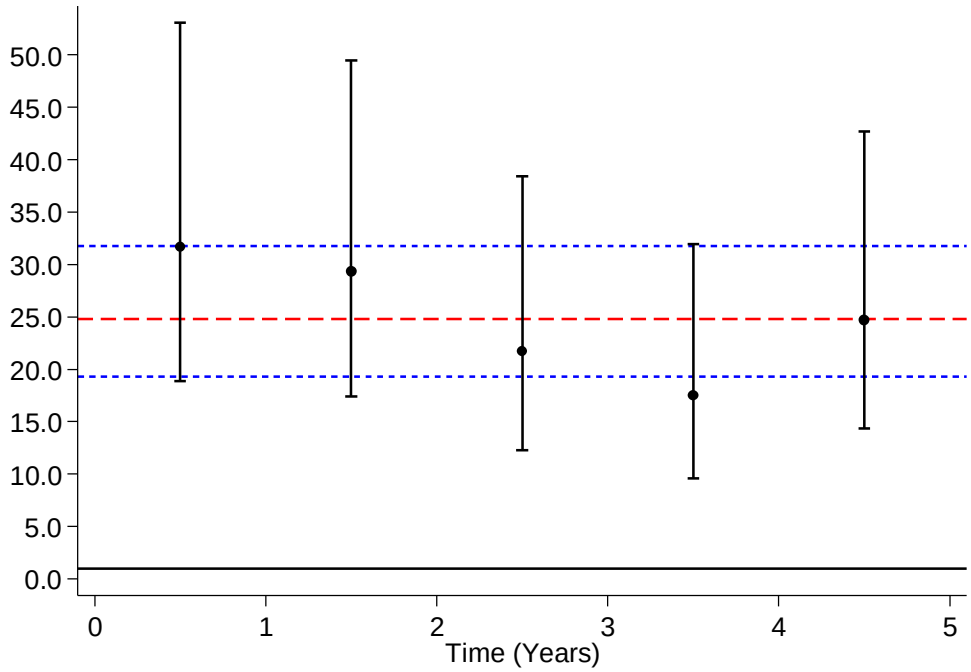
Figure 73 Kaplan-Meier survival curve: Epilepsy and all-cause mortality



Epilepsy was significantly associated with increased risk of mortality on univariate Cox regression analysis (HR 27.80, 95% CI 21.83-35.40) and following adjustment for sociodemographic (HR 27.61, 95% CI 21.68-35.16) and maternity (HR 24.77, 95% CI 19.32-31.77) confounders. However, the model did not satisfy the assumption of proportional hazards ($p<0.001$) and Poisson piecewise regression was used to more accurately model the hazard ratios over the follow up period. Figure 74 presents the fully adjusted incidence rate ratios from the Poisson piecewise model (black solid circles) alongside the static multivariate hazard ratio derived from the Cox proportional hazards model (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines).

The Cox model demonstrated that, over the full follow up period, children with epilepsy experienced an almost 25-fold risk of death compared to non-epileptic pupils. However, the Poisson piecewise model demonstrated that the elevated risk of death experienced by epileptic pupils compared to non-epileptic pupils actually decreased in magnitude with increasing follow-up time. Epileptic children experienced an approximately 30-fold risk of death within the first year of follow-up. However, the excess risk decreased to below 20-fold by four years.

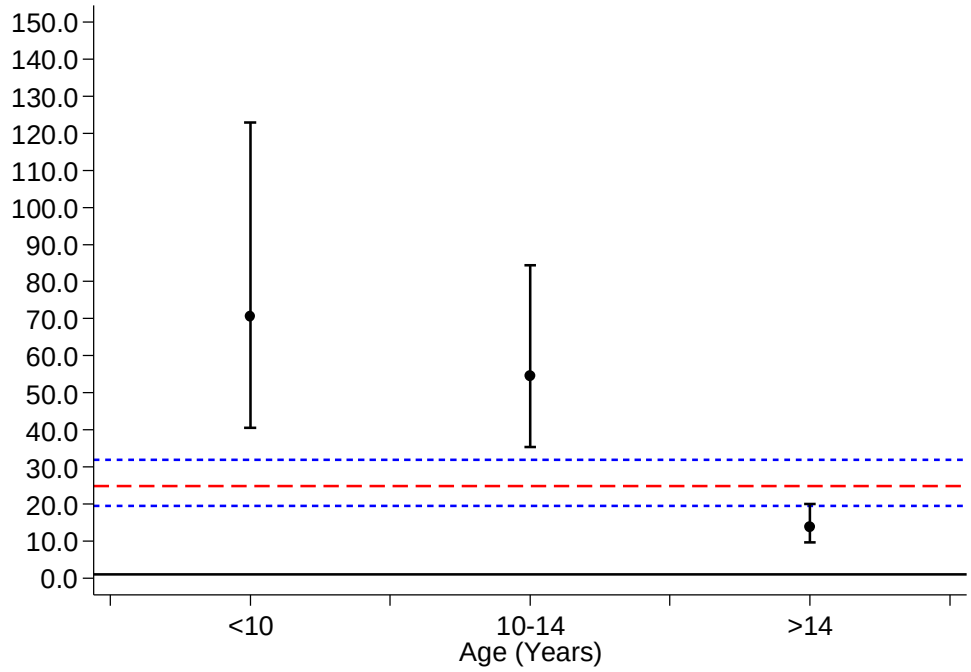
Figure 74 Epilepsy and risk of all-cause mortality modelled as time from diagnosis until death



The risk of death associated with epilepsy also varied depending on pupil age at the time of death. Figure 75 presents the fully adjusted incidence rate ratios from the Poisson piecewise model (black solid circles) alongside the static multivariate hazard ratio derived from the Cox proportional hazards model (red dashed line). Confidence intervals are shown for both models (black error bars and blue dashed lines). Epileptic children of all ages experienced an increased risk of death compared to non-epileptic children; however, there was a distinct trend whereby the association between epilepsy and risk of death became weaker for older children. The magnitude of the excess risk decreased from approximately 70 fold in children aged <10 years to approximately 13 fold in children

aged >14 years. Epilepsy can be more severe in younger children [271] and this may explain the greater risk of death associated with epilepsy in younger pupils.

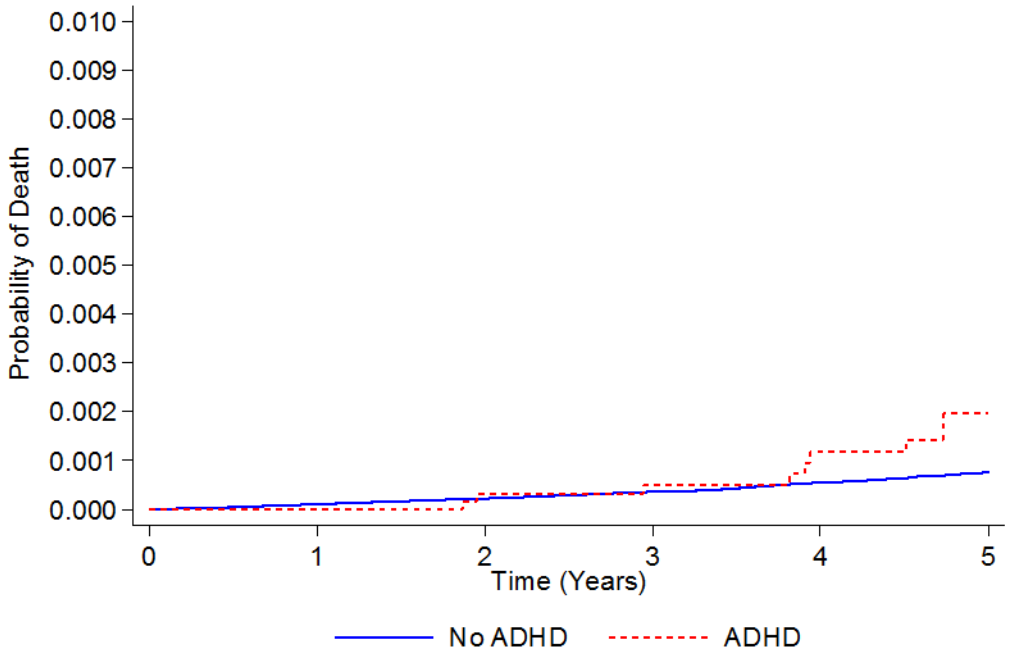
Figure 75 Epilepsy and risk of all-cause mortality modelled as age at death



5.10.4 Association between attention deficit hyperactivity disorder and mortality

766,244 pupils were followed up to assess the impact of ADHD on risk of all-cause mortality. The follow up period ranged from 1 year to 5 years and the mean follow up time was 4.33 years. 491 pupils (0.06%) died during the follow up period. Whilst the absolute probability of dying over follow up was very low, the probability was significantly higher for pupils with ADHD compared to those without ADHD (Figure 76).

Figure 76 Kaplan-Meier survival curve: Attention deficit hyperactivity disorder and all-cause mortality

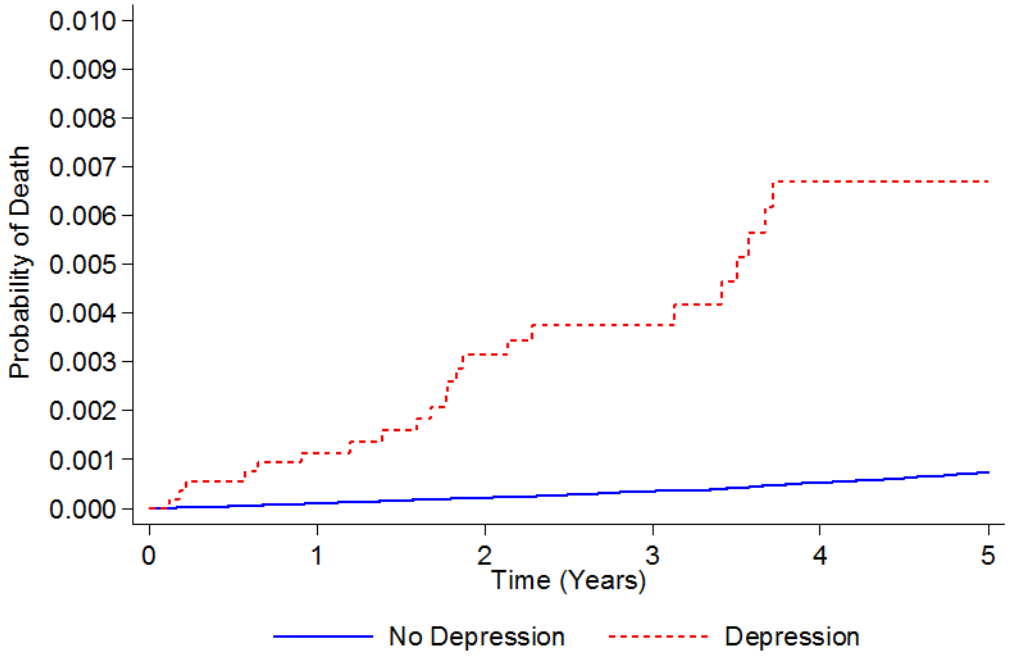


ADHD was significantly associated with increased risk of mortality on univariate Cox regression analysis (HR 2.11, 95% CI 1.09-4.08) but this disappeared after further adjustment for sociodemographic (HR 1.81, 95% CI 0.93-3.51) and maternity (HR 1.62, 95% CI 0.83-3.14) confounders. The model satisfied the proportional hazards assumption ($p<0.001$); therefore, the association was adequately described using a constant hazard ratio across the full follow up period.

5.10.5 Association between depression and mortality

7 pupils who received anti-depressant drugs which did not meet the depression definition outlined in section 4.3.2.2.5 were omitted. Therefore 766,237 pupils were followed up to assess the impact of depression on risk of all-cause mortality. The follow up period ranged from 1 year to 5 years and the mean follow up time was 4.33 years. 491 pupils (0.06%) died during the follow up period. Whilst the absolute probability of dying over follow up was very low, the probability was significantly higher for pupils with depression compared to those without depression (Figure 77).

Figure 77 Kaplan-Meier survival curve: Depression and all-cause mortality



Depression was significantly associated with increased risk of mortality on univariate Cox regression analysis (HR 5.79, 95% CI 3.73-8.99) and following adjustment for sociodemographic (HR 6.24, 95% CI 4.02-9.70) and maternity (HR 6.21, 95% CI 4.00-9.65) confounders. The model satisfied the proportional hazards assumption ($p<0.001$); therefore, the association was adequately described using a constant hazard ratio across the full follow up period. Children with depression had a 6.21 times greater risk of death over the five year follow up period compared to children without depression.

5.11 Association between each chronic condition and each health outcome after adjusting for other comorbid chronic conditions

Some children experienced more than one condition; therefore, the final fully adjusted models investigating risk of all-cause hospital admission and all-cause mortality were run again including all conditions together. The first two columns in Table 71 and Table 72 present incidence rate ratios and 95% confidence intervals from Poisson piecewise regression models, adjusted for sociodemographic and maternity confounders, investigating associations between each chronic condition and risk of all-cause hospital admission. The third and fourth columns in each table present the same information after adjusting each condition-specific model for presence of any of the other chronic conditions identified in this thesis. Table 71 presents data from regressions modelled by time from diagnosis until first admission. Table 72 presents data from regressions modelled by age at first admission. Adjusting for other chronic conditions had little impact on the effect size of the associations between either diabetes or asthma and risk of all-cause hospital admission. The associations between both epilepsy and ADHD and risk of all-cause hospital admission were attenuated after adjusting for other chronic conditions. However, the association between depression and risk of all-cause hospital admission was attenuated the most after adjusting for presence of other conditions. All of the attenuated associations remained significant.

Table 71 Association between each chronic condition and risk of all-cause hospital admission after adjusting for presence of other comorbid chronic conditions modelled as time from diagnosis until first admission

	multivariate model[1]		adjusted for comorbid conditions[2]	
	IRR	95% CI	IRR	95% CI
Diabetes				
0 - 1 years	4.43***	4.13-4.75	4.33***	4.03-4.64
1 - 2 years	4.23***	3.85-4.65	4.21***	3.83-4.63
2 - 3 years	3.47***	3.06-3.93	3.49***	3.08-3.95
3 - 4 years	3.41***	2.94-3.95	3.45***	2.98-4.00
4 - 5 years	2.92***	2.44-3.50	2.94***	2.45-3.53
Asthma				
0 - 1 years	1.98***	1.93-2.04	1.95***	1.90-2.01
1 - 2 years	1.69***	1.63-1.76	1.67***	1.60-1.73
2 - 3 years	1.49***	1.42-1.56	1.47***	1.40-1.54
3 - 4 years	1.47***	1.39-1.55	1.45***	1.37-1.53
4 - 5 years	1.38***	1.30-1.47	1.36***	1.28-1.45
Epilepsy				
0 - 1 years	5.52***	5.25-5.81	4.87***	4.63-5.13
1 - 2 years	2.93***	2.68-3.21	2.62***	2.39-2.87
2 - 3 years	2.52***	2.24-2.83	2.30***	2.04-2.58
3 - 4 years	2.23***	1.92-2.58	2.06***	1.78-2.38
4 - 5 years	1.92***	1.60-2.31	1.80***	1.49-2.16
ADHD				
1 - 5 years	1.33***	1.27-1.39	1.25***	1.19-1.31
Depression				
0 - 1 years	3.21***	3.01-3.41	2.57***	2.42-2.74
1 - 2 years	2.12***	1.92-2.35	1.73***	1.56-1.91
2 - 3 years	1.94***	1.69-2.22	1.58***	1.38-1.81
3 - 4 years	2.08***	1.75-2.47	1.72***	1.44-2.04
4 - 5 years	1.69***	1.30-2.19	1.40*	1.08-1.81

Exponentiated coefficients; 95% confidence intervals in brackets
** p<0.05, ** p<0.01, *** p<0.001*

[1]adjusted for sex, age, SIMD deprivation quintile, ethnic group, maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

[2] also adjusted for diabetes, asthma, epilepsy, ADHD and depression (as appropriate)

Table 72 Association between each chronic condition and risk of all-cause hospital admission after adjusting for presence of other comorbid chronic conditions modelled by age at first admission

	multivariate model[1]		adjusted for comorbid conditions[2]	
	IRR	95% CI	IRR	95% CI
Diabetes				
4 - 6 years	4.42***	3.75-5.22	4.44***	3.76-5.24
7 - 8 years	4.12***	3.46-4.90	4.05***	3.41-4.82
9 - 10 years	5.20***	4.53-5.97	5.23***	4.56-6.01
11 - 12 years	5.25***	4.67-5.91	5.23***	4.64-5.88
13 - 14 years	5.14***	4.65-5.70	5.07***	4.58-5.61
15 - 16 years	3.79***	3.40-4.22	3.73***	3.35-4.15
17 - 23 years	2.89***	2.59-3.24	2.88***	2.57-3.22
Asthma				
4 - 6 years	2.15***	2.06-2.25	2.13***	2.04-2.23
7 - 8 years	2.01***	1.91-2.11	1.99***	1.90-2.09
9 - 10 years	1.94***	1.84-2.04	1.93***	1.83-2.03
11 - 12 years	1.84***	1.75-1.94	1.82***	1.73-1.92
13 - 14 years	1.61***	1.54-1.69	1.58***	1.51-1.66
15 - 16 years	1.43***	1.37-1.51	1.39***	1.33-1.47
17 - 23 years	1.31***	1.25-1.38	1.27***	1.20-1.33
Epilepsy				
4 - 6 years	7.11***	6.45-7.83	6.87***	6.23-7.58
7 - 8 years	5.28***	4.67-5.97	5.09***	4.50-5.75
9 - 10 years	5.26***	4.68-5.91	5.05***	4.49-5.67
11 - 12 years	4.86***	4.36-5.41	4.51***	4.05-5.03
13 - 14 years	4.31***	3.91-4.75	3.80***	3.45-4.19
15 - 16 years	3.62***	3.32-3.96	3.09***	2.83-3.38
17 - 23 years	2.05***	1.85-2.26	1.77***	1.60-1.95
Depression				
4 - 6 years	3.46***	2.05-5.85	2.81***	1.66-4.75
7 - 8 years	2.09**	1.32-3.32	1.64*	1.03-2.60
9 - 10 years	2.79***	2.07-3.76	2.18***	1.61-2.94
11 - 12 years	3.24***	2.60-4.05	2.45***	1.96-3.07
13 - 14 years	4.39***	3.88-4.96	3.28***	2.90-3.71
15 - 16 years	3.73***	3.46-4.02	2.91***	2.70-3.14
17 - 23 years	2.18***	2.02-2.35	1.83***	1.69-1.97

Exponentiated coefficients; 95% confidence intervals in brackets
** p<0.05, ** p<0.01, *** p<0.001*

[1]adjusted for sex, age, SIMD deprivation quintile, ethnic group, maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

[2] also adjusted for diabetes, asthma, epilespy, ADHD and depression (as appropriate)

The first two columns in Table 73 present incidence rate ratios and 95% confidence intervals from Poisson piecewise regression models, adjusted for sociodemographic and maternity confounders, investigating associations between each chronic condition and risk of all-cause mortality. The third and fourth columns present the same information after adjusting each condition-specific model for presence of any of the other chronic conditions identified in this thesis. Adjusting for other chronic conditions attenuated the effect sizes of the associations between diabetes, asthma, epilepsy and depression and all-cause mortality. However all of these associations remained significant. The reduced effect size of the association between depression and all-cause mortality was particularly notable. Depressed children experienced a 6.2 times greater risk of death compared to pupils without depression; however, the excess risk reduced to 2.7 fold after adjusting for presence of other conditions.

Table 73 Association between each chronic condition and risk of all-cause mortality after adjusting for presence of other comorbid chronic conditions modelled as time from diagnosis until death

	multivariate model[1]		adjusted for comorbid conditions[2]	
	IRR	95% CI	IRR	95% CI
Diabetes				
1 - 5 years	3.84***	1.98-7.43	2.90**	1.49-5.66
Asthma				
1 - 5 years	1.77***	1.30-2.40	1.61**	1.18-2.19
Epilepsy				
0 - 1 years	31.65***	18.87-53.08	27.55***	16.31-46.54
1 - 2 years	29.33***	17.39-49.45	25.67***	15.13-43.57
2 - 3 years	21.72***	12.28-38.42	19.30***	10.86-34.29
3 - 4 years	17.50***	9.58-31.97	15.77***	8.61-28.91
4 - 5 years	24.73***	14.33-42.70	22.69***	13.11-39.28
Depression				
1 - 5 years	6.21***	4.00-9.65	2.72***	1.72-4.30

Exponentiated coefficients; 95% confidence intervals in brackets
** p<0.05, ** p<0.01, *** p<0.001*

[1]adjusted for sex, age, SIMD deprivation quintile, ethnic group, maternal age, maternal smoking during pregnancy, parity, mode of delivery, gestational age at delivery, sex-gestation-specific birthweight centile and 5 minute Apgar score

[2] also adjusted for diabetes, asthma, epilepsy, ADHD and depression (as appropriate)

6 Discussion

6.1 Conclusions and discussion of results

All five of the medical conditions investigated in this thesis were associated with adverse educational and health outcomes. The number of outcomes affected varied by condition. Table 74 presents a summary of the associations between each of the five chronic conditions and each of the nine outcomes of interest. It highlights whether there was an overall association (Y/N) and, where investigated interactions were significant, summarises how the strength of the association differed within sex, age and deprivation subgroups. Table 75 summarises the effect size of the main association for each of the five conditions against each of the nine outcomes. The effect size is summarised as a range in instances where it varied because the underlying model did not satisfy proportional odds or proportional hazards.

Table 74 Summary of the associations between each chronic condition and each outcome

	Diabetes		Asthma		Epilepsy		ADHD		Depression	
outcomes	Sig Assoc.	strength	Sig Assoc.	strength	Sig Assoc.	strength	Sig Assoc.	strength	Sig Assoc.	strength
Absence [1 - 3]	Y		Y	young > old affluent > deprived	Y	girls > boys young > old affluent > deprived	Y	old > young affluent > deprived	Y	old > young affluent > deprived
Exclusion [1 - 3]	N		Y		N		Y	girls > boys young > old affluent > deprived	Y	young > old
SEN [1 - 3]	Y	girls > boys young > old affluent > deprived	Y	girls > boys old > young	Y	girls > boys young > old affluent > deprived	Y	girls > boys old > young affluent > deprived	Y	boys > girls young and old > 11-14 affluent > deprived
Attainment [1 - 4]	N		Y [5, 6]		Y [5, 6]		Y [5]	girls > boys	Y [5, 6]	
Unemployment [1 - 4]	N		N		Y [5, 6]	affluent > deprived	Y [5]	girls > boys affluent > deprived	Y [6]	boys > girls affluent > deprived
Admission [1]	Y	girls > boys	Y		Y		Y	girls > boys	Y	girls > boys
Total admissions	N		N		Y		N		Y	
Total length of stay	N		N		Y		N		Y	
Mortality	Y		Y		Y		N		Y	

[1] investigated gender interaction
[2] investigated age ineraction
[3] investigated deprivation interaction
[4] investigated outcomes after excluding children with SEN and after adjusting for absenteeism
[5] association was attenuated after excluding children with SEN from the analyses
[6] association was attenuated after adjusting for absenteeism

Table 75 Summary of the main effect size between each chronic condition and each outcome

	Diabetes	Asthma	Epilepsy	ADHD	Depression
outcome	effect size	effect size	effect size	effect size	effect size
Absence [1]	1.34	1.25	1.50	1.18	1.95
Exclusion [1]	NS	0.91	NS	5.82	1.65
SEN [2]	2.45	1.28	10.11	8.77	2.24
Attainment [3]	NS	1.11	1.98 - 4.07	3.64	1.96 - 3.44
mediating for absenteeism		NS	1.23 - 3.07	4.07	NS - 1.89
omitting children with SEN		1.08	1.70 - 3.29	2.79	1.83 - 3.28
Unemployment [4]	NS	NS	1.99	1.42	1.98
mediating for absenteeism			1.70	1.44	1.42
omitting children with SEN			1.49	1.30	2.02
Admission [5]	2.92 - 4.43	1.38 - 1.98	1.92 - 5.52	1.33	1.69 - 3.21
Total admissions [6]	NS	NS	14.72	NS	3.31
Total length of stay [6]	NS	NS	7.55	NS	7.11
Mortality [5]	3.84	1.77	24.73 - 31.65	NS	6.21

[1] effect sizes are incidence rate ratios from GEE analyses with a negative binomial distribution and log link

[2] effect sizes are odds ratios from GEE analyses with a binomial distribution and logit link

[3] effect sizes are odds ratios from generalised ordinal regression analyses (range shows highest and lowest odds ratio where odds were non-proportional)

[4] effect sizes are odds ratios from binomial logistic regression analyses

[5] effect sizes are hazard ratios from proportional hazards Cox regression analyses (single number) or incidence rate ratios from Poisson piecewise regression analyses (range shows lowest and highest incidence rate ratio where hazards were non-proportional)

[6] effect sizes are incidence ratios from negative binomial regression analyses

NS = not significant

Depression, epilepsy and ADHD were associated with the most wide-ranging impact. Children treated for depression fared worse than their peers across all nine outcomes, and children treated for epilepsy and ADHD across eight and seven respectively. In contrast, children treated for asthma and diabetes fared worse than their peers in respect of around half the outcomes investigated.

Children treated for all five conditions had more frequent absenteeism from school and were more likely than their peers to be recorded as having SEN. However, only children treated for depression, epilepsy or ADHD experienced poorer academic attainment and increased risk of unemployment. Children treated for asthma also experienced poorer academic attainment but no increased risk of unemployment. Furthermore, children treated for depression or ADHD were significantly more likely to be excluded from school. Children treated for all of the five conditions were at higher risk of hospital admission and mortality; however, children treated for depression or epilepsy were also at higher risk of recurrent hospitalisation and longer periods of hospitalisation.

Some children experienced more than one condition; therefore, final statistical models included all diseases together. This attenuated the effect sizes within each model but did not change any of the associations initially observed. Hence, these conditions were associated with adverse outcomes independently of each other.

6.1.1 Epilepsy and depression

Depression and epilepsy were associated with poorer outcomes across all of the educational and health domains investigated with the exception of epilepsy and school exclusion. Therefore, these conditions were particularly detrimental to both the short and long-term aspects of a child's schooling and health. Both conditions were associated with increased SEN compared to peers although the association was stronger for epilepsy than for depression. Children with epilepsy or depression often experience comorbid behavioural and learning difficulties [114,117,124,282,283] and SEN [121] and, consistent with this, significant associations with all types of SEN were observed; most notably those attributed to social and behavioural difficulty, learning disability and autism. Whilst learning problems may lead directly to poor academic attainment, social and behavioural problems are likely to result in poorer conduct and increased school exclusion [284]. In this thesis, depressed children experienced increased risk of exclusion from school compared to peers; however, there was no such association with epilepsy. The association between depression and exclusion remained significant after adjusting for the other chronic conditions investigated, including comorbid ADHD which frequently presents alongside depression. However, in addition to ADHD, paediatric depression may present with a range of other behavioural comorbidities such as oppositional defiant disorder, conduct disorder or anxiety [285] and the presence of these unobserved comorbidities may have contributed to the increased risk of exclusion. Additionally, the definition of depression used in this thesis may have introduced selection bias because only children prescribed medication for depression, who were possibly more severe cases with more comorbid behaviour disorders, were identified.

Nevertheless, confirming the exact reasons for increased school exclusion associated with depression requires further investigation. It was perhaps surprising that, whilst epilepsy was also associated with increased behavioural difficulty, this did not similarly manifest in an increased risk of exclusion from school, particularly as children with epilepsy also commonly experience psychiatric comorbidity. This may be because children with epilepsy tend to suffer internalising psychiatric comorbidities such as depression and

anxiety as opposed to externalising ones such as oppositional defiant disorder or conduct disorder [286,287]. It is also likely that school services are more aware of children with a diagnosis of epilepsy rather than depression; evidenced by epilepsy being more strongly associated with SEN than any of the other chronic conditions investigated. Therefore children with epilepsy may receive more positive behavioural help throughout school than their peers with depression who go unnoticed until exclusions occur. The intricate relationships between epilepsy, depression and school exclusion definitely requires further consideration.

Compared to their peers, I demonstrated that children with depression or epilepsy also experienced significantly poorer academic attainment and a higher risk of subsequent unemployment after leaving school. Notably, these children not only had increased risk of poorer attainment overall, but were at particular risk of achieving grades specifically classed in the lowest category. Furthermore, even though SEN played a role as a mediator in the associations between epilepsy and academic attainment and unemployment, the associations remained significant after children with SEN were excluded from the cohort. The association between depression and attainment was also attenuated, albeit to a lesser extent, but remained significant after excluding children with SEN. This suggests that some children who have epilepsy or depression and who are not recorded as having SEN may, nonetheless, be in need of additional educational support which they are not currently receiving. These findings are consistent with previous research reporting poorer attainment in children with epilepsy [112-120] and depression [224-226]. Previous depression studies have similarly identified poorer leaver outcomes [14,15,213-215]

Absenteeism also plays a major role in the observed poorer school performance of children with epilepsy or depression. Increased absenteeism was observed for children with one of these conditions compared to peers although the association was strongest for children with depression. Inclusion of absenteeism attenuated the association between epilepsy and both attainment and unemployment. Therefore, the higher risk of poor attainment and unemployment among children with epilepsy was explained, in part, by their higher rates of absenteeism. The influence of absenteeism on school outcomes was even greater for children with depression. After adjusting for absenteeism, the association between depression and unemployment was heavily attenuated but remained significant. However, the association with attainment largely disappeared whereby only the lowest category of attainment remained significant. Therefore, absenteeism significantly influences poor attainment and future unemployment of children diagnosed with depression. Targeting

absenteeism, either through preventative measures or via interventions to assist with catching up with missed education, should be investigated to improve their academic grades and school leaver outcomes.

It is important to understand why children with epilepsy or depression are absent from school. In addition to suffering increased risk of any hospital admission, I observed that these children also experienced higher risk of recurrent hospitalisation and longer periods of hospitalisation compared to peers. This may indicate that they have particularly complex healthcare needs, which may have contributed to increased absenteeism. In the case of depression, it is possible that truancy or school refusal may additionally contribute to missed days from school as reported elsewhere [288]. Specific reasons for absenteeism are recorded by schools and should therefore be further investigated. Further analyses of the types of SEN should also be undertaken to understand whether identification of depressed and epileptic children within school services is appropriate and timely.

Children with depression or epilepsy experienced an increased risk of all-cause mortality and all-cause hospital admission within 5 years of their first known prescription date compared with peers. I also observed increased risk of injury and mental health admissions for children with depression and increased risk of admission related to the nervous system for children with epilepsy. The magnitude of the increased risk of hospital admission dropped sharply after the first year of follow-up for both conditions and continued to fall thereafter. This indicated that these children are most at risk of hospital admission around the time of diagnosis of their condition. The risk of mortality associated with epilepsy also became weaker over the follow up period. The relative increased risk of hospitalisation and mortality may reduce as children are stabilised on their respective medication and symptoms are better controlled [273]. Therefore, interventions which provide additional support to children during this specific early phase may be useful. Several studies have previously shown a strong association between epilepsy and risk of mortality [140-147]. The observed association between depression and injury admission is also consistent with existing literature [230,231].

Prevalence of epilepsy and depression observed in this thesis was 0.69% and 0.70% respectively. The prevalence of epilepsy is roughly consistent with 1% prevalence in children and adolescents reported elsewhere [106,107]. Worldwide prevalence of paediatric depression has been reported to be 2.8% [206]. Smaller studies have reported prevalence as high as 8% and 13% based on depressive symptoms. My thesis only identified children

actively receiving medication for depression which explains the lower percentage observed.

6.1.1.1 Differences by age, gender and deprivation

Although both epilepsy and depression were associated with greater social deprivation, the relative associations with absenteeism, SEN and unemployment were strongest for the least deprived pupils. This finding was explained by more deprived unaffected pupils already having higher absolute rates of SEN, absenteeism and unemployment. Therefore, the relative impact of epilepsy or depression on their outcomes was less pronounced. This highlights the persistence of long standing socioeconomic inequalities in health and education. The relative impact of epilepsy on SEN was stronger for girls and this was explained because, among unaffected children, boys had higher absolute rates of SEN. Despite higher absenteeism in girls among unaffected children, epilepsy still exerted a greater relative impact on girls than boys with respect to absenteeism.

Despite boys unaffected by depression already having higher SEN and more unemployment than unaffected girls, the relative impact of depression on SEN and unemployment was still greater for boys. This could indicate that depressed boys are more likely than depressed girls to receive special education. It is therefore important to understand why this is happening. Depressed boys may possibly present with more obvious symptoms than depressed girls perhaps due to presence of more comorbid conditions, increased disease severity or differences in behaviour patterns. Depressed boys tend to express violence toward themselves or others, show self-destructive behaviour, and misuse substances. Girls on the other hand tend to withdraw themselves from family and friends, admit feelings of hopelessness, lack motivation, show changes in sleeping patterns, and develop eating disorders. Depressed girls are also more likely to self-harm [289]. Therefore, despite being identified by health services and receiving treatment for their depression, girls may be less likely to be identified as depressed by their school and receive less help. Schools may also provide less input for girls because their presentation is perceived as less severe or because it is less disruptive to the school environment.

The observed relative impact of depression on risk of all-cause and injury-related hospital admission was greater for girls. The latter can be explained because, among unaffected children, boys had more injury admission. However the rate of all-cause hospital admission was similar for unaffected boys and girls. Therefore it is important to understand why

depression had a greater relative impact on girls as regards all-cause hospital admission. The relative impact on both outcomes was also greater for children between 11 and 16 years of age and decreased with time from diagnosis. This may reflect different behaviour patterns in teenage years such as increased instance of self-harm resulting in more admissions and more missed school days particularly amongst depressed girls and particularly around time of diagnosis. Again these findings may be partly explained by a combination of specific increased self-harm in depressed adolescents [277] and additional excess self-harm in adolescent girls [275,276,290] and may reflect the reported higher onset of depression in teenage girls with prevalence becoming twice as high in girls by early to mid-adolescence [278-281]. Teenage girls with depression should therefore be monitored particularly closely for increased risk of self-harm. It is interesting that depression impacted boys more than girls with respect to unemployment, yet there were no significant differences between genders in terms of the impact of depression on attainment, absenteeism or exclusion. The observed excess risk of unemployment in depressed boys may be due to the types of associated comorbidities and behaviours they experience. However, without further investigation, this is purely speculation.

The relative impact of depression on absenteeism and exclusion was strongest for older and younger pupils respectively. Depressed pupils who are older may be more likely to miss school through truancy or school refusal and therefore experience more absenteeism yet less opportunity for exclusion [288]. Conversely, younger depressed pupils may be more closely monitored by their parent and therefore be in school but with resulting opportunity for exclusion. Additionally, pupils medicated at a younger age may have more severe depression with a more complex range of behavioural comorbidities therefore manifesting in greater exclusion.

The association of epilepsy with absenteeism and SEN was strongest in younger pupils. Older children who were unaffected by epilepsy already experienced more absenteeism; therefore the relative impact of epilepsy was smaller in this age group. Epilepsy presenting at an early age may also be more severe and disruptive due to increased association with congenital neurological disability [271] and this may have resulted in the stronger observed association between epilepsy and both absenteeism and SEN in younger pupils. Whilst children with epilepsy experienced increased risk of all-cause hospital admission, admission related to the nervous system and all-cause mortality across the full follow up period, the relative impact on these outcomes was greatest close to time of diagnosis and for younger children. This is again consistent with more severe epilepsy in younger

children and may also indicate that the excess risk decreases as children become stabilised on their medication [273]. The association between epilepsy and admission related to the nervous system was stronger for boys than girls.

6.1.2 Attention deficit hyperactivity disorder

ADHD also severely impacted the educational and health outcomes of affected children. A diagnosis of ADHD had as detrimental an impact on school outcomes as depression, adversely affecting all five of the educational domains assessed. Affected children experienced an increased risk of all types of SEN compared to peers; most notably SEN attributed to social and behavioural difficulty, learning disability and autism. The observed associations with SEN [166-168,176,179,180] and learning disability [168,176] were consistent with previous studies. ADHD is primarily associated with increased impulsivity and hyperactivity, which can result in poor conduct and behaviour [13,166,173,179,180,195,291-296]. Children with ADHD can additionally present with a range of behavioural comorbidities such as anxiety, depression, conduct disorder and oppositional defiant disorder [154,297,298]. Therefore associations with all types of SEN were consistent with this and it was not surprising that I also observed a strong association between ADHD and increased school exclusion. Indeed, consistent with previous research [179,180], ADHD was more strongly associated with exclusion than any of the other conditions investigated.

Also consistent with previous literature observing poorer academic performance [166-168,170,172-175], attendance [172,182] and leaver outcomes [167,172,181], my thesis demonstrated that children with ADHD additionally missed more days of school compared to peers and suffered longer term adversity in the form of poorer school grades and greater unemployment. The associations with attainment and unemployment were attenuated but remained significant after excluding children with SEN. However, unlike children with depression or epilepsy, increased absenteeism did not appear to significantly contribute to the poorer long-term outcomes of children with ADHD. Indeed, adjusting for absenteeism had no tangible effect on the strength of the associations between ADHD and either attainment or unemployment. Therefore, children with ADHD did not experience poorer attainment and higher unemployment compared to their peers simply as a result of increased absenteeism. Notably, the association between ADHD and absenteeism was weaker than the corresponding associations between absenteeism and either depression or epilepsy. A possible explanation may be that children with ADHD have absences due to truancy or school refusal, which are more common yet shorter in duration, and therefore

have less additional impact on their grades over and above that directly due to their condition. Conditions that commonly present alongside ADHD have been associated with increased truancy and school refusal [288]. By comparison, children with depression or epilepsy may experience longer durations of absenteeism due to more complex needs which have a greater effect on their schooling. This, of course, is speculation and should be confirmed by further investigation.

Despite absenteeism having less of an impact on attainment and risk of unemployment, these long-term outcomes of children with ADHD were still poorer when compared to their peers. Therefore, it seems that previously reported behavioural and learning difficulties [168,173,176] primarily influence this observed deficit. Children with ADHD suffer two main types of behavioural problems both of which are equally detrimental to a child's ability to learn and assimilate information: hyperactivity/impulsivity and inattention [148,149]. Whilst hyperactivity and impulsivity are more commonly associated with unacceptable behaviour and unwillingness to learn, inattention may go more unnoticed in the classroom yet can have equally adverse effects on learning.

My thesis additionally demonstrated that children with ADHD also experienced poorer health outcomes in the form of increased risk of all-cause hospital admission and admission due to injury. However, consistent with previous literature [186,187], children with ADHD did not have an excess risk of all-cause mortality compared to peers. My findings are also consistent with previous studies reporting increased injury-related hospital admission [188,192-194]. Prevalence of ADHD in children has been reported to be around 5% both worldwide [159] and in Scotland [149]. However, prevalence of treated ADHD in Scotland ranges between 0.2% and 1.2% [149]. The 0.97% prevalence observed in my thesis was consistent with these figures.

6.1.2.1 Differences by age, gender and deprivation

The associations between ADHD and absenteeism, SEN, exclusion and unemployment were stronger among less deprived pupils. This can again be explained by the fact that more deprived unaffected children already had more absenteeism, exclusion, SEN and higher unemployment; therefore, the relative impact of ADHD was greater in less deprived pupils. Similarly, the relative impact of ADHD on exclusion, SEN, attainment and risk of all-cause and injury-related hospital admission was greater for girls than boys. For unaffected children, boys had more exclusion and SEN, poorer attainment and higher

absolute rates of injury admission. Therefore, the relative impact of ADHD on these outcomes was greater in girls. However, unaffected boys and girls had similar baseline rates of all-cause hospital admission. Therefore, it is important to understand why ADHD had a greater relative impact on girls with respect to risk of all-cause hospital admission. Whilst the increased risk of all-cause hospital admission remained constant across all ages, the relative risk of injury admission for children with ADHD compared to peers peaked in late teenage years. This may be partly explained by a combination of specific increased self-harm in children and adolescents with ADHD [274] and general increased self-harm in adolescence, particularly in girls [275,276].

The relative impact of ADHD on absenteeism and SEN was greater in older pupils whilst the relative impact on exclusions was greater in younger pupils. Similar patterns were observed in children with depression and it is possible that these common patterns are attributable to the same underlying mechanisms. Notably, that the increased risk of injury admission in late teenage years described above likely contributes to missed school days within this age group. Additionally, pupils with ADHD who are older may be more likely miss school through truancy or school refusal and therefore experience more absenteeism yet less opportunity for exclusion. On the other hand younger pupils may be more closely monitored by their parents and be in school with greater resulting opportunity for exclusion. It is also possible that younger children with ADHD may experience more exclusion because teachers are less aware of their condition and exclude them because they appear disruptive and poorly behaved rather than pupils with a chronic condition. As these pupils progress through school, their condition may become more apparent to school services resulting in more SEN and less exclusion, which is consistent with the observed greater relative impact of ADHD on SEN in older children. It is worth investigating the relationships between parents, clinicians and teachers in order to explore whether school teachers are made aware of ADHD diagnoses at the earliest opportunity and whether earlier communication and interventions could reduce some of the adverse outcomes.

6.1.3 Diabetes and asthma

Children with diabetes or asthma experienced a narrower range of adverse outcomes. Whilst this thesis demonstrated that children with diabetes experienced greater absenteeism and increased SEN compared to peers, they did not experience significantly poorer academic attainment or a significantly increased risk of exclusion or unemployment. Therefore, the short-term impact of greater absenteeism did not translate into long-term

adverse outcomes. The association between diabetes and increased absenteeism reported in my thesis is consistent with previous research [29,30,33,35,38,39]. Whilst some previous studies have reported poorer academic performance of diabetic children compared to peers or siblings [9,10,29-33] others, consistent with this thesis, have reported no difference [34-37].

In contrast with the other chronic conditions investigated, diabetes was only associated with SEN attributed specifically to learning difficulty, physical motor disability or a physical health condition. The association between diabetes and learning difficulty but not learning disability is notable. This could indicate that, whilst experiencing difficulty, diabetic children have less severe learning impairments compared to children with the other chronic conditions of interest. Their difficulties may therefore largely arise due to their increased absence and missed learning which may explain in part why long-term attainment and employment is largely unaffected by their condition. However, it is important to understand fully the reasons why diabetic children fair better than children with other chronic conditions investigated in this thesis. The findings may reflect a real difference because children with diabetes are likely to be unwell when first diagnosed with diabetes. It may take some time to stabilise them on insulin and for them, and their parents, to become confident in managing the condition. However, thereafter, with good control of their condition, most diabetic children can avoid future symptoms such as hypoglycaemic episodes and hospitalisations. Additionally, compared to children with ADHD, epilepsy or depression, children with diabetes suffer less comorbidity; in particular less psychiatric comorbidity although they do suffer increased risk of depression and cognitive disturbance compared to their non-diabetic peers [299]. A third explanation may be that teachers are more empathetic to physical problems rather than mental or behavioural problems. Alternatively, the difference may be partially attributed to selection bias in the ascertainment of these conditions. Children with type 1 diabetes require insulin; therefore, all children with type 1 diabetes will receive medication. By comparison, only children with more severe asthma, depression, ADHD or epilepsy may require medication for their condition. It is possible that identification of more severe cases may have inflated the effect sizes of the associations for the other conditions but not diabetes.

I demonstrated that pupils with asthma experienced greater absenteeism and increased risk of all types of SEN compared to peers. The observed association with increased absenteeism is consistent with all but one [78] of the previous comparable studies identified [69,70,73,79,81-85,88]. Previous research has also observed increased use of special education services [70] and increased learning disability among asthmatic children [84]. Whilst asthmatic

children in this thesis experienced significantly poorer attainment compared to peers, they did not, however, experience a significantly increased risk of unemployment. Furthermore, the association between asthma and attainment was slightly attenuated after excluding children with SEN but disappeared completely after adjusting for absenteeism. It was therefore apparent that the poorer attainment of asthmatic children compared to peers was largely attributable to their increased absenteeism. Similar to depressed children, developing interventions to reduce absenteeism or mitigate their impact could radically improve the academic grades of asthmatic children. Whilst some previous studies have reported poorer attainment for asthmatic children [70,79], others have reported no difference in performance compared to healthy peers [80-82,85]. Some [70,71] but not all [73,80] have adjusted for increased absenteeism. It is important to understand why absenteeism plays such a major role in the attainment of asthmatic children. Absence levels have previously been associated with more severe asthma [70,73,82] and more requirements for healthcare [73].

Interestingly, unlike children with depression, the lower academic grades observed for asthmatic children did not seem to translate into increased unemployment. Whilst their school grades were adversely affected compared to peers, the association between asthma and attainment was weaker than the corresponding associations between attainment and ADHD, depression or epilepsy. Therefore it seems that, whilst asthmatic children's school grades were impacted, this was a significantly modest effect which might explain why it was not enough to alter their outcomes after leaving school. It is important to understand why absenteeism among asthmatic children has such a detrimental impact on their educational attainment in comparison to children with ADHD or epilepsy. Understanding patterns of absenteeism would be a starting point for further investigation. It might be the case that asthmatic pupils experience longer durations of absence which impact more significantly on their ability to catch up in class.

It is unclear why children with asthma observed in this thesis had fewer exclusions compared to healthy peers even after adjusting for potential sociodemographic and maternity confounders. This finding is counter-intuitive particularly in light of one previous study reporting the association to be the opposite way around [84]. However, on subgroup analyses, the findings were only significant within boys and those aged between 11 and 14 years and within the most deprived 40% of the population. Unaffected children within these groups already had the highest rates of exclusion; therefore it is possible that small differences, large sample size and residual confounding played a part in these findings. Further investigation is needed to confirm whether this is a real finding.

My thesis demonstrated that both diabetic and asthmatic children experienced a greater risk of all-cause hospital admission and all-cause mortality compared to peers but they did not experience more hospital admissions overall or an increased length of stay per admission. The associations between diabetes and all-cause admission [62,63] and all-cause mortality [28,56-60] are consistent with previous literature. The magnitude of the increased risk of all-cause and cause-specific admission associated with these conditions decreased with time from first receipt of relevant medication. This is not surprising since it may take several months following diagnosis for these conditions to be stabilised and for affected children and their parents to become knowledgeable and confident in managing them. These patterns in risk of hospital admission are consistent my other demonstrated findings previously discussed because they may contribute towards the increased absenteeism experienced by diabetic and asthmatic pupils; particularly younger asthmatic pupils.

My thesis reported diabetes and asthma prevalence of 0.43% and 6.00% respectively. The observed diabetes prevalence is consistent with a prevalence of 0.39% reported by a comparable population-wide study in Sweden [9]. Previous asthma studies have reported asthma prevalence between 4.9% and 18.6% however methodology used to diagnose cases ranges greatly across studies and most have based diagnosis on reported symptoms. My thesis only identified children actively receiving medication for asthma which explains why my observed prevalence was at the lower end of the aforementioned range.

6.1.3.1 Differences by age, gender and deprivation

The relative impact of diabetes and asthma on risk of SEN was stronger for girls and was explained because, among unaffected children, boys had higher absolute rates of SEN. The relative impact of diabetes and asthma on SEN was also stronger for younger and older pupils respectively. Previous studies have reported that the negative impact of diabetes on cognition and academic skills is most apparent when onset is before 7 years of age [9,10,49]. Within unaffected children, older and more deprived pupils had more absenteeism which explains why the association between asthma and absenteeism was strongest in these groups. It is also possible that parents may be more over-protective of their children when they are younger and suffering from asthma which may result in more absenteeism. Finally, it may take children and parents time to become accustomed to dealing with their condition resulting in greater absenteeism among younger pupils.

The relative impact of diabetes on risk of all-cause hospital admission was greater for girls than boys. Both had similar rates of baseline hospital admissions; therefore, it is important to understand why diabetes had a greater relative impact on girls with respect to risk of hospital admission. Furthermore, the relative impact of asthma on risk of respiratory admission was greater for boys than girls. The relative impact of asthma on risk of admission was greater for younger pupils, which is consistent with younger asthmatic children also experiencing a relatively greater risk of absenteeism. The relative impact of diabetes was greater for children aged between 9 and 14 years of age which is likely to reflect the fact that diabetic children in their early teens have poorer adherence to their insulin medication and are more difficult to manage around puberty [272].

6.1.4 Wider context and future policy

These findings clearly demonstrate that children with chronic conditions have significantly poorer educational and health outcomes compared to their peers. Although I have examined subsequent health outcomes in a paediatric cohort over a relatively short follow up period of five years the detrimental health effects of having a chronic condition are felt over the course of a lifetime. Similarly, the detrimental impact on educational outcomes observed over a relatively short period of follow-up are likely to have long-term sequelae for the child; including being an additional risk factor for adverse outcomes in adulthood.

Dahlgren and Whitehead's model (1991) [2] clearly illustrates the wider determinants of health (Figure 78). Permission to reproduce this figure has been granted by the Institute for Future Studies. The model recognises that individual factors are important determinants of health: both non-modifiable factors such as age, gender, ethnicity and genotype and modifiable lifestyle factors such as smoking, alcohol intake, physical activity and diet. However, it also highlights the wide range of social and environmental determinants that lie outside of the health sector and largely beyond the control of the individual, including education. Whilst better education can improve health directly in terms of health promotion and disease prevention, it also impacts health indirectly through higher income, better working conditions, better living conditions and greater access to healthcare and leisure activities [2,3]. Therefore, the relationship between health and education is bi-directional. Evans and Stoddart (1990) [3] more clearly demonstrated the complex inter-relationships between the factors which influence health and wellbeing (Figure 79). Permission to reproduce this figure has been granted by Elsevier. Their model highlighted that diseases can impair function and therefore prosperity but also that prosperity can itself

influence lifestyle behaviours that then impact on health. These models suggest that the long-term sequelae of adverse educational outcomes will include additional risks to health.

Figure 78 Wider determinants of health – Dahlgren and Whitehead model
figure reproduced from Dahlgren and Whitehead (1991) [2] with permission from the Institute for Future Studies [300]

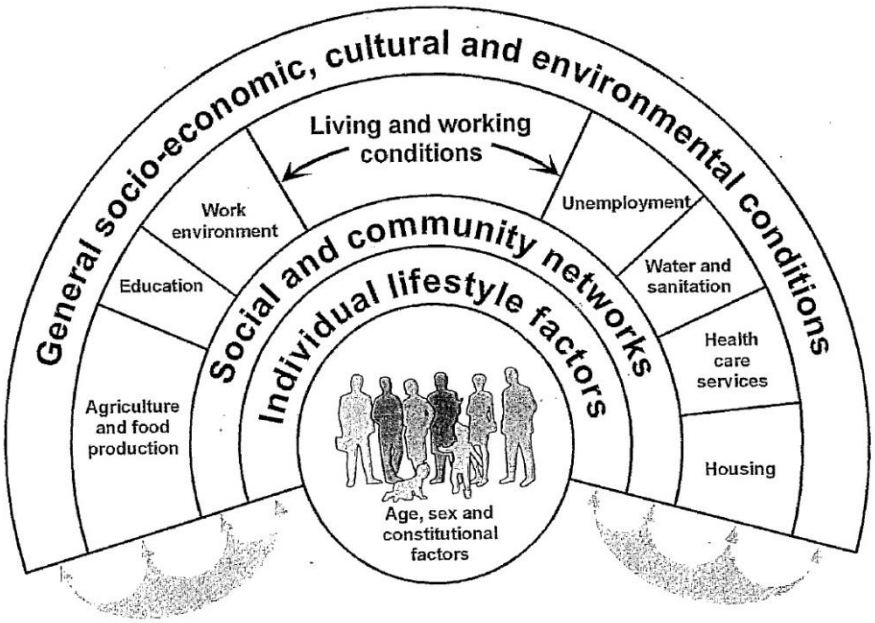
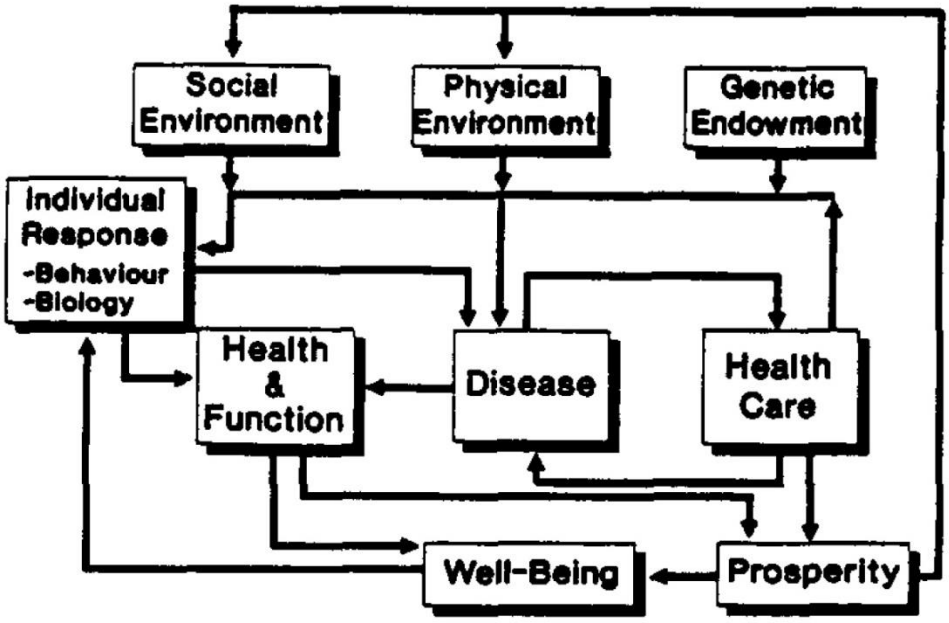


Figure 79 Wider determinants of health – Evans and Stoddart model
figure reproduced from Evans and Stoddart (1990) [3] with permission from Elsevier [301]



With this in mind, it is imperative that all children, regardless of their background or circumstances, receive the best standard of education possible and have equal opportunity to be able to perform well at school. All children are different; therefore, policies are required to ensure that school services and resources are tailored to individual pupil needs. This can help ensure that adequate additional support is given to disadvantaged pupils for example children from poorer socioeconomic backgrounds, children with troubled home lives, children with poor health and children with learning difficulties or behavioural problems.

Getting it right for every child (GIRFEC) is Scotland's national approach to improving outcomes and supporting wellbeing of children and young people and has been in place for a number of years [302]. GIRFEC aims to ensure that correct support is provided by the correct people when needed. It is focussed on tackling needs early and promoting pupil wellbeing and is child centred and committed to fostering joined up working between children, parents, teachers and wider services. Pupil wellbeing is outlined within the Children and Young People (Scotland) Act 2014 [303] and consists of eight indicators knows as SHANARRI: safe, healthy, achieving, nurtured, active, respected, responsible and included. The act itself makes it a requirement for local authorities to promote, support and safeguard the wellbeing of children via tailored services. From 2018, all children who need extra support which is not generally available will have that support given in the form of a tailored child's plan [304]. Following a review of guidance services in 2004, new policies are also in place to implement personal support for all pupils on the understanding that children are more likely to perform well when they feel happy, safe and supported. The Scottish Government is currently developing its Named Person Scheme [305] to ensure that someone is responsible as a point of contact for each child younger than 18 years of age. These contacts will aim to provide information, support and advice for every pupil when needed. A third part of GIRFEC is focussed on services working together to provide children with the best support possible. The Education (Additional Support for Learning) (Scotland) Act 2004 provides the legal framework for additional support for learning to be provided [306]. This act requires educational authorities and health and social work to work together to provide the necessary support for children with complex additional support needs.

Schools currently work in partnership with Allied Health Professionals (AHPs); most commonly speech and language therapists, physiotherapists and occupational therapists. Currently, all children receiving medication at school should be recognised on the school

system and be allocated a child healthcare plan. However; children who are only medicated at home do not have a healthcare plan. Depending on school resources, children may also have access to a school nurse, a local authority doctor, a school paediatrician, psychiatric services, or child and adolescent mental health services (CAHMS). Selected schools may have other services such as autistic units. However, these services vary greatly across schools as do recording and definitions of SEN.

Other recent policy changes include Curriculum for Excellence [256] which aims to transform education so that learning is relevant, inspiring and engaging. The Scottish Attainment Challenge was implemented in 2015 [307] and was committed to addressing and narrowing the attainment gap related to poverty. The Scottish Government has funded £750 million towards this venture and this is linked to the Pupil Equity Fund [308] which allocates funding to schools most affected by the poverty related attainment gap. The Scottish Government has also prioritised improving mental health services across Scotland via the mental health strategy 2012-15 [309]. This aims to promote mental wellbeing and prevent mental illness and one of the key areas concerns child and adolescent mental health. The National Improvement Framework for Scottish Education, launched in 2016, aims to develop and improve data collection and intelligence across primary and secondary schools in order to improve support of pupils and inform policymakers where future improvement may be needed.

Whilst some of these policies have yet to be implemented, others have been in place for several years. However, in spite of these policies, this thesis demonstrated that children suffering from chronic illness continue to experience educational and health disadvantage compared to peers. The introduction of many of these policies pre-dates the study period; therefore, it appears that they have not fully helped bridge the deficit experienced by chronically ill pupils. This may not be surprising because, regardless of additional input, some chronically ill children may simply never be able to perform as well as their healthy peers. However, in general, many of the educational policies focus on health and wellbeing and the impact of social disadvantage; particularly poverty. Whilst poverty accounts for a large part of the disparity in educational opportunities, chronic ill health should not be overlooked. New policies such as the mental health strategy may have an impact on chronic conditions such as ADHD and depression and may enable better communication between schools and health care providers. However, given the continuing performance gap between chronically ill pupils and their peers, additional and more specific policies

may need to be introduced to help identify, treat and care for children with chronic conditions in school.

6.2 Strengths and limitations

6.2.1 Strengths

I have analysed data from a large, non-selective study population of schoolchildren across the whole of Scotland. In total, 2,793,185 records pertaining to 766,244 pupils were analysed across five school years. This exceeds the sample sizes used in previous studies investigating chronic conditions and educational outcomes. Indeed only Dalquist et al [9], who investigated associations between diabetes and school outcomes in more than 1.3 million schoolchildren, included a comparably large number of participants. The majority of diabetes studies analysing school outcomes have used small population samples [29,30,32-37,52] ranging between 60 [30,32] and 500 [35] subjects in total. Persson et al [10] matched 9,940 controls to 2,845 diabetic children whilst other studies have analysed in excess of 20,000 [31] and 100,000 [38] children in total but have only included in the region of 100 diabetic children within these samples. The largest study to assess the effects of asthma during school years analysed 300,000 school records but only investigated grade point average [79]. Whilst smaller studies used samples in the region of 1,500 [73,88], 4,000 [82] or 10,000 [70,83,84] subjects, others used samples of between fifty and a few hundred [71,78,80,81,85,89] and compared matched comparison subjects to asthmatic children [80,81,85,89]. Studies investigating educational outcomes for children with epilepsy compared to peers have used small samples, with the largest study matching 131 siblings to 219 children with epilepsy [120] and it has not been uncommon for investigators to analyse matched cohorts totalling between 40 and 100 participants [114,117-119,122]. The largest ADHD study included almost 9,500 children with and without ADHD[178] whilst sample size in other studies has ranged between 60 [168] and 1000 [172] children in total. Previous studies assessing the impact of depression on educational outcomes have generally used larger cohorts with the largest sampling in the region of 14,000 pupils [14,15,214]. However, earlier studies have used small population samples [219-221].

This thesis only included local authority maintained schools; however, less than 5% of Scottish children attend private schools [245]. Therefore, data from the pupil census covered more than 95% of schoolchildren in Scotland and comprised all children from primary 1 (P1) through to final year of secondary school (S6) as well as pupils attending special

schools. The full school-age population was included; therefore, the cohort comprised schoolchildren of both genders, between 4 and 19 years of age, from all socioeconomic and ethnic backgrounds attending urban, suburban and rural schools. Previous studies have been limited by including only boys [176,182], predominately children from particular ethnic or socioeconomic backgrounds [82,83,166,168,172], or pupils of particular ages [99]. Others have not included children attending special schools [9]. The large size of my thesis population meant that I had sufficient power to test for statistical interactions and undertake sub-group analyses where appropriate. I have been able to investigate differences by gender, socioeconomic status and age for the majority of outcomes whereas previous studies with small sample size have not been able to explore these potential effect modifiers.

This thesis analysed a wide range of outcomes covering both the education and health sectors. No other studies in the UK have investigated educational outcomes of schoolchildren with chronic conditions on a population-wide scale. Furthermore, to the best of my knowledge, no previous studies worldwide have investigated as wide a range of educational outcomes for children with chronic conditions using population-wide data, and none have dually investigated educational outcomes, hospitalisations and mortality. Previous studies have linked Scotland-wide education data to national maternity data to investigate associations between obstetric factors and SEN [16,17]. This research extended the approach by additionally linking these datasets to prescribing data, hospital admissions and deaths to investigate a wider array of predictive factors and outcomes whilst adjusting for a range of important confounders relating to sociodemographic status, pregnancy and delivery. The large population size and wide array of available outcomes allowed investigation of academic attainment and unemployment after adjusting for absenteeism and excluding children with SEN. Some [70,71] but not all [73,80] of the previous asthma studies investigating attainment have explored absenteeism as a possible mediator. Previous studies investigating the attainment of children with diabetes, ADHD, depression and epilepsy have not adjusted for absenteeism. Furthermore, previous studies have not investigated school absenteeism as a possible mediator for future unemployment. Whilst some previous studies have excluded children with learning difficulties from their analyses [173,225], few have excluded children with any type of SEN. I was also able to adjust for children experiencing more than one chronic condition by including all conditions together in final statistical models. Therefore, each chronic condition specific analysis was adjusted for presence of any of the other identified chronic conditions. By comparison, few previous studies have adjusted for such a range of confounders and even fewer have included obstetric and maternity confounders [9,73].

My population cohort included repeated measures for children attending school across multiple years during the study period. Therefore I used generalised estimating equations for longitudinal data to account for correlations between outcomes recorded for the same pupils across different years. Few other studies have analysed longitudinal data [14,15,213-215,218,223,224] or taken account of such correlations. Indeed several previous studies have analysed outcomes using cross-sectional study designs [219-222] limited by potential reverse causation. Robust analyses were performed such as Poisson piecewise regression, to account for non-proportional hazards when investigating health outcomes, and generalised ordinal logistic regression, to account for non-parallel odds when evaluating academic attainment. In addition to identifying poorer outcomes for children with chronic conditions, these techniques enabled more intricate relationships between each condition and the respective outcomes to be revealed. For example, children with chronic conditions experienced increased hospitalisation; however the excess risk decreased in magnitude over the follow up period and varied by age at admission. Similarly, whilst children with asthma, epilepsy, ADHD or depression were observed to have poorer academic attainment compared to peers, I demonstrated that children with epilepsy or depression were particularly adversely affected and had an increased risk of having attainment specifically classed as poor.

All of the education and health outcomes evaluated in this thesis were obtained directly from routine administrative databases and these databases undergo regular quality assurance checks [310]. The linkage of education and health records relied on probabilistic matching; however a previous validation of this methodology demonstrated that more than 99% of the linkages that were made attached the pupil to his/her correct health records for singleton births [19]. Attainment was evaluated using grades achieved in national examinations recorded across the full syllabus of school subjects and across multiple years of schooling from S4 through to S6. By comparison, some previous studies have used standardised test scores [33,35,71,80,117,120,166,168,219], parental, teacher or pupil reports [29,31,36-38], or have only based school attainment on particular subjects [32,33,85,173,221,223,226]. Other educational outcomes assessed in this research were equally robust. Absenteeism and exclusions were recorded annually by schools and collated at the end of each year. Leaver destination is recorded 6-months after a child leaves school by Skills Development Scotland [311] and is submitted to the Scottish Government. The NEET definition used for unemployment in this study was based on that used by the Scottish Government and the Office for National Statistics [255]. Records of SEN were recorded by the schools and included individual types of SEN enabling analyses of more specific outcomes such as

learning disability, autism, and mental and physical disability. Few other studies have investigated SEN, and these have relied on parental report or undefined methods of ascertainment. Furthermore, no other studies have reported on the full range of different types of SEN observed in my study. Previous studies have assessed subsequent outcomes using questionnaire responses and some have lost subjects in the follow up period [57,172,181]. However, my study obtained follow up for children's health outcomes using record linkage to administrative health records. Therefore, pupils should not have been lost to follow up unless they emigrated.

Many previous studies investigating educational outcomes of children with chronic conditions only identified chronically ill subjects attending medical clinics [29,33,36,37,112,116,117,122,166,168,171,181]. However, this research analysed a population-wide cohort of schoolchildren, rather than a subgroup identified using health records. Therefore, in this respect, inclusion was not restricted to the most severely affected cases. In addition, previous studies have relied on parental, patient or teacher questionnaire responses with varying clinical input and few have used prescribing data to identify children with chronic conditions [140]. I identified pupils with chronic conditions based on receipt of medication as prescribed by a clinician. Therefore the method of case ascertainment used in this research, whilst experiencing limitations outlined in the next section, was based on clinical prescribing data and may be more accurate than previous studies relying on questionnaire responses.

6.2.2 Limitations

My thesis also had several limitations. I identified children with chronic medical conditions from their prescribed medication. The first main assumption was that some drugs are used to treat only one disease: anti-epileptics for epilepsy; insulin for type I diabetes; bronchodilator and steroid inhalers for asthma; CNS stimulants for ADHD; anti-depressants for depression. Therefore, children in receipt of these drugs were assumed to have these conditions. However these medications may also be prescribed to treat other conditions not related to their primary use. For example, stage 2 asthma treatment may be issued to treat chronic obstructive pulmonary disease (COPD). However, this disease only occurs in adults [312] and therefore no children in my population would be on medication for this reason. Insulin may also be used to treat insulin insufficiency associated with cystic fibrosis. However, prevalence of cystic fibrosis is only 5 per 10,000 people and therefore the number of children with cystic fibrosis in Scotland during the study period would have

been very small (approximately 400) compared to the number of children with type 1 diabetes [313]. My thesis did not identify specific type of diabetes. However, type 2 diabetes most commonly presents in adulthood and is very rare in children [23]. Therefore the vast majority of children in my cohort would have had type 1 diabetes. Some anti-depressant medications may be used to treat anxiety or other psychiatric disorders such as obsessive compulsive disorder; however these commonly occur alongside depression. Although not first line, specific tricyclic antidepressants such as imipramine may also be prescribed for bed wetting. However the number of prescriptions for imipramine in my cohort was small in absolute terms and the majority were for children over 12 years of age making bed wetting a less likely reason. Antidepressant and anti-epileptic drugs may also be used to treat chronic pain syndromes; however, the number of children using these agents for pain syndromes is again small. The extended use of these medications is significantly more common in the adult population.

The second main assumption was that if a child suffered from one of these chronic conditions then they were on medication. However, in reality, only people with type 1 diabetes definitely require medication to survive. It is possible for children with less severe forms of the other conditions to be un-medicated until a diagnosis is made or until symptoms deteriorate. These issues provided some challenges. Firstly, observed prevalence may have been overestimated resulting in type 1 errors (false positive) whereby children identified as having a particular chronic condition actually received their medication for another reason. Conversely the observed prevalence may have been underestimated resulting in type 2 errors (false negative) whereby children with milder forms of a condition did not require medication and were not identified. Secondly, I was only able to study the disease specific outcomes of children on medication, who were likely to have more severe forms of each chronic condition and poorer outcomes. This may have overestimated the association of each condition with the outcomes of interest. The comparison group, however, will have included children with milder forms of each condition who were undiagnosed or not on medication and this may have underestimated the effect of each condition.

Medications can have adverse side effects; therefore children receiving prescription medication may experience poorer school outcomes due to the side effects of their treatment in addition to the primary effects of their condition. This may have biased the results by inflating the effect size of associations between chronic conditions and outcomes. Previous studies have assessed the secondary effects of prescribed medication

by comparing outcomes for medicated and non-medicated chronically ill children. I was unable to identify untreated children with these conditions; therefore I could not investigate whether adverse outcomes were due to the medication or the underlying condition. Previous studies have also investigated the effects of different medications and doses. This was not the aim of my thesis; however, the available data will be used to explore this area in future work.

Whilst education data were available from 2006 onwards, prescribing data were only available for linkage from 2009 onwards. Furthermore, at the time the data were extracted, school records were only available until 2013. Therefore the cohort was restricted to school records between 2009 and 2013 and comprised records pertaining to all schoolchildren attending school during this period. Children were diagnosed with a chronic condition based on receipt of medication between 2009 and 2013; however, I did not have information on formal clinical diagnoses or dates of diagnoses for each condition. For children who started medication subsequent to their first available pupil census date post 2009, it was reasonable to assume that their first prescription date roughly equated to their date of diagnosis. However, I was unable to estimate dates of diagnoses, based on commencement of treatment, for children who were already on medication by 2009 or on starting school. I was therefore unable to investigate effects of age at diagnosis or duration of illness on any of the outcomes.

Limitations around date of diagnoses also impacted analyses investigating all-cause and cause-specific hospital admissions and all-cause mortality. These outcomes were analysed by following up children for five years using Cox proportional hazards and Poisson piecewise regression models. However, I only had school census data, prescription information, admissions and deaths from 2009 onwards regardless of whether children started school, had been prescribed medication or had been admitted to hospital before this period. Therefore the survival analyses only modelled time until first hospital admission and death following first school census date or first relevant prescription date from 2009 onwards. The models quantified risk of hospital admission and death at different ages. Whilst this was informative, the model did not provide any information on age at actual clinical diagnosis because this was not known for any of the children and all of the pupils had different lengths of follow up.

Despite the large population size, I was unable to investigate sub-group analyses by ethnic group for any of the educational or health outcomes due to small numbers of non-white

children. Additionally, deprivation sub-group analyses were more difficult to interpret for the outcomes of academic attainment and risk of hospital admission because I performed more complex non-parallel odds and non-proportional hazards models respectively. Performing less complicated analyses would have enabled easier interpretation between deprivation subgroups but these inferior fitting models would have resulted in less robust conclusions and more limited understanding within the respective main analyses. Finally, the numbers of deaths observed across the study period were small which meant that there was insufficient statistical power to study specific causes of death or investigate differences in all-cause mortality by gender, socioeconomic status or age.

This study compared children with each of the chronic conditions of interest in turn against unaffected peers. However, for each condition investigated, children in either the chronically ill or healthy peer group may have additionally experienced any of the other conditions investigated. Therefore, after studying each condition individually, the main analyses were reran including all of the chronic conditions in the same models (diabetes, asthma, epilepsy, ADHD and depression) to see if the effects of each were independent of the others. However, it was not possible to identify children with additional comorbid conditions and adjust accordingly or exclude these children from the analyses. My analyses adjusted for a wealth of sociodemographic and maternity confounders; however, as with any observational study there is always potential for residual confounding. I did not adjust for other potentially important confounding factors such as parental influence, childhood morbidity or prior hospitalisation. These factors will be investigated in future work.

Absence and exclusion data from ScotXed have historically been recorded annually; however, from 2010 onwards the data were only submitted every two years with outcomes recorded for the second year. These data gaps meant that absences and exclusions could only be analysed for census years 2009, 2010 and 2012; however, there are unlikely to have been systematic differences in the two missing years. A further limitation of the absence data concerned an inability to investigate durations of individual absence episodes. The data included counts of total numbers of days absent per year but did not include numbers of individually occurring episodes of absenteeism per year. This meant that comparisons could not be made between children with multiple short duration absences and those with less frequent but longer duration absences.

Improvements in recording of SEN over the study period may have biased results. 6.4% of pupils had a recorded SEN in 2009; however, this percentage increased every year and, by

2013, 14.8% had a recorded SEN. This meant that some children who had school difficulties in earlier years may not have been as likely to have been identified as having a SEN compared to children in later years. Therefore, the associations between chronic conditions and SEN in earlier years may have been underestimated. Additionally, some new SEN codes were recorded from 2010 onwards. Many of these were not relevant or used in my analyses; however, SEN attributed to communication problems was analysed despite not being recorded before 2010.

All of the school measures were recorded by individual schools; therefore differences in recording may have occurred across different institutions resulting in recording bias. For example, some schools may have had a more rigorous approach to exclusions or better ability to record SEN. Additionally schools in less deprived areas may have had higher average grades or better outcomes than schools in more deprived areas. For this reason it may be worthwhile to investigate multilevel modelling in future to take into account differences at school level. In order to perform meaningful multilevel analyses, one would require information at pupil level and also at school level. I did not perform multilevel modelling because I did not have any information at the school level and at present this information is still not accessible for research purposes.

1.6% of children in the pupil census could not be linked to the CHI register. ScotXed held full pupil names on the pupil census but were not permitted to share these for linkage purposes. Therefore records were probability matched using pupil sex, date of birth and postcode as linkage identifiers. Absence of full names in the linkage resulted in twins being omitted from the final cohort. For twins with the same postcode, sex and date of birth, it was impossible to decipher which link was the correct one; therefore the cohort was restricted to singleton births. Whilst probability matching is a widely used and accepted method, there is always potential for some residual bad links to occur. However, a previous validation of this methodology demonstrated that more than 99% of the linkages that were made attached the pupil to his/her correct health records for singleton births [237,314]. Furthermore, additional bad links were identified and removed wherever possible.

A further 12.3% of children could not be linked to their maternity records and were therefore omitted from the cohort. SMR02 covers 98% of births in Scotland excluding home births; therefore the most likely explanation is that children who could be linked to CHI but did not have an SMR02 record were not born in Scotland. According to the 2011 Scottish Census, 11% of Scottish residents aged 5-19 years were born outside of Scotland

and therefore should not be linkable to Scottish maternity records [315]. The cohort therefore was restricted to Scottish births; however it is unlikely that exclusion of children born in other countries introduced systematic bias. The prevalence of each condition within the 12.3% of pupils who could not be linked to maternity data (diabetes 0.32%; asthma 4.24%; epilepsy 0.60%; ADHD 0.86%; depression 0.66%) was similar to the prevalence within those who could (diabetes 0.43%; asthma 6.00%; epilepsy 0.69%; ADHD 0.97%; depression 0.70%).

The introduction of Curriculum for Excellence (CfE) [256] ensured that S4 and S5 pupils sat a new set of national 4 and 5 examinations in 2013 which replaced standard grade and higher awards. This meant that S4 and S5 attainment data in 2013 was not comparable to previous years. This was not an issue for this study because my derivation of attainment only included 2013 exam results for S6 pupils who had not moved on to CfE by that point (section 4.3.1.2.5). However, Curriculum for Excellence will provide future challenges when analysing attainment data across multiple years due to lack of consistency in recording either side of 2013.

6.3 Further work

My thesis uncovered some very interesting associations between the respective chronic conditions and educational and health outcomes. Whilst some associations, for example increased absenteeism and increased SEN in chronically ill children, may not be surprising, the risks have not previously been quantified using population-wide data either in Scotland or the UK. However, several more intricate aspects of the uncovered associations were less intuitive and merit further investigation. For example, it is not immediately clear why the academic performances of depressed and asthmatic children seem to be particularly adversely affected by their increased absenteeism. Whilst information on length of absenteeism is not available from ScotXed at this point, investigating specific reasons for absence among the different chronically ill children, particularly those with depression and asthma would be informative.

Whilst most of the associations between chronic conditions and poor outcomes were stronger for girls than boys, the opposite was true for depression. Associations between depression and both SEN and unemployment were stronger for boys which perhaps reflects different gender specific comorbidities and patterns of behaviour in depressed children. Investigating the types of SEN experienced by depressed boys and depressed girls may

help uncover some of the differences. Younger children with depression and ADHD also experienced greater risk of exclusion suggesting that perhaps teachers are less aware of their condition in early years and are more likely to exclude them. The weaker association between ADHD and SEN in early years confirms that there may indeed be issues with early diagnoses of ADHD within the school setting. Future dialogue with schools would be worthwhile in order to understand what constitutes the different types of SEN and whether there are indeed issues with early diagnoses.

Further work is also required to understand why children with diabetes and asthma experience less long term adversity compared to children with the other chronic conditions. In particular, children with asthma suffer poorer grades but this doesn't translate into future risk of unemployment whereas children with depression experience adversity on both counts. The association between asthma and exclusion was counterintuitive and should also be revisited.

Risk of hospital admission became weaker over follow up time for all of the conditions with the exception of ADHD. Therefore, it seems that chronically ill children are particularly at risk close to their time of diagnosis. Changes in practice may be needed to provide support over this difficult time. Gender differences were also apparent and my thesis uncovered increased risk of injury admission in children with depression or ADHD; particularly adolescent girls. Diabetic teenagers also experienced increased risk of hospital admission probably due to decreased levels of medication adherence. Future interventions should therefore focus on these subgroups of children.

Various avenues are open to investigation within wider future work. This thesis identified children receiving medications to treat specific chronic conditions and assessed the main effect of those conditions on a range of educational and health outcomes. It was not within the scope of this thesis to investigate medication specific effects such as type of medication or dose. However this information is readily available within PIS data. Exploring type or dose of medication may provide information on disease severity which may contribute towards poor outcomes. Specific medications may affect children in different ways, both in terms of treating their condition and producing adverse side effects.

I adjusted condition-specific analyses for the presence of other chronic conditions investigated in this thesis (diabetes, asthma, epilepsy, ADHD or depression). It would be insightful to specifically model the increased risk of poor outcomes for children suffering

from specific combinations of these conditions compared to unaffected peers. Future work should also aim to identify additional comorbid conditions which may influence poor outcomes. These comorbidities may be identifiable from retrospective linkage to previous acute or psychiatric hospital admissions or may become available in the future from linkage to Scotland-wide primary care data. Several studies have matched children with chronic conditions to their unaffected siblings in order to assess associations between chronic ill health and educational outcomes whilst adjusting for family and neighbourhood confounders which are common to siblings [29,35,38,39,80,114,120,121,137,176,214]. My cohort contained non-twin siblings; therefore similar investigation is feasible and worthwhile.

I analysed attainment based on all school subjects; however, the dataset contained attainments and grades from specific individual subjects and so further analyses could explore potential differences in, for example, language based subjects versus mathematical or science based subjects. Multilevel analyses could be incorporated in future to adjust for potential school level differences if school level data becomes available. Finally, I defined presence of each chronic condition based on receipt of relevant medication; however, I had no information on dates of diagnoses. Future linkage to other sources such as primary care data or disease registers may provide this information and allow analyses of age at diagnosis and duration of diagnosis.

6.4 Challenges, lessons learned and final thoughts

I have really enjoyed undertaking this PhD and have found it a very rewarding experience. Naturally, there have been challenges along the way. An early challenge involved obtaining the various governance approvals to link education data to health data. This project was the first to link education data to prescribing data and therefore the governance process took a lot longer than expected. The first PAC application was rejected but was approved on re-submission after further discussion on the proviso that subsequent assurances would be given around patient consent for the study. Dealing with large amounts of complex linked data provided subsequent challenges at the analysis stage. I had not used prescribing data or education data before and the education data in particular required a lot of manipulation and data cleaning to enable me to analyse the data the way I wanted. I also encountered several issues with the safe haven in the early stages of the project which delayed my analyses. The safe haven infrastructure was fairly new which meant that there were inevitably some teething problems. Most of the problems centred on waiting for data to be signed out of the safe haven and software issues whereby, for

example, Stata ado files needed for the various statistical analyses had to be manually uploaded by staff at ATOS origin. The latter issue was resolved after the safe haven was transferred to EPCC. From a more personal point of view, my mum passed away nine months into my PhD and this was a huge setback which I did not envisage dealing with. More general challenges included trying not to procrastinate too much, staying organised and focussed and keeping the main goal in mind. I've learned that doing a PhD is a marathon and not a sprint!

I have learned lots of new skills throughout the PhD. As a statistician within ISD Scotland, I had previously worked in the area of record linkage where I analysed a wide variety of health data including SMR data. This PhD allowed me to broaden my horizons by working with educational data which I found very interesting. Having previously used SPSS and SAS, I have now become a very competent user of Stata and I have learned lots of new and interesting analytical and statistical techniques. I undertook a sizeable literature review which has provided valuable experience of using Endnote and the various search tools. My organisational, time management and project management skills have improved amongst others and I have had the opportunity to attend various courses and conferences, present my work, take part in tutoring and even deliver a lecture. The analytical aspects of this PhD came more naturally to me as a statistician whilst some more challenging aspects included the literature review and putting my findings into the wider context. I feel that the challenges I have encountered and the experiences I have gained throughout the PhD will be invaluable for my career going forward.

List of References

1. World Health Organisation. Global Health Risks. Mortality and burden of disease attributable to selected major risks. 2004; http://www.who.int/healthinfo/global_burden_disease/GlobalHealthRisks_report_full.pdf?ua=1&ua=1. Accessed 17/06/2017.
2. Dahlgren G, Whitehead M. Policies and strategies to promote social equity in health. Stockholm: Institute for future studies. 1991.
3. Evans RG, Stoddart GL. Producing health, consuming health care. Soc Sci Med. 1990;31(12):1347-1363.
4. Institute of Public Health in Ireland. Health Impacts of Education. A review. 2008; https://www.publichealth.ie/files/file/Health%20Impacts%20of%20Education.pdf. Accessed 17/06/2017.
5. Connelly R, Sullivan A, Jerrim J. Primary and secondary education and poverty review. York, Joseph Rowntree Foundation 2014; http://www.cls.ioe.ac.uk/library-media/documents/Primary%20and%20secondary%20education%20and%20poverty%20review%20August%202014.pdf. Accessed 21/06/2017.
6. Mokkink LB, van der Lee JH, Grootenhuis MA, Offringa M, Heymans HS. Defining chronic diseases and health conditions in childhood (0-18 years of age): national consensus in the Netherlands. Eur J Pediatr. 2008;167(12):1441-1447.
7. Crump C, Rivera D, London R, Landau M, Erlendson B, Rodriguez E. Chronic health conditions and school performance among children and youth. Annals of epidemiology. 2013;23(4):179-184.
8. Taras H, Potts-Datema W. Chronic Health Conditions and Student Performance at School. Journal of School Health. 2005;75(7):255-266.
9. Dahlquist G, Kallen B, Swedish Childhood Diabetes Study G. School performance in children with type 1 diabetes--a population-based register study. Diabetologia. 2007;50(5):957-964.
10. Persson S, Dahlquist G, Gerdtham UG, Steen Carlsson K. Impact of childhood-onset type 1 diabetes on schooling: a population-based register study. Diabetologia. 2013;56(6):1254-1262.
11. Reilly C, Neville BG. Academic achievement in children with epilepsy: a revie w. Epilepsy Research. 2011;97(1-2):112-123.
12. Taras H, Potts-Datema W. Childhood asthma and student performance at school. Journal of School Health. 2005;75(8):296-312.
13. Frazier TW, Youngstrom EA, Glutting JJ, Watkins MW. ADHD and achievement: meta-analysis of the child, adolescent, and adult literatures and a concomitant study with college students. Journal of Learning Disabilities. 2007;40(1):49-65.
14. Fletcher JM. Adolescent depression: diagnosis, treatment, and educational attainment. Health economics. 2008;17(11):1215-1235.
15. Needham BL. Adolescent depressive symptomatology and young adult educational attainment: an examination of gender differences. Journal of Adolescent Health. 2009;45(2):179-186.
16. MacKay DF, Smith GC, Dobbie R, Pell JP. Gestational age at delivery and special educational need: retrospective cohort study of 407,503 schoolchildren. PLoS medicine. 2010;7(6):e1000289.
17. Mackay DF, Smith GCS, Dobbie R, Cooper SA, Pell JP. Obstetric factors and different causes of special educational need: retrospective cohort study of 407 503 schoolchildren. BJOG: An International Journal of Obstetrics & Gynaecology. 2013;120(3):297-308.
18. Diabetes UK. Diabetes in the UK 2012. 2012; https://www.diabetes.org.uk/Documents/Reports/Diabetes-in-the-UK-2012.pdf. Accessed 15/06/2017.

19. International Diabetes Federation 2013. IDF Diabetes Atlas Sixth edition. 2013; http://www.idf.org/sites/default/files/EN_6E_Atlas_Full_0.pdf. Accessed 15/06/2017.

20. Diabetes UK. Diabetes Facts and Stats. 2014; https://www.diabetes.org.uk/Documents/About%20Us/Statistics/Diabetes-key-stats-guidelines-April2014.pdf. Accessed 15/06/2017.

21. Maahs DM, West NA, Lawrence JM, Mayer-Davis EJ. Epidemiology of type 1 diabetes. Endocrinol Metab Clin North Am. 2010;39(3):481-497.

22. Kucera M, Sullivan AL. The educational implications of type I diabetes mellitus: A review of research and recommendations for school psychological practice. Psychology in the Schools. 2011;48(6):587-603.

23. Scottish Diabetes Survey Monitoring Group. Scottish Diabetes Survey 2013. 2013; http://www.diabetesinscotland.org.uk/Publications/SDS2013.pdf. Accessed 15/06/2017.

24. Tuomilehto J. The emerging global epidemic of type 1 diabetes. Curr Diab Rep. 2013;13(6):795-804.

25. Royal College of Paediatrics and Child Health. National Paediatric Diabetes Audit 2015-16 Report 1: Care Processes and Outcomes. 2017; http://www.rcpch.ac.uk/system/files/protected/page/Complete%20NPDA%202015-16%20report%20FINAL_0.pdf. Accessed 15/06/2017.

26. American Diabetes Association. Diabetes care in the school and day care setting. Diabetes Care. 2014;37 Suppl 1:S91-96.

27. Holmes CS, Cant MC, Fox MA, Lampert NL, Greer T. Disease and demographic risk factors for disrupted cognitive functioning in children with insulin-dependent diabetes mellitus (IDDM). School Psychology Review. 1999;28(2):215-227.

28. Dahlquist G, Kallen B. Mortality in childhood-onset type 1 diabetes: a population-based study. Diabetes Care. 2005;28(10):2384-2387.

29. Parent KB, Wodrich DL, Hasan KS. Type 1 diabetes mellitus and school: a comparison of patients and healthy siblings. Pediatr Diabetes. 2009;10(8):554-562.

30. Ryan C, Longstreet C, Morrow L. The effects of diabetes mellitus on the school attendance and school achievement of adolescents. Child: Care, Health & Development. 1985;11(4):229-240.

31. Wennick A, Hallstrom I, Lindgren B, Bolin K. Attained education and self-assessed health later in life when diagnosed with diabetes in childhood: a population-based study. Pediatric Diabetes. 2011;12(7):619-626.

32. Meo SA, Alkahlan MA, Al-Mubarak MA, et al. Impact of type 1 diabetes mellitus on academic performance. The Journal of international medical research. 2013;41(3):855-858.

33. Yu SL, Kail R, Hagen JW, Wolters CA. Academic and social experiences of children with insulin-dependent diabetes mellitus. Children's Health Care. 2000;29(3):189-208.

34. Erkolahti R, Ilonen T. Academic Achievement and the Self-image of Adolescents with Diabetes Mellitus Type-1 and Rheumatoid Arthritis. Journal of Youth and Adolescence. 2005;34(3):199-205.

35. McCarthy AM, Lindgren S, Mengeling MA, Tsalikian E, Engvall JC. Effects of diabetes on learning in children. Pediatrics. 2002;109(1):E9.

36. Lloyd CE, Robinson N, Fuller JH. Education and employment experiences in young adults with type 1 diabetes mellitus. Diabetic Medicine. 1992;9(7):661-666.

37. Robinson N, Stevens LK, Protopapa LE. Education and employment for young people with diabetes. Diabet Med. 1993;10(10):983-989.

38. Glaab LA, Brown R, Daneman D. School attendance in children with Type 1 diabetes. Diabet Med. 2005;22(4):421-426.

39. Vetiska J, Glaab L, Perlman K, Daneman D. School attendance of children with type 1 diabetes. Diabetes Care. 2000;23(11):1706-1707.

40. Shehata G, Eltayeb A. Cognitive function and event-related potentials in children with type 1 diabetes mellitus. Journal of Child Neurology. 2010;25(4):469-474.

41. Naguib JM, Kulinskaya E, Lomax CL, Garralda ME. Neuro-cognitive performance in children with type 1 diabetes—A meta-analysis. Journal of Pediatric Psychology. 2009;34(3):271-282.

42. McCarthy AM, Lindgren S, Mengeling MA, Tsalikian E, Engvall J. Factors associated with academic achievement in children with type 1 diabetes. Diabetes Care. 2003;26(1):112-117.

43. Hannonen R, Komulainen J, Eklund K, Tolvanen A, Riikonen R, Ahonen T. Verbal and academic skills in children with early-onset type 1 diabetes. Developmental Medicine and Child Neurology. 2010;52(7):e143-e147.

44. Hannonen R, Komulainen J, Riikonen R, et al. Academic skills in children with early-onset type 1 diabetes: The effects of diabetes-related risk factors. Developmental Medicine and Child Neurology. 2012;54(5):457-463.

45. Perantie DC, Lim A, Wu J, et al. Effects of prior hypoglycemia and hyperglycemia on cognition in children with type 1 diabetes mellitus. Pediatric Diabetes. 2008;9(2):87-95.

46. Desrocher M, Rovet J. Neurocognitive Correlates of Type 1 Diabetes Mellitus in Childhood. Child Neuropsychology. 2004;10(1):36-52.

47. Lin A, Northam EA, Rankins D, Werther GA, Cameron FJ. Neuropsychological profiles of young people with type 1 diabetes 12 yr after disease onset. Pediatric Diabetes. 2010;11(4):235-243.

48. Schoenle EJ, Schoenle D, Molinari L, Largo RH. Impaired intellectual development in children with type I diabetes: Association with hba1c, age at diagnosis and sex. Diabetologia. 2002;45(1):108-114.

49. Gaudieri PA, Chen R, Greer TF, Holmes CS. Cognitive function in children with type 1 diabetes: a meta-analysis. Diabetes Care. 2008;31(9):1892-1897.

50. Blasetti A, Chiuri RM, Tocco AM, et al. The effect of recurrent severe hypoglycemia on cognitive performance in children with type 1 diabetes: a meta-analysis. Journal of Child Neurology. 2011;26(11):1383-1391.

51. Rovet JF, Ehrlich RM, Hoppe M. Specific intellectual deficits in children with early onset diabetes mellitus. Child development. 1988;59(1):226-234.

52. Overstreet S, Holmes CS, Dunlap WP, Frentz J. Sociodemographic risk factors to intellectual and academic functioning in children with diabetes. Intelligence. 1997;24(2):367-380.

53. Fletcher JM, Richards MR. Diabetes's 'health shock' to schooling and earnings: increased dropout rates and lower wages and employment in young adults. Health Aff (Millwood). 2012;31(1):27-34.

54. Morgan E, Cardwell CR, Black CJ, McCance DR, Patterson CC. Excess mortality in Type 1 diabetes diagnosed in childhood and adolescence: a systematic review of population-based cohorts. Acta Diabetologica. 2015;52(4):801-807.

55. Angus VC, Waugh N. Hospital admission patterns subsequent to diagnosis of type 1 diabetes in children : a systematic review. BMC Health Services Research. 2007;7:199.

56. Patterson CC, Dahlquist G, Harjutsalo V, et al. Early mortality in EURODIAB population-based cohorts of type 1 diabetes diagnosed in childhood since 1989. Diabetologia. 2007;50(12):2439-2442.

57. Secrest AM, Becker DJ, Kelsey SF, LaPorte RE, Orchard TJ. All-cause mortality trends in a large population-based cohort with long-standing childhood-onset type 1 diabetes: the Allegheny County type 1 diabetes registry. Diabetes Care. 2010;33(12):2573-2579.

58. O'Grady MJ, Delaney J, Jones TW, Davis EA. Standardised mortality is increased three-fold in a population-based sample of children and adolescents with type 1 diabetes. Pediatric Diabetes. 2013;14(1):13-17.

59. Gagnum V, Stene LC, Sandvik L, et al. All-cause mortality in a nationwide cohort of childhood-onset diabetes in Norway 1973-2013. Diabetologia. 2015;58(8):1779-1786.

60. Skrivarhaug T, Bangstad HJ, Stene LC, Sandvik L, Hanssen KF, Joner G. Long-term mortality in a nationwide cohort of childhood-onset type 1 diabetic patients in Norway. Diabetologia. 2006;49(2):298-305.

61. McNally PG, Raymond NT, Burden ML, et al. Trends in mortality of childhood-onset insulin-dependent diabetes mellitus in Leicestershire: 1940-1991. Diabetic Medicine. 1995;12(11):961-966.

62. Icks A, Rosenbauer J, Holl RW, et al. Hospitalization among diabetic children and adolescents and the general population in Germany. German Working Group for Pediatric Diabetology. Diabetes Care. 2001;24(3):435-440.

63. Sayers A, Thayer D, Harvey JN, et al. Evidence for a persistent, major excess in all cause admissions to hospital in children with type-1 diabetes: results from a large Welsh national matched community cohort study. BMJ Open. 2015;5(4):e005644.

64. Global Asthma Network. The Global Asthma Report 2014. 2014; http://www.globalasthmareport.org/resources/Global_Asthma_Report_2014.pdf. Accessed 15/06/2017.

65. Subbarao P, Mandhane PJ, Sears MR. Asthma: epidemiology, etiology and risk factors. Canadian Medical Association Journal. 2009;181(9):E181-E190.

66. American Thoracic Society Patient Information Series. What is asthma? 2013; https://www.thoracic.org/patients/patient-resources/resources/asthma.pdf. Accessed 21/06/2017.

67. Healthcare Improvement Scotland. SIGN/BTS British guideline on the management of asthma. 2013; https://www.asthma.org.uk/globalassets/campaigns/healthcare-improvement-scotland_asthma-priorities.pdf. Accessed 04/06/2017.

68. British Lung Foundation. Asthma Statistics. 2017; https://statistics.blf.org.uk/asthma. Accessed 15/06/2017.

69. Collins JE, Gill TK, Chittleborough CR, Martin AJ, Taylor AW, Winefield H. Mental, emotional, and social problems among school children with asthma. Journal of Asthma. 2008;45(6):489-493.

70. Kohen DE. Asthma and school functioning. Health Reports. 2010;21(4):35-45.

71. Liberty KA, Pattemore P, Reid J, Tarren-Sweeney M. Beginning school with asthma independently predicts low achievement in a prospective cohort of children. Chest. 2010;138(6):1349-1355.

72. Mazurek JM, Schleiff PL, Henneberger PK. Is childhood asthma associated with educational level and longest-held occupation? American Journal of Epidemiology. 2012;175(4):279-288.

73. Tsakiris A, Iordanidou M, Paraskakis E, et al. The presence of asthma, the use of inhaled steroids, and parental education level affect school performance in children. BioMed research international. 2013;2013:762805.

74. Mohammadbeigi A, Hassanzadeh J, Mousavizadeh A. Prevalence of asthma in elementary school age children in Iran--a systematic review and meta analysis study. Pakistan Journal of Biological Sciences. 2011;14(19):887-893.

75. Nelson KA, Meadows L, Yan Y, Schootman M, Strunk RC. Asthma prevalence in low-income urban elementary school students in St. Louis, 1992 and 2004. Journal of Pediatrics. 2009;154(1):111-115.

76. Sharma BS, Kumar MG, Chandel R. Prevalence of asthma in urban school children in Jaipur, Rajasthan. Indian Pediatrics. 2012;49(10):835-836.

77. Werneck G, Ruiz S, Hart R, White M, Romieu I. Prevalence of asthma and other childhood allergies in Brazilian schoolchildren. Journal of Asthma. 1999;36(8):677-690.

78. Millard MW, Johnson PT, Hilton A, Hart M. Children with asthma miss more school: fact or fiction? Chest. 2009;135(2):303-306.

79. Moonie S, Cross CL, Guillermo CJ, Gupta T. Grade retention risk among children with asthma and other chronic health conditions in a large urban school district. Postgraduate Medicine. 2010;122(5):110-115.

80. Lindgren S, Lokshin B, Stromquist A, et al. Does asthma or treatment with theophylline limit children's academic performance? New England Journal of Medicine. 1992;327(13):926-930.

81. Silverstein MD, Mair JE, Katusic SK, Wollan PC, O'Connell EJ, Yunginger JW. School attendance and school performance: a population-based study of children with asthma. Journal of Pediatrics. 2001;139(2):278-283.

82. Moonie S, Sterling DA, Figgs LW, Castro M. The relationship between school absence, academic performance, and asthma status. Journal of School Health. 2008;78(3):140-148.

83. Moonie SA, Sterling DA, Figgs L, Castro M. Asthma status and severity affects missed school days. Journal of School Health. 2006;76(1):18-24.

84. Fowler MG, Davenport MG, Garg R. School functioning of US children with asthma. Pediatrics. 1992;90(6):939-944.

85. Krenitsky-Korn S. High school students with asthma: attitudes about school health, absenteeism, and its impact on academic achievement. Pediatric Nursing. 2011;37(2):61-68.

86. Rietveld S, Colland VT. The impact of severe asthma on schoolchildren. J Asthma. 1999;36(5):409-417.

87. Roder I, Kroonenberg PM, Boekaerts M. Psychosocial functioning and stress-processing of children with asthma in the school context: differences and similarities with children without asthma. Journal of Asthma. 2003;40(7):777-787.

88. Bonilla S, Kehl S, Kwong KY, Morphew T, Kachru R, Jones CA. School absenteeism in children with asthma in a Los Angeles inner city school. Journal of Pediatrics. 2005;147(6):802-806.

89. Carpentier MY, Mullins LL, Van Pelt JC. Psychological, academic, and work functioning in college students with childhood-onset asthma. Journal of Asthma. 2007;44(2):119-124.

90. Sibbald B, Anderson HR, McGuigan S. Asthma and employment in young adults. Thorax. 1992;47(1):19-24.

91. Akinbami LJ, Schoendorf KC. Trends in childhood asthma: prevalence, health care utilization, and mortality. Pediatrics. 2002;110(2 Pt 1):315-322.

92. Anderson HR, Gupta R, Kapetanakis V, et al. International correlations between indicators of prevalence, hospital admissions and mortality for asthma in children. International Journal of Epidemiology. 2008;37(3):573-582.

93. To T, Dell S, Tassoudji M, Wang C. Health outcomes in low-income children with current asthma in Canada. Chronic diseases in Canada. 2009;29(2):49-55.

94. Saynajakangas O, Valmari P, Tuuponen T, Keistinen T. Trends in hospitalization for childhood asthma in Finland in 1996-2004. Acta Paediatrica. 2007;96(6):919-923.

95. Chua KL, Ma S, Prescott S, Ho MH, Ng DK, Lee BW. Trends in childhood asthma hospitalisation in three Asia Pacific countries. Journal of Paediatrics & Child Health. 2011;47(10):723-727.

96. Prietsch SO, Zhang L, Catharino AR, Vauchinski L, Rodrigues FE. Asthma mortality among Brazilian children up to 19 years old between 1980 and 2007. Jornal de Pediatria. 2012;88(5):384-388.

97. Hasegawa K, Tsugawa Y, Brown DF, Camargo CA, Jr. Childhood asthma hospitalizations in the United States, 2000-2009. Journal of Pediatrics. 2013;163(4):1127-1133.e1123.

98. Fitzgerald DA, Gillis J. Asthma deaths in children in New South Wales 2004-2013: Could we have done more? Journal of Paediatrics & Child Health. 2015;51(11):1127-1133.

99. Schwebel DC, Brezausek CM. Injury risk in children with asthma. Journal of Asthma. 2009;46(6):560-563.

100. Kuo CJ, Chen VC, Lee WC, et al. Asthma and suicide mortality in young people: a 12-year follow-up study. American Journal of Psychiatry. 2010;167(9):1092-1099.

101. Fisher RS, Acevedo C, Arzimanoglou A, et al. ILAE Official Report: A practical clinical definition of epilepsy. Epilepsia. 2014;55(4):475-482.

102. Epilepsy Action. Epilepsy: An introduction. 2012. https://www.epilepsy.org.uk/sites/epilepsy/files/Epilepsy-an-introduction.pdf. Accessed 15/06/2017.

103. International League Against Epilepsy. Epilepsy 2015: A practical guide to epilepsy. 2015; http://ilaebritish.org.uk/wp-content/uploads/2016/03/Epilepsy-2015-A-Practical-Guide.pdf. Accessed 15/06/2017.

104. Epilepsy action. Epileptic seizures explained. 2014; https://www.epilepsy.org.uk/sites/epilepsy/files/H125%20-%20BOOKLET%20-%20SEIZURES%20EXPLAINED.pdf. Accessed 28/06/2017.

105. Epilepsy Action Australia. Seizure smart. 2012; http://www.epilepsy.org.au/sites/default/files/Seizure%20Smart%20-%20Classification%20of%20Seizures%20(focal)_0.pdf. Accessed 21/06/2017.

106. Joint Epilepsy Council. Epilepsy prevalence, incidence and other statistics. 2011; http://www.epilepsyscotland.org.uk/pdf/Joint_Epilepsy_Council_Prevalence_and_Incidence_September_11_(3).pdf. Accessed 15,06,2017.

107. Berg AT, Jallon P, Preux PM. The epidemiology of seizure disorders in infancy and childhood: definitions and classifications. Handbook of clinical neurology. 2013;111:391-398.

108. Camfield P, Camfield C. Incidence, prevalence and aetiology of seizures and epilepsy in children. Epileptic Disord. 2015;17(2):117-123.

109. Epilepsy Scotland. Epilepsy and Treatment. 2008; http://www.epilepsyscotland.org.uk/pdf/EpilepsyTreatment.pdf. Accessed 15/06/2017.

110. Fastenau PS, Jianzhao S, Dunn DW, Austin JK. Academic underachievement among children with epilepsy: proportion exceeding psychometric criteria for learning disability and associated risk factors. Journal of Learning Disabilities. 2008;41(3):195-207.

111. Mitchell WG, Chavez JM, Lee H, Guzman BL. Academic underachievement in children with epilepsy. Journal of Child Neurology. 1991;6(1):65-72.

112. Adewuya AO, Oseni SB, Okeniyi JA. School performance of Nigerian adolescents with epilepsy. Epilepsia. 2006;47(2):415-420.

113. Nuhu FT, Yusuf AJ, Sheikh TL, Eseigbe EE. Poor academic performance among adolescents with epilepsy in Kaduna, northern Nigeria: A case-control study. SAJCH South African Journal of Child Health. 2012;6(4):132-134.

114. Singh H, Aneja S, Unni KE, Seth A, Kumar V. A study of educational underachievement in Indian children with epilepsy. Brain & Development. 2012;34(6):504-510.

115. Austin JK, Huberty TJ, Huster GA, Dunn DW. Academic achievement in children with epilepsy or asthma. Developmental Medicine & Child Neurology. 1998;40(4):248-255.

116. Suurmeijer T. Treatment, seizure-free periods, and educational achievements: a follow-up study among children with epilepsy and healthy children. Family Practice. 1991;8(4):320-328.

117. Melbourne Chambers R, Morrison-Levy N, Chang S, Tapper J, Walker S, Tulloch-Reid M. Cognition, academic achievement, and epilepsy in school-age children: a case-control study in a developing country. Epilepsy & Behavior. 2014;33:39-44.

118. Oliveira EP, Neri ML, Medeiros LL, Guimaraes CA, Guerreiro MM. School performance and praxis assessment in children with Rolandic Epilepsy. Profono. 2010;22(3):209-214.

119. Miziara CS, de Manreza ML, Mansur L, et al. Impact of benign childhood epilepsy with centrotemporal spikes (BECTS) on school performance. Seizure. 2012;21(2):87-91.

120. Dunn DW, Johnson CS, Perkins SM, et al. Academic problems in children with seizures: relationships with neuropsychological functioning and family variables during the 3 years after onset. Epilepsy & Behavior. 2010;19(3):455-461.

121. Berg AT, Hesdorffer DC, Zelko FA. Special education participation in children with epilepsy: what does it reflect? Epilepsy & Behavior. 2011;22(2):336-341.

122. Ibekwe RC, Ojinnaka NC, Iloeje SO. Factors influencing the academic performance of school children with epilepsy. Journal of Tropical Pediatrics. 2007;53(5):338-343.

123. Germano E, Gagliano A, Magazu A, et al. Benign childhood epilepsy with occipital paroxysms: neuropsychological findings. Epilepsy Research. 2005;64(3):137-150.

124. Bailet LL, Turk WR. The impact of childhood epilepsy on neurocognitive and behavioral performance: A prospective longitudinal study. Epilepsia. 2000;41(4):426-431.

125. Oostrom KJ, Smeets-Schouten A, Kruitwagen CL, Peters AC, Jennekens-Schinkel A, Dutch Study Group of Epilepsy in C. Not only a matter of epilepsy: early problems of cognition and behavior in children with "epilepsy only"--a prospective, longitudinal, controlled study starting at diagnosis. Pediatrics. 2003;112(6 Pt 1):1338-1344.

126. Williams J, Phillips T, Griebel ML, et al. Factors Associated with Academic Achievement in Children with Controlled Epilepsy. Epilepsy & behavior : E&B. 2001;2(3):217-223.

127. Matthews WS, Barabas G, Ferrari M. Achievement and school behavior among children with epilepsy. Psychology in the Schools. 1983;20(1):10-12.

128. Conant LL, Wilfong A, Inglese C, Schwarte A. Dysfunction of executive and related processes in childhood absence epilepsy. Epilepsy & Behavior. 2010;18(4):414-423.

129. Chambers RM, Morrison-Levy N, Chang S, Tapper J, Walker S, Tulloch-Reid M. Cognition, academic achievement, and epilepsy in school-age children: A case–control study in a developing country. Epilepsy & Behavior. 2014;33:39-44.

130. Pavlou E, Gkampeta A. Learning disorders in children with epilepsy. Childs Nervous System. 2011;27(3):373-379.

131. Seidenberg M, Beck N, Geisser M, et al. Academic achievement of children with epilepsy. Epilepsia. 1986;27(6):753-759.

132. Rathouz PJ, Zhao Q, Jones JE, et al. Cognitive development in children with new onset epilepsy. Developmental Medicine & Child Neurology. 2014;56(7):635-641.

133. Croona C, Kihlgren M, Lundberg S, Eeg-Olofsson O, Eeg-Olofsson KE. Neuropsychological findings in children with benign childhood epilepsy with centrotemporal spikes. Developmental Medicine & Child Neurology. 1999;41(12):813-818.

134. Goldberg-Stern H, Gonen OM, Sadeh M, Kivity S, Shuper A, Inbar D. Neuropsychological aspects of benign childhood epilepsy with centrotemporal spikes.[Erratum appears in Seizure. 2010 Apr;19(3):202-3]. Seizure. 2010;19(1):12-16.

135. Tosun A, Gokcen S, Ozbaran B, et al. The effect of depression on academic achievement in children with epilepsy. Epilepsy & Behavior. 2008;13(3):494-498.

136. Farwell JR, Dodrill CB, Batzel LW. Neuropsychological abilities of children with epilepsy. Epilepsia. 1985;26(5):395-400.

137. Baca CB, Vickrey BG, Vassar SD, Cook A, Berg AT. Injuries in adolescents with childhood-onset epilepsy compared with sibling controls. Journal of Pediatrics. 2013;163(6):1684-1691.e1684.

138. Kirsch R, Wirrell E. Do cognitively normal children with epilepsy have a higher rate of injury than their nonepileptic peers? Journal of Child Neurology. 2001;16(2):100-104.

139. Prasad V, Kendrick D, Sayal K, Thomas SL, West J. Injury among children and young adults with epilepsy. Pediatrics. 2014;133(5):827-835.

140. Ackers R, Besag FM, Hughes E, Squier W, Murray ML, Wong IC. Mortality rates and causes of death in children with epilepsy prescribed antiepileptic drugs: a retrospective cohort study using the UK General Practice Research Database. Drug safety. 2011;34(5):403-413.

141. Callenbach PM, Westendorp RG, Geerts AT, et al. Mortality risk in children with epilepsy: the Dutch study of epilepsy in childhood. Pediatrics. 2001;107(6):1259-1263.

142. Camfield CS, Camfield PR, Veugelers PJ. Death in children with epilepsy: a population-based study. Lancet. 2002;359(9321):1891-1895.

143. Christensen J, Pedersen CB, Sidenius P, Olsen J, Vestergaard M. Long-term mortality in children and young adults with epilepsy--A population-based cohort study. Epilepsy Research. 2015;114:81-88.

144. Berg AT, Nickels K, Wirrell EC, et al. Mortality risks in new-onset childhood epilepsy. Pediatrics. 2013;132(1):124-131.

145. Harvey AS, Nolan T, Carlin JB. Community-based study of mortality in children with epilepsy. Epilepsia. 1993;34(4):597-603.

146. Nickels KC, Grossardt BR, Wirrell EC. Epilepsy-related mortality is low in children: a 30-year population-based study in Olmsted County, MN. Epilepsia. 2012;53(12):2164-2171.

147. Sillanpaa M, Shinnar S. Long-term mortality in childhood-onset epilepsy. New England Journal of Medicine. 2010;363(26):2522-2529.

148. Scottish Intercollegiate Guidelines Network. Management of attention deficit and hyperkinetic disorders in children and young people: A national clinical guideline. 2009; http://www.sign.ac.uk/sign-112-management-of-attention-deficit-and-hyperkinetic-disorders-in-children-and-young-people.html. Accessed 21/06/2017.

149. Healthcare Improvement Scotland. Attention deficit and hyperactivity disorders, Services over Scotland, Final report. 2012; http://www.healthcareimprovementscotland.org/our_work/mental_health/adhd_services_over_scotland/stage_3_adhd_final_report.aspx. Accessed 21/06/2017.

150. DSM-5 American Psychiatric Association. Diagnostic and statistical manual of mental disorders. Arlington: American Psychiatric Publishing. 2013.

151. Rowland AS, Lesesne CA, Abramowitz AJ. The epidemiology of attention-deficit/hyperactivity disorder (ADHD): a public health view. Mental retardation and developmental disabilities research reviews. 2002;8(3):162-170.

152. Cantwell DP. Attention Deficit Disorder: A Review of the Past 10 Years. Journal of the American Academy of Child & Adolescent Psychiatry. 1996;35(8):978-987.

153. Biederman JN, J.; Sprich, S. Comorbidity of attention deficit hyperactivity disorder with conduct, depressive, anxiety, and other disorders. American Journal of Psychiatry. 1991;148(5):564-577.

154. Pliszka SR. Comorbidity of Attention-Deficit/Hyperactivity Disorder With Psychiatric Disorder: An Overview. The Journal of clinical psychiatry. 1998.

155. Global Burden of Disease Study 2013 Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet. 2015;386(9995):743-800.

156. Polanczyk GV, Willcutt EG, Salum GA, Kieling C, Rohde LA. ADHD prevalence estimates across three decades: an updated systematic review and meta-regression analysis. International Journal of Epidemiology. 2014;43(2):434-442.

157. Skounti M, Philalithis A, Galanakis E. Variations in prevalence of attention deficit hyperactivity disorder worldwide. European Journal of Pediatrics. 2007;166(2):117-123.

158. Willcutt EG. The Prevalence of DSM-IV Attention-Deficit/Hyperactivity Disorder: A Meta-Analytic Review. Neurotherapeutics : the journal of the American Society for Experimental NeuroTherapeutics. 2012;9(3):490-499.

159. Guilherme Polanczyk MD, Maurício Silva de Lima MD, Ph.D., Bernardo Lessa Horta MD, Ph.D., Joseph Biederman MD, Luis Augusto Rohde MD, Ph.D. The Worldwide Prevalence of ADHD: A Systematic Review and Metaregression Analysis. American Journal of Psychiatry. 2007;164(6):942-948.

160. Amiri S, Fakhari A, Maheri M, Mohammadpoor Asl A. Attention deficit/hyperactivity disorder in primary school children of Tabriz, North-West Iran. Paediatric and Perinatal Epidemiology. 2010;24(6):597-601.

161. Rowland AS, Umbach DM, Stallone L, Naftel AJ, Bohlig EM, Sandler DP. Prevalence of Medication Treatment for Attention Deficit–Hyperactivity Disorder Among Elementary School Children in Johnston County, North Carolina. American Journal of Public Health. 2002;92(2):231-234.

162. Dalsgaard S, Nielsen HS, Simonsen M. Consequences of ADHD medication use for children's outcomes. Journal of health economics. 2014;37:137-151.

163. Beau-Lejdstrom R, Douglas I, Evans SJW, Smeeth L. Latest trends in ADHD drug prescribing patterns in children in the UK: prevalence, incidence and persistence. BMJ Open. 2016;6(6).

164. Advokat C, Lane SM, Luo C. College Students With and Without ADHD: Comparison of Self-Report of Medication Usage, Study Habits, and Academic Achievement. Journal of attention disorders. 2011;15(8):656-666.

165. Scheffler RM, Brown TT, Fulton BD, Hinshaw SP, Levine P, Stone S. Positive Association Between Attention-Deficit/ Hyperactivity Disorder Medication Use and Academic Achievement During Elementary School. Pediatrics. 2009;123(5):1273-1279.

166. Barkley RA, Anastopoulos AD, Guevremont DC, Fletcher KE. Adolescents with ADHD: Patterns of behavioral adjustment, academic functioning, and treatment utilization. Journal of the American Academy of Child & Adolescent Psychiatry. 1991;30(5):752-761.

167. Bussing R, Porter P, Zima BT, Mason D, Garvan C, Reid R. Academic outcome trajectories of students with ADHD: Does exceptional education status matter? Journal of Emotional and Behavioral Disorders. 2012;20(3):131-143.

168. DeShazo Barry T, Lyman RD, Klinger LG. Academic underachievement and attention-deficit/hyperactivity disorder: The negative impact of symptom severity on school performance. Journal of school psychology. 2002;40(3):259-283.

169. Kakouros E, Maniadaki K, Papaeliou C. How Greek teachers perceive school functioning of pupils with ADHD. Emotional & Behavioural Difficulties. 2004;9(1):41-53.

170. McConaughy SH, Volpe RJ, Antshel KM, Gordon M, Eiraldi RB. Academic and social impairments of elementary school children with attention deficit hyperactivity disorder. School Psychology Review. 2011;40(2):200-225.

171. Massetti GM, Lahey BB, Pelham WE, et al. Academic achievement over 8 years among children who met modified criteria for attention-deficit/hyperactivity disorder at 4-6 years of age. Journal of abnormal child psychology. 2008;36(3):399-410.

172. Barbaresi WJ, Katusic SK, Colligan RC, Weaver AL, Jacobsen SJ. Long-term school outcomes for children with attention-deficit/hyperactivity disorder: A population-based perspective. Journal of Developmental and Behavioral Pediatrics. 2007;28(4):265-273.

173. Wu SY, Gau SS. Correlates for academic performance and school functioning among youths with and without persistent attention-deficit/hyperactivity disorder. Research in Developmental Disabilities. 2013;34(1):505-515.

174. Ek U, Westerlund J, Holmberg K, Fernell E. Academic performance of adolescents with ADHD and other behavioural and learning problems - A population-based longitudinal study. Acta Paediatrica, International Journal of Paediatrics. 2011;100(3):402-406.

175. Powers RL, Marks DJ, Miller CJ, Newcorn JH, Halperin JM. Stimulant treatment in children with attention-deficit/hyperactivity disorder moderates adolescent academic outcome. Journal of Child & Adolescent Psychopharmacology. 2008;18(5):449-459.

176. Faraone SV, Biederman J, Lehman BK, et al. Intellectual performance and school failure in children with attention deficit hyperactivity disorder and in their siblings. Journal of abnormal psychology. 1993;102(4):616-623.

177. Sievertsen HH, Gino F, Piovesan M. Cognitive fatigue influences students' performance on standardized tests. Proceedings of the National Academy of Sciences of the United States of America. 2016;113(10):2621-2624.

178. Taanila A, Ebeling H, Tiihala M, et al. Association between childhood specific learning difficulties and school performance in adolescents with and without ADHD symptoms: a 16-year follow-up. Journal of attention disorders. 2014;18(1):61-72.

179. LeFever GB, Villers MS, Morrow AL, Vaughn ES. Parental perceptions of adverse educational outcomes among children diagnosed and treated for ADHD: A call for improved school/provider collaboration. Psychology in the Schools. 2002;39(1):63-71.

180. Rohde LA, Biederman J, Busnello EA, et al. ADHD in a school sample of Brazilian adolescents: A study of prevalence, comorbid conditions, and impairments. Journal of the American Academy of Child & Adolescent Psychiatry. 1999;38(6):716-722.

181. Kuriyan AB, Pelham WE, Jr., Molina BS, et al. Young adult educational and vocational outcomes of children diagnosed with ADHD. Journal of abnormal child psychology. 2013;41(1):27-41.

182. Kent KM, Pelham WE, Jr., Molina BS, et al. The academic experience of male high school students with ADHD. Journal of abnormal child psychology. 2011;39(3):451-462.

183. McCarthy S, Cranswick N, Potts L, Taylor E, Wong ICK. Mortality Associated with Attention-Deficit Hyperactivity Disorder (ADHD) Drug Treatment. Drug safety. 2009;32(11):1089-1096.

184. DiScala C, Lescohier I, Barthel M, Li G. Injuries to children with attention deficit hyperactivity disorder. Pediatrics. 1998;102.

185. Dalsgaard S, Ostergaard SD, Leckman JF, Mortensen PB, Pedersen MG. Mortality in children, adolescents, and adults with attention deficit hyperactivity disorder: a nationwide cohort study. Lancet. 2015;385(9983):2190-2196.

186. Barbaresi WJ, Colligan RC, Weaver AL, Voigt RG, Killian JM, Katusic SK. Mortality, ADHD, and psychosocial adversity in adults with childhood ADHD: a prospective study. Pediatrics. 2013;131(4):637-644.

187. Schelleman H, Bilker WB, Strom BL, et al. Cardiovascular events and death in children exposed and unexposed to ADHD agents. Pediatrics. 2011;127(6):1102-1110.

188. van den Ban E, Souverein P, Meijer W, et al. Association between ADHD drug use and injuries among children and adolescents. European child & adolescent psychiatry. 2014;23(2):95-102.

189. Guevara J, Lozano P, Wickizer T, Mell L, Gephart H. Utilization and Cost of Health Care Services for Children With Attention-Deficit/Hyperactivity Disorder. Pediatrics. 2001;108(1):71-78.

190. Birnbaum HG, Kessler RC, Lowe SW, et al. Costs of attention deficit–hyperactivity disorder (ADHD) in the US: excess costs of persons with ADHD and their family members in 2000. Current Medical Research and Opinion. 2005;21(2):195-205.

191. Hakkaart-van Roijen L, Zwirs BWC, Bouwmans C, et al. Societal costs and quality of life of children suffering from attention deficient hyperactivity disorder (ADHD). European child & adolescent psychiatry. 2007;16(5):316-326.

192. Hurtig T, Ebeling H, Jokelainen J, Koivumaa-Honkanen H, Taanila A. The Association Between Hospital-Treated Injuries and ADHD Symptoms in Childhood and Adolescence: A Follow-Up Study in the Northern Finland Birth Cohort 1986. Journal of attention disorders. 2016;20(1):3-10.

193. Shilon Y, Pollak Y, Aran A, Shaked S, Gross-Tsur V. Accidental injuries are more common in children with attention deficit hyperactivity disorder compared with their non-affected siblings. Child: Care, Health & Development. 2012;38(3):366-370.

194. Pastor PN, Reuben CA. Identified attention-deficit/hyperactivity disorder and medically attended, nonfatal injuries: US school-age children, 1997-2002. Ambulatory Pediatrics. 2006;6(1):38-44.

195. Loe IM, Feldman HM. Academic and educational outcomes of children with ADHD. Journal of Pediatric Psychology. 2007;32(6):643-654.

196. Daley D, Birchwood J. ADHD and academic performance: why does ADHD impact on academic performance and what can be done to support ADHD children in the classroom? Child: Care, Health & Development. 2010;36(4):455-464.

197. Thapar A, Collishaw S, Pine DS, Thapar AK. Depression in adolescence. The Lancet. 2012;379(9820):1056-1067.

198. Melnyk BM, Moldenhauer Z, Tuttle J, Veenema TG, Jones D, Novak J. Improving child and adolescent mental health. An evidence-based approach. Advance for nurse practitioners. 2003;11(2):47-52.

199. Richardson LA, Keller AM, Selby-Harrington ML, Parrish R. Identification and treatment of children's mental health problems by primary care providers: A critical review of research. Archives of psychiatric nursing. 1996;10(5):293-303.

200. Merikangas KR, Nakamura EF, Kessler RC. Epidemiology of mental disorders in children and adolescents. Dialogues in Clinical Neuroscience. 2009;11(1):7-20.

201. Hauser M, Galling B, Correll CU. Suicidal ideation and suicide attempts in children and adolescents with bipolar disorder: a systematic review of prevalence and incidence rates, correlates, and targeted interventions. Bipolar disorders. 2013;15(5):507-523.

202. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry. 2005;62(6):593-602.

203. Polanczyk GV, Salum GA, Sugaya LS, Caye A, Rohde LA. Annual Research Review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. Journal of Child Psychology and Psychiatry. 2015;56(3):345-365.

204. Birrer RB. Depression and aging too often do mix: Accurate recognition is half the battle. Postgraduate medicine. 1998;104(3):143-164.

205. Costello EJ, Egger H, Angold A. 10-Year Research Update Review: The Epidemiology of Child and Adolescent Psychiatric Disorders: I. Methods and Public Health Burden. Journal of the American Academy of Child & Adolescent Psychiatry. 2005;44(10):972-986.

206. Costello EJ, Erkanli A, Angold A. Is there an epidemic of child or adolescent depression? Journal of Child Psychology and Psychiatry. 2006;47(12):1263-1271.

207. Hyde JS, Mezulis AH, Abramson LY. The ABCs of depression: integrating affective, biological, and cognitive models to explain the emergence of the gender difference in depression. Psychological review. 2008;115(2):291-313.

208. Mojtabai R, Olfson M, Han B. National Trends in the Prevalence and Treatment of Depression in Adolescents and Young Adults. Pediatrics. 2016.

209. Nemeroff CB. The burden of severe depression: A review of diagnostic challenges and treatment alternatives. Journal of Psychiatric Research. 2007;41(3–4):189-206.

210. Birmaher B, Ryan ND, Williamson DE, Brent DA, Kaufman J. Childhood and Adolescent Depression: A Review of the Past 10 Years. Part II. Journal of the American Academy of Child & Adolescent Psychiatry. 1996;35(12):1575-1583.

211. Grunhaus L, Dannon PN, Schreiber S, et al. Repetitive transcranial magnetic stimulation is as effective as electroconvulsive therapy in the treatment of nondelusional major depressive disorder: an open study. Biological Psychiatry. 2000;47(4):314-324.

212. McCarty CA, Mason WA, Kosterman R, Hawkins JD, Lengua LJ, McCauley E. Adolescent School Failure Predicts Later Depression Among Girls. Journal of Adolescent Health. 2008;43(2):180-187.

213. Fergusson DM, Woodward LJ. Mental health, educational, and social role outcomes of adolescents with depression. Archives of General Psychiatry. 2002;59(3):225-231.

214. Fletcher JM. Adolescent depression and educational attainment: results using sibling fixed effects. Health economics. 2010;19(7):855-871.

215. Jonsson U, Bohman H, Hjern A, von Knorring L, Olsson G, von Knorring AL. Subsequent higher education after adolescent depression: a 15-year follow-up register study. European Psychiatry: the Journal of the Association of European Psychiatrists. 2010;25(7):396-401.

216. Arnett JJ. Adolescent storm and stress, reconsidered. The American psychologist. 1999;54(5):317-326.

217. Verboom CE, Sijtsema JJ, Verhulst FC, Penninx BW, Ormel J. Longitudinal associations between depressive problems, academic performance, and social functioning in adolescent boys and girls. Developmental psychology. 2014;50(1):247-257.

218. Dekker MC, Ferdinand RF, van Lang ND, Bongers IL, van der Ende J, Verhulst FC. Developmental trajectories of depressive symptoms from early childhood to late adolescence: gender differences and adult outcome. Journal of Child Psychology & Psychiatry & Allied Disciplines. 2007;48(7):657-666.

219. Brumback RA, Jackoway MK, Weinberg WA. Relation of intelligence to childhood depression in children referred to an educational diagnostic center. Perceptual and Motor Skills. 1980;50(1):11-17.

220. Hamilton EB, Asarnow JR, Tompson MC. Social, academic, and behavioral competence of depressed children: Relationship to diagnostic status and family interaction style. Journal of Youth and Adolescence. 1997;26(1):77-87.

221. Vincenzi H. Depression and reading ability in sixth-grade children. Journal of school psychology. 1987;25(2):155-160.

222. Tesiny EP, Lefkowitz MM, Gordon NH. Childhood depression, locus of control, and school achievement. Journal of educational psychology. 1980;72(4):506-510.

223. Briere FN, Janosz M, Fallu JS, Morizot J. Adolescent Trajectories of Depressive Symptoms: Codevelopment of Behavioral and Academic Problems. Journal of Adolescent Health. 2015;57(3):313-319.

224. Shahar G, Henrich CC, Winokur A, Blatt SJ, Kuperminc GP, Leadbeater BJ. Self-criticism and depressive symptomatology interact to predict middle school academic achievement. Journal of Clinical Psychology. 2006;62(1):147-155.

225. Lundy SM, Silva GE, Kaemingk KL, Goodwin JL, Quan SF. Cognitive Functioning and Academic Performance in Elementary School Children with Anxious/Depressed and Withdrawn Symptoms. The open pediatric medicine journal. 2010;4:1-9.

226. Marcotte D, Levesque N, Fortin L. Variations of cognitive distortions and school performance in depressed and non-depressed high school adolescents: A two-year longitudinal study. Cognitive Therapy and Research. 2006;30(2):211-225.

227. Miller DK, Constance HL, Brennan PA. Health Outcomes Related to Early Adolescent Depression. The Journal of adolescent health : official publication of the Society for Adolescent Medicine. 2007;41(3):256-262.

228. Min HJ, Jon DI, Jung MH, et al. Depression, aggression, and suicidal ideation in first graders: a school-based cross-sectional study. Comprehensive psychiatry. 2012;53(8):1145-1152.

229. Han PP, Holbrook TL, Sise MJ, et al. Postinjury depression is a serious complication in adolescents after major trauma: injury severity and injury-event factors predict depression and long-term quality of life deficits. J Trauma. 2011;70(4):923-930.

230. Asbridge M, Azagba S, Langille DB, Rasic D. Elevated depressive symptoms and adolescent injury: examining associations by injury frequency, injury type, and gender. BMC Public Health. 2014;14:190.

231. You J, Leung F. The role of depressive symptoms, family invalidation and behavioral impulsivity in the occurrence and repetition of non-suicidal self-injury in Chinese adolescents: a 2-year follow-up study. Journal of adolescence. 2012;35(2):389-395.

232. Winkler E. Overview of record linkage and current research directions. 2006; https://www.census.gov/srd/papers/pdf/rrs2006-02.pdf. Accessed 27/05/2017.

233. Statistical Methods Team. Data Integration Manual. 2006; http://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwjDr-71xc_UAhULIsAKHQciAosQFggiMAA&url=http%3A%2F%2Fwww.stats.govt.nz%2F~%2Fmedia%2FStatistics%2Fabout-us%2Fpolicies-protocols-guidelines%2Fdata-integration-further-technical-info%2FDataIntegrationManual.pdf&usg=AFQjCNGTNFATXC0iJ5lGLLpzZ3D4hbWujw. Accessed 21/06/2017.

234. Kendrick S, Clarke J. The Scottish record linkage system. Health bulletin. 1993;51(2):72.

235. Kendrick S, Douglas M, Gardner D, Hucker D. Best-link matching of Scottish health data sets. Methods of information in medicine. 1998;37(1):64-68.

236. Fleming M, Kirby B, Penny KI. Record linkage in Scotland and its applications to health research. Journal of Clinical Nursing. 2012;21(19pt20):2711-2721.

237. Wood R, Clark D, King A, Mackay D, Pell J. Novel cross-sectoral linkage of routine health and education data at an all-Scotland level: a feasibility study. The Lancet. 2013;382:S10.

238. The Farr Institute. Who we are. 2017; http://www.farrinstitute.org/about/who-we-are. Accessed 27/05/2017.

239. NHS National Services Scotland. How NSS works. 2017; https://nhsnss.org/how-nss-works/. Accessed 20/06/2017.

240. The Scottish Government. Scottish Exchange of Data (ScotXed). 2017; http://www.gov.scot/Topics/Statistics/ScotXed.

241. The Scottish Government. Pupil Census Results. 2017; http://www.gov.scot/Topics/Statistics/Browse/School-Education/PubPupilCensus. Accessed 09/06/2017.

242. Scottish Qualifications Authority. 2017; http://www.sqa.org.uk/sqa/70972.html. Accessed 15/05/2017.

243. The Scottish Government. Destinations of leavers from Scottish schools. 2017; http://www.gov.scot/Topics/Statistics/Browse/School-Education/PubDestinationLeavers. Accessed 09/06/2017.

244. The Scottish Government. Initial destinations of senior phase school leavers. 2017; http://www.gov.scot/Publications/2017/03/2421/0. Accessed 09/06/2017.

245. Scottish Council of Independent Schools. Pupils numbers. 2016; http://www.scis.org.uk/facts-and-figures/pupil-numbers/. Accessed 22/06/2017.

246. Information Services Division. 2017; http://www.isdscotland.org/. Accessed 15/06/2015.

247. Scottish Government eHealth Division. The use of the CHI (Community Health Index) to support integrated care across the NHS in Scotland. 2013; http://www.ehealth.nhs.scot/wp-content/uploads/sites/7/documents/Health-and-social-care-CHI-Guidance-version-1-1-Strategy-Board-Approved-June-20131.pdf. Accessed 21/06/2017.

248. Public Benefit and Privacy Panel for Health and Social Care. http://www.informationgovernance.scot.nhs.uk/pbpphsc/. Accessed 09/06/2017.

249. Information Services Division. eDRIS Products and Services. 2017; http://www.isdscotland.org/Products-and-Services/EDRIS/. Accessed 09/06/2017.

250. Citrix Systems. 2017; https://www.citrix.co.uk/about/. Accessed 22/06/2017.

251. University of Edinburgh. About EPCC. 2016; https://www.epcc.ed.ac.uk/about. Accessed 22/06/2017.

252. NHS National Services Scotland. Data Quality Assurance. Assessment of Maternity Data (SMR02) 2008-09. 2010; http://www.isdscotland.org/Products-and-Services/Data-Quality/Previous-Projects/DQA-Assessment-of-Maternity-Data-SMR02-2008-to-2009.pdf. Accessed 21/06/2017.

253. The Scottish Government. Attendance and absence in Scottish schools. 2015; http://www.gov.scot/Topics/Statistics/Browse/School-Education/PubAttendanceAbsence. Accessed 21/06/2017.

254. Department for Education. Special Educational Needs Code of Practice. 2001; http://webarchive.nationalarchives.gov.uk/20130401151715/https://www.education.gov.uk/publications/eorderingdownload/dfes%200581%20200mig2228.pdf. Accessed 21/06/2017.

255. Office for National Statistics. Statistical bulletin: Young people not in education, employment or training (NEET), UK: Feb 2017. 2017; https://www.ons.gov.uk/employmentandlabourmarket/peoplenotinwork/unemployment/bulletins/youngpeoplenotineducationemploymentortrainingneet/feb2017. Accessed 09/06/2017.

256. The Scottish Government. Curriculum for Excellence. 2016; http://www.gov.scot/Topics/Education/Schools/curriculum. Accessed 09/06/2017.

257. Scottish Credit and Qualifications Framework. 2017; http://scqf.org.uk/. Accessed 15/06/2017.

258. The Scottish Government. Scottish Index of Multiple Deprivation (SIMD) Background and methodology. 2016; http://www.gov.scot/Topics/Statistics/SIMD/BackgroundMethodology. Accessed 21/06/2017.

259. World Health Organization. The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines. Vol 1: World Health Organization; 1992.

260. Twisk JW. Applied longitudinal data analysis for epidemiology: a practical guide: Cambridge University Press; 2013.

261. Hardin J. Generalized estimating equations. Hardin JW, Hilbe J. Generalized estimating equations: New York: Chapman & Hall; 2003.

262. Hubbard AE, Ahern J, Fleischer NL, et al. To GEE or not to GEE: comparing population average and mixed models for estimating the associations between neighborhood risk factors and health. Epidemiology. 2010;21(4):467-474.

263. Cui J. QIC program and model selection in GEE analyses. Stata Journal. 2007;7(2):209-220.

264. Hilbe JM. Modeling count data: Springer; 2011.

265. Williams R. Generalized ordered logit/partial proportional odds models for ordinal dependent variables. Stata Journal. 2006;6(1):58-82.

266. Kuha J. AIC and BIC: Comparisons of Assumptions and Performance. Sociological methods & research. 2004;33(2):188-229.

267. Royston P, Lambert P. Flexible parametric survival analysis using Stata: beyond the Cox model. 2011. College Station, Tx: Stata Press Google Scholar. 2011.

268. Canchola AJ, Stewart SL, Bernstein L, et al. Cox regression using different time-scales. Western Users of SAS Software. San Francisco, California. 2003.

269. Griffin BA, Anderson GL, Shih RA, Whitsel EA. Use of alternative time scales in Cox proportional hazard models: implications for time-varying environmental exposures. Statistics in medicine. 2012;31(27):3320-3327.

270. Kleinbaum DG, Klein M. Survival analysis: a self-learning text: Springer Science & Business Media; 2006.

271. Hauser WA. Seizure Disorders: The Changes With Age. Epilepsia. 1992;33:6-14.

272. Borus JS, Laffel L. Adherence challenges in the management of type 1 diabetes in adolescents: prevention and intervention. Current opinion in pediatrics. 2010;22(4):405-411.

273. Cockerell OC, Sander J, Hart YM, Shorvon SD, Johnson A. Remission of epilepsy: results from the National General Practice Study of Epilepsy. The Lancet. 1995;346(8968):140-144.

274. Lam LT. Attention deficit disorder and hospitalization owing to intra- and interpersonal violence among children and young adolescents. Journal of Adolescent Health. 2005;36(1):19-24.

275. Hawton K, Rodham K, Evans E, Weatherall R. Deliberate self harm in adolescents: self report survey in schools in England. Bmj. 2002;325(7374):1207-1211.

276. O'Connor RC, Rasmussen S, Miles J, Hawton K. Self-harm in adolescents: self-report survey in schools in Scotland. The British Journal of Psychiatry. 2009;194(1):68-72.

277. Brunner R, Parzer P, Haffner J, et al. Prevalence and psychological correlates of occasional and repetitive deliberate self-harm in adolescents. Archives of pediatrics & adolescent medicine. 2007;161(7):641-649.

278. Hankin BL, Abramson LY, Moffitt TE, Silva PA, McGee R, Angell KE. Development of depression from preadolescence to young adulthood: Emerging gender differences in a 10-year longitudinal study. Journal of abnormal psychology. 1998;107(1):128-140.

279. Nolen-Hoeksema S, Girgus JS. The emergence of gender differences in depression during adolescence. Psychol Bull. 1994;115(3):424-443.

280. Twenge JM, Nolen-Hoeksema S. Age, gender, race, socioeconomic status, and birth cohort differences on the children's depression inventory: a meta-analysis. Journal of abnormal psychology. 2002;111(4):578-588.

281. Wade TJ, Cairney J, Pevalin DJ. Emergence of gender differences in depression during adolescence: national panel results from three countries. J Am Acad Child Adolesc Psychiatry. 2002;41(2):190-198.

282. Lhatoo SD, Sander JWAS. The Epidemiology of Epilepsy and Learning Disability. Epilepsia. 2001;42:6-9.

283. Huntington DD, Bender WN. Adolescents with Learning Disabilities at Risk? Emotional Well-Being, Depression, Suicide. Journal of Learning Disabilities. 1993;26(3):159-166.

284. Zhang D, Katsiyannis A, Herbst M. Disciplinary exclusions in special education: A 4-year analysis. Behavioral Disorders. 2004:337-347.

285. Angold A, Costello E. Depressive comorbidity in children and adolescents. American journal of Psychiatry. 1993;150(12):1779-1791.

286. Ettinger A, Reed M, Cramer J. Depression and comorbidity in community-based patients with epilepsy or asthma. Neurology. 2004;63(6):1008-1014.

287. Gaitatzis A, Trimble M, Sander JW. The psychiatric comorbidity of epilepsy. Acta Neurologica Scandinavica. 2004;110(4):207-220.

288. Egger HL, Costello EJ, Angold A. School refusal and psychiatric disorders: a community study. J Am Acad Child Adolesc Psychiatry. 2003;42(7):797-807.

289. Abela JRZ, Hankin BL. Handbook of depression in children and adolescents. 2008.

290. Laye-Gindhu A, Schonert-Reichl KA. Nonsuicidal Self-Harm Among Community Adolescents: Understanding the "Whats" and "Whys" of Self-Harm. Journal of Youth and Adolescence. 2005;34(5):447-457.

291. Fletcher J, Wolfe B. Long-term Consequences of Childhood ADHD on Criminal Activities. The Journal of Mental Health Policy and Economics. 2009;12(3):119-138.

292. Flory K, Molina BSG, Pelham JWE, Gnagy E, Smith B. Childhood ADHD Predicts Risky Sexual Behavior in Young Adulthood. Journal of Clinical Child & Adolescent Psychology. 2006;35(4):571-577.

293. Garland AF, Hough RL, McCabe KM, Yeh MAY, Wood PA, Aarons GA. Prevalence of Psychiatric Disorders in Youths Across Five Sectors of Care. Journal of the American Academy of Child & Adolescent Psychiatry. 2001;40(4):409-418.

294. Lambert NM, Hartsough CS. Prospective Study of Tobacco Smoking and Substance Dependencies Among Samples of ADHD and Non-ADHD Participants. Journal of Learning Disabilities. 1998;31(6):533-544.

295. Pratt TC, Cullen FT, Blevins KR, Daigle L, Unnever JD. The Relationship of Attention Deficit Hyperactivity Disorder to Crime and Delinquency: A Meta-Analysis. International Journal of Police Science & Management. 2002;4(4):344-360.

296. Thompson AL, Molina BSG, Pelham W, Gnagy EM. Risky Driving in Adolescents and Young Adults with Childhood ADHD. Journal of Pediatric Psychology. 2007;32(7):745-759.

297. McIntosh D, Kutcher S, Binder C, Levitt A, Fallu A, Rosenbluth M. Adult ADHD and comorbid depression: A consensus-derived diagnostic algorithm for ADHD. Neuropsychiatric Disease and Treatment. 2009;5:137-150.

298. Spencer T, Biederman J, Wilens T. Attention-deficit/hyperactivity disorder and comorbidity. Pediatric Clinics of North America. 1999;46(5):915-927.

299. Grey M, Whittemore R, Tamborlane W. Depression in Type 1 diabetes in children: Natural history and correlates. Journal of Psychosomatic Research. 2002;53(4):907-911.

300. Institute for Future Studies. 2017; https://www.iffs.se/en/. Accessed 01/11/2017.

301. Elsevier. 2017; https://www.elsevier.com/. Accessed 01/11/2017.

302. The Scottish Government. What is GIRFEC? 2017; http://www.gov.scot/Topics/People/Young-People/gettingitright/what-is-girfec. Accessed 20/06/2017.

303. Children and Young People (Scotland) Act 2004,. http://www.legislation.gov.uk/asp/2014/8/pdfs/asp_20140008_en.pdf. Accessed 20/06/2017.

304. The Scottish Government. Child's Plan. 2017; http://www.gov.scot/Topics/People/Young-People/gettingitright/childs-plan. Accessed 20/06/2017.

305. The Scottish Government. What is a named person? 2017; http://www.gov.scot/Topics/People/Young-People/gettingitright/named-person. Accessed 20/06/2017.

306. The Education (Additional Support for Learning) (Scotland) Act 2014. http://www.legislation.gov.uk/asp/2004/4/pdfs/asp_20040004_en.pdf. Accessed 20/06/2017.

307. The Scottish Government. The Scottish Attainment Challenge. 2017; http://www.gov.scot/Topics/Education/Schools/Raisingeducationalattainment. Accessed 20/06/2017.

308. The Scottish Government. Pupil Equity Funding. 2017; http://www.gov.scot/Topics/Education/Schools/Raisingeducationalattainment/pupilequityfund. Accessed 20/06/2017.

309. The Scottish Government. Mental Health Strategy for Scotland: 2012-2015. 2012; http://www.gov.scot/Publications/2012/08/9714. Accessed 20/06/2017.

310. Information Services Division. Data Quality Assurance Assessments. 2017; http://www.isdscotland.org/Products-and-Services/Data-Quality/Assessments/. Accessed 21/06/2017.

311. Skills Development Scotland. School Leaver Destinations. 2017; http://www.skillsdevelopmentscotland.co.uk/publications-statistics/statistics/archived-school-leaver-destinations/?page=1&statisticCategoryId=8&order=date-desc. Accessed 22/06/2017.

312. Littner MR. Chronic Obstructive Pulmonary Disease. Annals of internal medicine. 2008;148(5):ITC3-1.

313. Farrell PM. The prevalence of cystic fibrosis in the European Union. Journal of Cystic Fibrosis. 2008;7(5):450-453.

314. Dave Clark. Report on Quality of Linkage of ScotXed to Community Health Index (Incorporating Historic CHI Postcode Archive). 2015; http://www.isdscotland.org/Products-and-Services/EDRIS/Docs/20150421-Linking-ScotXed-Data.pdf. Accessed 22/06/2017.

315. Scotland's Census. Standard Outputs. 2011; http://www.scotlandscensus.gov.uk/ods-analyser/jsf/tableView/tableView.xhtml. Accessed 25/06/2017.

Appendices

Appendix 1

Letter from the Privacy Advisory Committee approving the study

Appendix1

Information Services Division
Area 151A
Gyle Square
1 South Gyle Crescent
EDINBURGH
EH12 9EB
Telephone 0131 275 6000
Fax 0131 275 7606
www.isdscotland.org

NHS National Services Scotland

Mr Michael Fleming
Senior Statistician
University of Glasgow
Institute of Health and Wellbeing
1 Lilybank Gardens
GLASGOW
G12 8RZ

Date 3 September 2014
Your Ref
Our Ref PAC 51/13

Enquiries to Janet Murrey
Extension 6954
Direct Line 0131 275 6954
Email janet.murray1@nhs.net

Dear Mr Fleming

Investigating the relationship between health and educational outcomes in children

The Privacy Advisory Committee has considered and approved your application for a data linkage in support of the above study.

Conditions applied: None

Time period: As specified

Points highlighted: The sensitivity and wide volume of data requested was noted. It was noted that this is a pathfinder project and that part of the output should include an analysis of the privacy risks and a demonstration of how these were justified by the learning obtained. For example that all the detail requested was in fact necessary.

The lack of lay involvement in this specific project and a reliance on generic information leaflets and public consultation through the Farr Institute was noted. It was suggested that engagement of the public for example through parent groups and other support groups might ensure measured use of data, raise public awareness of the purposes of research, seek feedback and demonstrate how privacy risks have been recognised and minimised in project design.

It would be useful to have further discussion in six months regarding how these points have been addressed.

The approval of the Committee is for a period of 5 years from the date of this letter. Any change to the terms of your application, including changes in data user(s), additional data fields or extension of the time period approved must be requested through Susan Kerr, PAC Administrator on 0131 275 6445 or nss.pac@nhs.net.

Please note that the access to data facilitated by this approval is subject to the satisfactory completion of approved information governance training, which must be updated every 3 years.

Appendix1

Please note that the following details about your application will be published under the following headings on the PAC website at http://www.nhsnss.org/pages/corporate/pac_meetings_and_decision_making.php later this year:

No	Title	Type	Summary	Date sent to PAC	PAC Responses	NSS Decision	Date Completed

In order to progress your request please contact the eDRIS team on telephone 0131 275 7333 or email nss.eDRIS@nhs.net.

Yours sincerely

Dr Janet Murray
Consultant in Public Health Medicine

cc eDRIS