UNIVERSITY OF LEEDS

Incidence and burden of allergic conditions and the effects of birthweight and growth on wheezing disorders in the Born in Bradford cohort

Teumzghi Fisseha Mebrahtu

BSc, MIHMEP, MSc

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

School of Medicine

December, 2015

The candidate confirms that the work submitted is his/her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Acknowledgements

I am grateful to my supervisors, Roger Parslow and Richard Feltbower for providing patient and invaluable support in the production of the thesis.

I thank the co-authors of the four published papers: my supervisors Roger Parslow and Richard Feltbower for their valuable and constructive advice and comments; Emily Petherick for her comments; and Darren Greenwood for statistical advice on meta-analysis and valuable comments in relation to the manuscripts.

Data collection for the Born in Bradford (BiB) cohort was funded by the National Institute for Health Research (NIHR) under the Collaborations for Leadership in Applied Health Research and Care (CLAHRC) programme for Leeds, York and Bradford and the Programme Grants for Applied Research funding scheme. I am grateful to all the participants, health professionals and researchers who made it possible for BiB to operate and for me to present analyses of these data in this thesis.

I also thank the Hall Dorman family for funding my PhD research scholarship.

Publications arising from this thesis

Mebrahtu TF; Feltbower RG; Parslow RC. Incidence and burden of wheezing disorders, rhinitis and eczema in pre-school children: findings from the Born in Bradford cohort. 2016; **Submitted**.

Contributors' statement: Teumzghi Mebrahtu, Richard Feltbower and Roger Parslow conceived the idea. Teumzghi Mebrahtu designed the statistical approach and performed the analyses, and produced the draft and finalised the manuscript. Richard Feltbower and Roger Parslow revised and commented on the manuscript. All authors approved the final version of the manuscript.

Mebrahtu TF; Feltbower RG; Parslow RC. The effects of birthweight and growth on wheezing disorders: findings from the Born in Bradford cohort. *BMJ Open*. 2015; 5 (11) e009553; Doi: 10.1136/bmjopen-2015-009553.

Contributors' statement: Teumzghi Mebrahtu, Richard Feltbower and Roger Parslow conceived the idea. Teumzghi Mebrahtu designed the statistical approach and performed the analyses, and produced the draft and finalised the manuscript. Richard Feltbower and Roger Parslow revised and commented on the manuscript. All authors approved the final version of the manuscript.

Mebrahtu TF; Feltbower RG; Greenwood DC; Parslow RC. Childhood body mass index and wheezing disorders: A systematic review and meta-analysis. *Pediatric Allergy and Immunology*. 2015; 26 (1):62-72. Doi: 10.1111/pai.12321

Contributors' statement: Teumzghi Mebrahtu, Richard Feltbower and Roger Parslow conceived the idea. Teumzghi Mebrahtu performed the literature Search. Teumzghi Mebrahtu, Richard Feltbower and Roger Parslow contributed to study design, literature search, and data extraction and study selection. Teumzghi Mebrahtu performed all the statistical analyses, interpretation of results and drafted the manuscript. Darren

Greenwood provided advice on statistical analyses and interpretation of results. Richard Feltbower, Roger Parslow and Darren Greenwood revised and commented on the manuscript. All authors approved the final version of the manuscript.

Mebrahtu TF; Feltbower RG; Petherick ES; Parslow RC. Growth patterns of white British and Pakistani children in the born in Bradford cohort: A latent growth modelling approach. *Journal of Epidemiology and Community Health*. 2015; 69 (4):368-373. Doi: 10.1136/jech-2014-204571

Contributors' statement: Teumzghi Mebrahtu, Richard Feltbower and Roger Parslow conceived the idea. Teumzghi Mebrahtu designed the statistical approach and performed the analyses, and produced the draft and finalised the manuscript. Richard Feltbower and Roger Parslow revised and commented on the manuscript. Emily Petherick provided the data and advice on its interpretation, and commented on the manuscript. All authors approved the final version of the manuscript.

Mebrahtu TF; Feltbower RG; Greenwood DC; Parslow RC. Birth weight and childhood wheezing disorders: a systematic review and meta-analysis. *Journal of Epidemiology and Community Health*. 2015; 69:500-508. Doi: 10.1136/jech-2014-204783

Contributors' statement: Teumzghi Mebrahtu, Richard Feltbower and Roger Parslow conceived the idea. Teumzghi Mebrahtu performed the literature Search. Teumzghi Mebrahtu, Richard Feltbower and Roger Parslow contributed to study design, literature search, and data extraction and study selection. Teumzghi Mebrahtu performed all the statistical analyses, interpretation of results and drafted the manuscript. Darren Greenwood provided advice on statistical analyses and interpretation of results. Richard Feltbower, Roger Parslow and Darren Greenwood revised and commented on the manuscript. All authors approved the final version of the manuscript.

Abstract

Past epidemiologic studies have claimed that birthweight, body mass index, and childhood growth are associated with childhood wheezing disorders although the findings are inconsistent. The aim of this thesis was to investigate the effects of birthweight body mass index and childhood growth on wheezing disorders through meta-analyses of past epidemiologic studies and using contemporary cohort data.

An online search of published papers linking childhood wheezing disorders with birthweight, BMI, and growth was carried out using EMBASE and Medline medical research databases. Risk estimates were pooled using a *random-effects* method. Data from 13,734 Born in Bradford (BiB) cohort children were used to investigate the incidence and burden of allergic diseases, and the effects of birthweight on wheezing disorders. Data of 1,598 BiB1000 children were used to investigate the effects of weight at the age of 3 years and childhood growth on wheezing disorders. Birthweight was categorised using the World Health Organisation and Centre for Disease Prevention and Control guidelines. Weight Standardised Scores were derived using World Health Organisation growth standards. Body mass index was categorised based on Centre for Disease Prevention and Control guideline.

Based on a total of 77 studies that comprised more than 3 million children, the summary risk estimates indicated that low birthweight children have an increased risk of wheezing disorders when compared with the normal birthweight children. In addition, underweight children have a reduced risk of wheezing disorders whilst overweight and obese children have an increased risk when compared with normal body mass index children.

Based on the cohort data, the results indicate that the burden of allergic conditions is higher than previously reported in earlier studies. In addition, there is an increased risk of wheezing disorders for low birthweight, slow growth during the first three months, and fast growth between 3 and 12 months.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Description
ABIC	Sample size adjusted Bayesian Information Criterion
AIC	Akaike Information Criterion
ALSPAC	Avon Longitudinal Study of Parents and Children Cohort
BiB	Born in Bradford
BiB1000	A sub-cohort of Born in Bradford
BIC	Bayesian Information Criterion
BLRT	Bootstrapped Likelihood Ratio Test
CDC	Centres for Diseases Prevention and Control
CFI	Comparative Fit Index
CI	Confidence Interval
DAGs	Direct Acyclic Graphs
df	Degrees of freedom
FIML	Full Information Maximum Likelihood
GMM	Growth Mixture Models
IOTF	International Obesity Task Force
LGCM	Latent Growth Curve Model
LL	Log-likelihood
LMR LRT	Lo-Mendell-Rubin Likelihood Ratio Test
MAAS	Manchester Asthma and Allergy Study
MAR	Missing data at Random
MCAR	Missing data Completely at Random
MICE	Multiple Imputations by Chained Equations
MI	Multiple Imputation
ML	Maximum Likelihood
OR	Odds Ratio
RMSEA	Root Mean Square Error Approximation
RR	Relative Risk
SDS	Standardised Scores
SES	Socioeconomic Status
SRMR	Standardised Root Mean Square Residuals

TLI	Tucker-Lewis Index
WHO	World Health Organisation

List of Equations

Equation 3.7 $$\Psi=\begin{bmatrix} \sigma^2_{0i} & & \\ \sigma^2_{01i} & \sigma^2_{1i} & \\ \sigma^2_{02i} & \sigma^2_{12} & \sigma^2_{2i} \end{bmatrix}$$ ……..…..…..……119

Equation 3.8 $y_{ti} = \eta_{0i} + \eta_{1i}\lambda_{1ti} + \eta_{2i}\lambda_{2ti} + \varepsilon_{ti}$ ……..……..…120

Equation 3.9 $\eta_{1i} = \alpha_1 + \zeta_{1i}$ ……..……..……120

Equation 3.10 $\eta_{2i} = \alpha_2 + \zeta_{2i}$ ……..……..……120

Equation 3.11 $y^k_{ti} = \eta^k_{0i} + \eta^k_{1i}\lambda^k_{ti} + \varepsilon^k_{ti}$ ….……...211

Equation 3.12 $$\Psi^k=\begin{bmatrix} \sigma^2_{0i}{}^k & \\ \sigma^2_{01i}{}^k & \sigma^2_{1i}{}^k \end{bmatrix}$$ ……121

CHAPTER 1

INTRODUCTION

1.1 What are allergic diseases?

Allergic diseases are inflammatory reaction of body cells mainly caused by abnormal immune system response to harmless environmental antigens or allergens (Warner and Warner, 2002; Galli et al., 2008; Grammatikos, 2008).When an allergen or foreign substance enters or makes direct contact with the body, the immune system reacts by mobilizing immunoglobulin E (IgE) and other T-cell populations as a process of normal defence mechanism (Kabesch and Von Mutius, 2002; Galli et al., 2008). In most people, this is a normal process by which the body protects against some parasitic infections, however, in those who are susceptible, the body overreacts or becomes hypersensitive, including to harmless substances, which then leads to allergic reactions (Grammatikos, 2008). This immediate reaction results in an acute inflammation around the area of insult that subsides in a short period of time if the exposure is transient. However, if the exposure persists, it can develop into chronic inflammation, that is, chronic allergic disease (Galli et al., 2008).

The primary cause of allergic diseases is unknown; however, genetic and environmental factors may play a key role (Grammatikos, 2008). The common substances that provoke allergic reactions are pollen, dust mites, foods such as nuts and peanuts, drugs and some organic compounds from plants (Galli et al., 2008). Some of the signs and symptoms of allergic reactions are wheezing, cough and shortness of breath, runny nose, itchy and watery eyes, skin rashes or eczema, vomiting and diarrhoea (Stanley, 1952; Galli et al., 2008).

In children, the common allergic disorders are asthma, rhinitis and eczema (Asher et al., 1995; Asher et al., 2006). Asthma is defined as a chronic disease of the passage of airways, characterised by smooth muscle contraction, accumulation of mucous and debris in the lumen, vascular congestion and airway wall oedema which leads to breathlessness and wheezing (Roche and Jeffery, 2002). Rhinitis is an inflammation of the inside part of the nose characterised by a runny nose, stiffness and sneezing in the absence of cold or flu (Asher et al., 1995). Eczema can be defined as a dry skin

condition characterised by itchy and red skin accompanied by rashes mostly in the flexural areas in the absence of external causes (Asher et al., 1995).

1.2 Prevalence and burden of childhood allergic diseases

Childhood allergic diseases are global health problem and their prevalence was observed to rise in the last decades (Masoli et al., 2004; Asher et al., 2006; Pearce et al., 2007). Based on self-reported symptoms, the International Study of Asthma and Allergies in Childhood (ISAAC) reported that the global prevalence of asthma, rhinitis and eczema for 6-7 age group during 2002-2003 was 12.6% (range =2.8–37.6%), 8.5% (range = 2.2–24.2%) and 8.9% (range =2–22.3%) , respectively (Asher et al., 2006). The respective prevalence for the 13-14 age groups was 14.1% (range = 3.4–31.2%), 14.8% (range = 4.5–45.1%) and 8.1% (range = 1.4–21.8%). The UK was among the highest affected countries in the world with 12-months period prevalence of 20.9%, 10.1% and 16% for asthma, rhinitis and eczema symptoms in the 6-7 age group. The respective prevalence for the 13-14 age group was 24.7% 15.3% and 14.7% (Asher et al., 2006).

In a recent retrospective cohort study of 43, 473 children using the national General Practice Research Database (GPRD), 18-years period prevalence of asthma, eczema and rhinitis in the UK was 22.9%, 36.5% and 11.4%, respectively (Punekar and Sheikh, 2009). The same study also estimated that, in 2008, there were 3.7 million, 2.2 million and 0.8 million under 18 years of age children diagnosed with eczema, asthma and rhinitis, respectively.

In the UK, it is estimated that 1 in 5 and 1 in 11 children suffer from eczema (Eczema-UK, 2015) and asthma (Asthma-UK, 2014), respectively. Although figures for eczema and rhinitis are not available, there were 1.1 million childhood asthma cases, and around 25,000 emergency hospital admissions in 2012 (Asthma-UK, 2014). It is also estimated that the National Health Service (NHS) spends £1 billion a year treating and caring for childhood and adult asthma cases (Asthma-UK, 2014).

Confirmation of rhinitis and eczema cases in young and under-five children may not be difficult, however, diagnosis of asthma is problematic because ‘wheezing’ which is the key symptom of asthma can also occur due to other causes such as, developmental anomalies (e.g. polyps), recurrent aspiration (e.g. gastroesophageal reflex), perinatal disorders (e.g. congenital infection), genetic disorders (e.g. cystic

fibrosis) and viral infections (Mckenzie, 2002; Silverman, 2002). In addition, although there are various asthma confirmatory tests available (Bush and Fleming, 2015), young children can be less cooperative in participating in such tests that may lead to an under-diagnosis of true asthma cases. Therefore, the word ‘asthma’ may not be an adequate term for what can be described as a spectrum of respiratory problems.

1.3 Childhood wheezing disorders

Childhood wheezing disorders comprise a variety of respiratory problems that share a common symptom—*wheezing* (Silverman, 2002). Wheezing disorders can be categorised into a variety of phenotypes based on the temporal pattern and duration of wheeze (Brand et al., 2008). Based on temporal patterns, wheezing disorders can be: a) episodic, often related with viral cold or b) multiple-trigger where wheezing occurs between episodes and shows exacerbations of symptoms during episodes due to triggers other than cold (e.g. tobacco smoke). Based on the duration, wheezing disorders can be: a) transient where wheezing symptoms occur before the age of three years and disappear by the age of six; b) persistent where wheezing symptoms start before the age of three years and continue to manifest until the age six and afterwards; or c) late-onset where wheezing symptoms start after the age of three years (Brand et al., 2008).

1.3.1 Contributing factors for childhood wheezing disorders

Although childhood wheezing disorders are diverse, the biological contributing factors can be broadly categorised into three: *anatomical* and *physiological development*, *inflammatory response*, and *smooth muscle remodelling*.

1.3.1.1 Anatomical and physiological development

Anatomical development of the lungs starts from the *embryonic phase* (0-7 weeks gestation) and continues until *postnatal phase* (up to 18 months after birth) (Hislop and Pandya, 2002). Each anatomical component of the lung has its own timetable of development although they are not independent to each other (Hislop and Pandya, 2002). At the time of birth, the lungs need to have enough number of alveoli in order to carry out air breathing and gas exchange functions which may not be the case for babies with intrauterine growth restriction and born premature (Willet and Sly, 2002).

Maternal smoking during pregnancy is believed to be associated with low birthweight (Jaddoe et al., 2008) due to constriction of utero-placental circulation induced by nicotine (Willet and Sly, 2002). Lower birthweight babies are likely to have smaller lungs and lower number of alveoli, which in turn could cause breathing difficulty and wheezing symptoms (Hislop and Pandya, 2002). Babies with less structurally developed lungs would also be more likely to have respiratory infections that can lead to wheezing symptoms (Hislop and Pandya, 2002).

1.3.1.2 Inflammatory response

Inflammation is a protective reaction of any vascularised tissue of the body against foreign substances (Grigg, 2002). In response to inhaled antigens, a cascade of inflammatory process takes place (Renauld, 2001). First, *CD4+* helper T cells produce interleukins (*IL-4*, *IL-5*, *IL-9*, and *IL-13*). Second, naive T-cells differentiate into Th2 cells and stimulate B-cells to produce *IgE* antibodies in response to *IL-4*. Fourth, the *IgE*s bind with mast cells. Finally, the mast cells release histamines and inflammatory cells (prostaglandins and leukotrienes) that cause bronchospasm and inflammation of the airways, respectively. At the same time, inflammatory cells such eosinophils are pulled towards the area of insult by the chemotactic factors released by mast cells (Renauld, 2001).

During the inflammatory reaction, airway tissue damage occurs mainly due to the presence of excess cytokines in the lung tissues (Holt, 2002). As a consequence to the tissue damage and inflammatory process, the airways become congested and narrow which then can lead to transient or persistent wheezing symptoms, based on the duration of the exposure (Balfour-Lynn and Openshaw, 2002).

1.3.1.3 Smooth muscle remodelling

The function of the airways is to regulate the flow of gas exchange and prevent harmful substances from reaching the air sacks (Renauld, 2001; Hislop and Pandya, 2002; Roche and Jeffery, 2002). However, this protective mechanism of the lung tissues may also cause damage to the tissues themselves.

As a healing process, damaged (injured) epithelial tissues are infiltrated by neutrophils and later by lymphocyte and macrophages in order to eradicate the dead cells before the process of repairing of tissues takes places (Roche and Jeffery,

2002). The healing process culminates by *remodelling* of the tissues , that is, formation of collagen and scar, and contraction of the surrounding tissues (Roche and Jeffery, 2002). *Inflammation* and *remodelling* are part of the healing process, in which the structure and function of the tissues is restored in the end. In wheezing disorders cases, however, both the inflammation and remodelling processes persist (Renauld, 2001; Roche and Jeffery, 2002; Galli et al., 2008).

The aftermath of the incessant inflammation and remodelling process is that the airways become thickened (Renauld, 2001), and much of the elasticity of the smooth muscles is reduced (Renauld, 2001; Roche and Jeffery, 2002) which then can lead to wheezing symptoms.

1.3.2 Hypotheses for the association of birthweight, body mass index and growth patterns with childhood wheezing disorders

1.3.2.1 Birthweight

Results from previous meta-analyses suggest that low birthweight is associated with wheezing disorders (Mu et al., 2014; Xu et al., 2014) although the mechanism of this relationship is less understood. However, it is reported that low birthweight children have increased risk of lower respiratory infections (Jackson et al., 2013; Lu et al., 2013). It was also suggested that low birthweight children could be susceptible to infectious diseases due to alteration in their immune function (Raqib et al., 2007).

Thus, it is possible that low birthweight children experience more viral respiratory infections than normal birthweight children due to either an immune function alteration or having lungs that can not carry out their function of air exchange and protection against harmful substance properly. Viral respiratory infections especially bronchiolitis from respiratory syncytial virus (RSV) and rhinovirus (RV) can develop into recurrent wheezing or childhood wheezing disorders (Singh et al., 2007; Sly et al., 2008). Recurrent viral infections can also cause damage to the respiratory airway tissues that can lead to stiffness of mucosal muscles and lack of elasticity, which then cause wheezing symptoms (Balfour-Lynn, 1996).

1.3.2.2 Body mass index

There are three potential mechanisms that may explain the association between body mass index (BMI) and wheezing disorders. First, overweight or obesity, due to the

presence of high visceral fat, can mechanically increase the pressure on the diaphragm and the airways which can lead to breathing difficulty and wheezing symptoms (Sontag, 2000). Second, obesity may cause gastro-oesophageal reflux and that the gastric content of the reflux may cause infection and inflammation of the airway that may then lead to wheezing symptoms (Sontag, 2000).

Third, overweight or obese people have excess pro-inflammatory hormones (e.g. adipokines) in their blood circulation (Guler et al., 2004; Castro-Rodríguez, 2007; Farah and Salome, 2012). This suggests that overweight or obesity is an inflammatory state where the cascading of inflammatory process can trigger asthmatic symptoms (Guler et al., 2004; Castro-Rodríguez, 2007; Farah and Salome, 2012).

1.3.2.3 Childhood growth patterns

The effect of growth patterns on wheezing disorders could perhaps be explained by effects of birthweight and BMI in combination. Growth starts from prenatal period and those who had adverse events during this period will most likely be low birthweight babies which then become susceptible to respiratory infections (Jackson et al., 2013; Lu et al., 2013). Postnatally, babies also may falter from normal growth to become obese which then leads to wheezing symptoms (Guler et al., 2004; Castro-Rodríguez, 2007; Farah and Salome, 2012).

1.3.3 The risk factors for wheezing disorders

The primary cause of wheezing disorders is not known, however, a combination of genetic predisposition with exposure to allergens are described to be the main risk factors (Grammatikos, 2008; WHO, 2013). Previous observational epidemiologic studies have also reported that factors such as low birthweight (Mu et al., 2014; Xu et al., 2014), overweight and obesity (Chen et al., 2013; Egan et al., 2013), fast growth during early age (Rzehak et al., 2013), family asthma (Lim et al., 2010), maternal smoking (Burke et al., 2012), breast feeding (Scholtens et al., 2009), child and family feeding habits (Chatzi et al., 2007), number of siblings (parity) and live births (McKeever et al., 2001), socioeconomic status (Kozyrskyj et al., 2010), gestational age (Jaakkola et al., 2006), gender (Osman et al., 2007) and ethnicity (Netuveli et al., 2005; Akinbami et al., 2014) are linked with the risk of wheezing disorders.

1.4 The Born in Bradford cohort project

The Born in Bradford (BiB) cohort study is based at the Bradford Institute for Health Research (BIHR) office, Braford Royal Infirmary, in Bradford. The study covers the Bradford district in West Yorkshire North of England with a total area of around 143 square miles. The district includes 30 electoral Wards with a population of around half a million residents mainly White British and Pakistani (mostly from Mirpuri region of Pakistan) (Raynor, 2008).

The district's population is younger than the national average (Raynor, 2008); and, while only around 18% of the residents are Pakistani origin (Wright et al., 2013), around 44% babies born in the district are from Pakistani origin mothers (Raynor, 2008).

Bradford district is among the most deprived cities in the United Kingdom and is known for its high childhood morbidity and twice the national average infant mortality (BDIMC, 2007; BDIMC, 2014). Babies born to Pakistani origin mothers are twice as likely as to die during their first year when compared with babies born from white mothers (BDIMC, 2007).

The BiB project was established in 2007 in response to the concern about the high childhood morbidity and infant mortality in the district (Wright et al., 2013). The aim of the project was to examine the impact of genetic, nutritional, environmental, behavioural and social factors on child health and development, and adult life (Wright et al., 2013).

1.5 Motivation for this thesis

1.5.1 Contemporary birth cohort data

In the UK, there are birth cohorts such as: Avon Longitudinal Study of Parents and Children (ALSPAC) cohort (Boyd et al., 2012), the Manchester Asthma and Allergy Study (MAAS) cohort, (Custovic et al., 2002), and the Millennium Cohort study (Dex and Joshi, 2005). However, in these cohorts, minority ethnic groups are underrepresented and much of the diseases outcome data are collected through questionnaires. Therefore, comparative statistical analyses among different ethnic groups could be under powered. In addition, data collected through questionnaires are also prone to recall and measurement error biases.

On the other hand, the BiB cohort is made up of almost homogenous bi-ethnic population (i.e. Pakistani and white British origin), and all child anthropometric measurements are collected by trained workers. In addition, the project also provides a unique opportunity for utilising primary care data in SystmOne (http://www.tpp-uk.com/products/systmone), that is, electronic patient record (i.e. disease and drug prescription information). Therefore, any comparative analyses among the two ethnicities using the BiB cohort data will provide strong statistical power. Results based on the cohort data will also be less prone to recall and measurement error biases.

1.5.2 The requirement for update

The effects of birthweight on childhood wheezing disorders have been extensively studied although results remained inconsistent (Chatkin and Menezes, 2005). Until March 2015, 83 studies that investigated birth weight and childhood wheezing disorders were published. However, only 20% of them contributed to the last two systematic reviews and meta-analyses (Flaherman and Rutherford, 2006; Mu et al., 2014).

Previous observational epidemiologic studies also suggested that overweight/obesity and childhood asthma are associated although results remain inconsistent. Meta-analyses were carried out in the past, however, one of the meta-analyses is too old (Flaherman and Rutherford, 2006), and the other two recent meta-analyses focused only on cohort studies, though not all previous cohort studies were included (Chen et al., 2013; Egan et al., 2013).

Therefore, there is a need to carry out an up-to-date investigation of the association of birthweight and BMI with childhood wheezing disorders through a systematic review and meta-analysis of past epidemiologic studies.

1.5.3 The need to use novel statistical analytic techniques

Even though a large volume of research has been carried out on childhood wheezing disorders and modifiable risk factors, the accuracy and precision of the findings may be questionable. This is because, rigorous thought and proper planning prior to data analysis was lacking in the past studies. For instance, critical thinking of what variables to include and exclude was rarely in evidence. Most publications consider all variables entered into the model as potential confounders although the term confounding strictly refers to a variable causally related with but not directly

affected by the outcome and exposure variables (McNamee, 2003). The consequence of including a variable which is directly affected by an exposure is shown to exhibit a reversal effect, leading to incorrect effect estimates and erroneous inferences (Tu et al., 2005). In order to minimise biases due to such problems, one has to use novel tools, Direct Acyclic Graphs (DAGs), in order to properly identify confounding and confounders (Greenland et al., 1999; Shrier and Platt, 2008).

Childhood weight change and growth patterns have been reported as predictors of health during childhood and adult life (Eriksson et al., 2003; Hardy et al., 2004; Baker et al., 2007; Eriksson et al., 2007; Owen et al., 2009; Halldorsson et al., 2011). However, much of the evidence is based on multiple regression models, which are prone to collinearity problems caused by the repeated weight measurements that can lead to biased coefficient estimates (Duncan and Duncan, 2004; Tu et al., 2013). To avoid such collinearity problem, generalized estimating equations (Ballinger, 2004; Hwang and Takane, 2005), multilevel linear models (Bryk and Raudenbush, 1987) or latent growth models (Muthen, 2001; Duncan and Duncan, 2004; Muthén, 2004) are recommended.

1.6 Aims of this thesis

This thesis mainly focuses on the investigation of the effects of birthweight, BMI, and childhood growth through systematic review and meta-analysis of past observational epidemiologic studies and using BiB cohort data. In so doing, the thesis has the following aims:

- To conduct an up-to-date systematic review and meta-analysis of studies on the effects of birthweight, BMI, and growth on childhood wheezing disorders;
- To investigate the incidence and burden of childhood wheezing disorders, eczema and rhinitis in the BiB cohort population; and,
- To investigate the effects of birthweight, weight at the age of 3 years and childhood growth on childhood wheezing disorders in the BIB cohort.

1.7 Structure of this thesis

The thesis is structured in four parts. Chapter 1 forms the first which provided a summary background for the thesis.

Chapter 2 forms part 2, presents the summary of past epidemiologic studies that investigated the effects of birthweight, BMI and childhood growth patterns on wheezing disorders through systematic literature reviews and meta-analyses. Methods for literature search, data extraction and standardisation are described. Finally, summary results for each of the anthropometric measures (i.e. birthweight, BMI and growth patterns) are presented, followed by discussions of the results and concluding remarks.

Part three consists of Chapters 3, 4 and 5. Chapter 3 discusses methodological issues in research and details the materials and methodologies used in a series of analyses using the BiB cohort data, namely: the incidence and burden of allergic conditions; describing growth patterns of white British and Pakistani children; and investigating the effects of birthweight, weight at the age of 3 years and childhood growth patterns on wheezing disorders. In Chapter 4, results for each analyses carried out in chapter 3 are presented and described in detail; and Chapter 5 includes discussions of the results in Chapter 4 together with conclusions.

Final part consists of Chapter 6 and presents a summary of the key findings from the past (i.e. systematic review and meta-analysis) and current (i.e. results from the BiB cohort data) epidemiological research and highlights areas for further investigation.

CHAPTER 2

SYSTEMATIC REVIEW AND META-ANALYSIS

2.1 Chapter overview

This chapter investigated the effects of birthweight, BMI and childhood growth patterns on childhood wheezing disorders using past observational epidemiologic studies' data. In section 2.2, a brief introductory discussion of systematic review and meta-analysis methods (i.e. *fixed-effect* and *random-effects*) is presented.

In section 2.3; a critique of past systematic reviews and meta-analyses of birthweight and childhood wheezing disorders is presented. Then, methods used and results are presented. The section then ends with critical discussion of the results and concluding remarks.

Likewise, section 2.4 presents a critique of past systematic reviews and meta-analyses of BMI and childhood wheezing disorders, and a detailed description of methods and results, and a critical discussion of the findings. Finally, in section 2.5, a descriptive summary and discussion of childhood growth patterns and childhood wheezing disorders studies is presented.

2.2 Overview of systematic review and meta-analysis

Systematic literature review is a process that includes systematically locating, appraising, selecting, and synthesising data from past literature with an aim of creating generalisations or answering a research question (Hedges, 2009a). Meta-analysis is a process of statistically synthesising or combining results from past studies (Hedges, 2009a). There are two types of modelling approaches in meta-analysis, that is, *fixed-effect* or *random-effects* (Borenstein et al., 2009a; Hedges, 2009b).

Under *fixed-effect* modelling, the analyst only wants to make inferences in the effect-size parameter of the included studies (Hedges, 2009b). The assumption is that there is one true effect which underlines all the studies in the analysis and that all differences in the observed effects are due to sampling error within each study (Borenstein et al., 2009a). Therefore, the relationship between the observed effect and unknown true effect can be written as:

$$Y_i = \theta + \varepsilon_i \qquad (2.1)$$

where Y_i and ε_i are the individual observed effect and the sampling error for study (i), respectively, and theta (θ) is the unknown true effect or population mean (Borenstein et al., 2009a). In order to obtain the most precise estimation of the population effect size, a weighted mean is calculated where the weight assigned to each of the studies is the inverse of that study's variance (Borenstein et al., 2009a; Shadish and Haddock, 2009). So, the weight given to each study in a *fixed-effect* meta-analysis (W_{fi}) is:

$$W_{f_i} = \frac{1}{V_{Y_i}} \qquad (2.2)$$

where V_{Y_i} is the within-study variance, that is, the sampling error for study (i). Therefore the weighted mean (or the summary risk estimate) can be calculated as:

$$M = \frac{\sum_{i=1}^{k} W_{fi} Y_i}{\sum_{i=1}^{k} W_{fi}} \qquad (2.3)$$

where M is the summary risk estimate, W_{fi} is the *fixed effect* weight assigned for study (i), Y_i is the observed effect size for study (i). The variance of the summary effect is estimated as:

$$V_M = \frac{1}{\sum_{i=1}^{k} W_{fi}} \qquad (2.4)$$

Under *random-effects* modelling, the assumption is that the true effects of studies are not identical but randomly distributed (Borenstein et al., 2009a). Therefore, we would have two sources of variation, that is, the within-study (sampling error) and the between-study variation. Then, equation 2.1 is written as

$$Y_i = \mu + \zeta_i + \varepsilon_i \qquad (2.5)$$

Where Y_i is the observed effect and mu (μ) is the mean of the distribution of true effects among a population of studies; and zeta (ζ) is the difference between true study and population effect (i.e. between-study error), and epsilon (ε) is the difference between the observed study effect and the unknown true study effect (within-study error). Subsequently, equation 2.2 is modified as:

$$W_{r_i} = \frac{1}{V_{Yi} + \tau^2} \qquad (2.6)$$

Where W_{ri} and τ^2 are the random effects assigned weight for study *i* and the between-study variation, respectively. The between-study variance and is estimated as:

$$\tau^2 = \frac{Q - (k-1)}{\sum_i w_{f_i} - \left(\frac{\sum_i w_{f_i}^2}{\sum_i w_{f_i}} \right)} \qquad (2.7)$$

where W_{fi} is weight from the fixed-effect inverse variances, Q is the *heterogeneity* test statistic, *k* is the number of studies and *k*-1 is the degrees of freedom (*df*). Once the weights given to each study are calculated, the summary effect and its variance are estimated as in the *fixed-effect* model (Borenstein et al., 2009a; Shadish and Haddock, 2009).

The choice between the *fixed-effect* and *random-effects* meta-analyses depends on the assumption about the studies and the generalisation that one wants to make from

the results (Borenstein et al., 2009a; Hedges, 2009b). Suppose that in scenario (a), a big interventional study is planned with 5000 participants. However, due to personnel and logistics, it is decided that the study be split into 10 studies. It is also decided that the results of each study are to be aggregated for a summary effect size. In this circumstance, it can be assumed that all the studies included are functionally identical, so a *fixed-effect* meta-analysis is the plausible method (Borenstein et al., 2009a).

In scenario (b), suppose that there are 10 studies to be conducted in different countries or continents by different researchers and there is a plan to aggregate the results for a summary effect size. It is very unlikely that the studies will be functionally identical so one cannot calculate a common effect size. Hence, in this scenario, a *random-effects* meta-analysis becomes more sensible choice (Borenstein et al., 2009a). In scenario (b), based on the summary effect size, one can also make generalisations about other populations, but not in scenario (a).

The ultimate aim of performing a meta-analysis is not only to calculate summary effect size, but also, to assess the variation of the effect sizes of studies (*heterogeneity),* and identify its sources (Borenstein et al., 2009c). There are two sources of variation or *heterogeneity* of effect sizes, that is, true variation and random error. The task is to isolate the true heterogeneity and quantify it. There are three tools to identify heterogeneity: (a) the Q statistic that measures the weighted squared deviations; (b) the tau-squared (τ^2) that measures the between study variance and (c) the I^2 which is the ratio of true heterogeneity to the total observed variation. Once the presence of true heterogeneity is confirmed, its sources are investigated through (a) a subgroup analysis where the mean effect size for different subgroups are compared (Borenstein et al., 2009e) or (b) meta-regression where the relationship between study level covariates and the effect size are assessed (Lau et al., 1998; Borenstein et al., 2009d).

2.3 Birthweight and wheezing disorders

2.3.1 Critique of past systematic reviews and meta-analyses

The effects of birthweight on childhood wheezing disorders have been extensively studied although results remained inconsistent (Chatkin and Menezes, 2005). Syntheses of studies have been carried out in the past (Flaherman and Rutherford, 2006; Mu et al., 2014; Xu et al., 2014), however, the results were inconsistent and the methodologies applied by the authors were less rigorous. For example, in a meta-analysis of 9 observational epidemiologic studies, it was reported that there was an increase of 20% (RR=1.2, 95% CI : 1.1 to 1.3) in childhood asthma risk for high birthweight children (Flaherman and Rutherford, 2006). However, the studies included in this meta-analysis of high birthweight and childhood asthma used a variety of definitions for high birthweight and risk estimations. One of the studies used 3.8kg (Schwartz et al., 1990), four used 4.0kg (Fergusson et al., 1997; Gregory et al., 1999; Leadbitter et al., 1999; Bolte et al., 2004) and another used 4.5kg (Sin et al., 2004) as cut-off points for high birthweight, whilst three others used different birthweight measurements (Rasanen et al., 2000; Xu et al., 2002; Yuan et al., 2003); four used relative risk (Schwartz et al., 1990; Fergusson et al., 1997; Yuan et al., 2002; Sin et al., 2004) and five used odds ratios (Gregory et al., 1999; Leadbitter et al., 1999; Rasanen et al., 2000; Xu et al., 2002; Bolte et al., 2004) which could potentially affect the summary risk estimates.

From a meta-analysis of nine studies, Mu et al. (2014) also reported that low birthweight increases the risk of asthma by 28% (OR=1.28 , 95% CI: 1.09 to 1.50) and 34% (OR=1.34, 95% CI:1.13 to1.60) for studies that used two and three birthweight categories respectively. However, the population's age and birthweight categorisation were not consistent across the studies included. For example, one of the studies used data-driven quartile birthweight categories (Taveras et al., 2006), another had a mixture of child and adult populations (Gregory et al., 1999), and three others were treated as adult studies (Steffensen et al., 2000; Yuan et al., 2002; Remes et al., 2008) although the participants were children. And also, one other included study (Rona et al., 1993) used 'asthma attack' as an outcome measure for asthma while this may underestimate the true number of cases as many asthmatics may not experience any 'attack' at all.

2.3.2 Sythematic literature review methods

2.3.2.1 Literature Search Strategy

The reviews were carried out in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009) whilst a protocol was also registered with PROSPERO (Mebrahtu et al., 2014). Online search was carried out using the EMBASE and MEDLINE databases. Prior to the searches, a list of terms and phrases was constructed. Table 2.1 gives the details of terms and phrases used for the literature search.

2.3.2.2 Inclusion criteria

Eligible papers were those published as an article, in English, and reported original research about the effects of birthweight on wheezing disorders in children 0-19 years of age. No lower limit for a time of publication was set. However, the literature covered the period to March 2015. Case-control, cohort, and cross-sectional studies were included.

2.3.2.3 Exclusion criteria

Papers were excluded if:

a) Birthweight was modelled as a continuous variable as an assumption was made that the risk of outcome (wheezing disorder) is higher in the lower and higher ends of weight bands. Thus, 'standard' categorical variable of birthweight was considered to be more appropriate than its continuous format;
b) Authors claimed birthweight was included in their analyses but no comparison group or risk estimates were presented in text or table of the papers;
c) Authors used data-driven multiple categories of birthweight that cannot be converted into the 'standard' categories;
d) Studies included an adult population with no separate data available for children and adolescents; and,
e) Authors used data-driven multiple birthweight categories and if the number of categories presented were generally too few (<4) to allow combination

with studies through estimating nonlinear dose-response curves (Orsini et al., 2012).

Table 2.1 Terms and phrases used during literature search

	Birthweight
1	birthweight
2	low birthweight
3	high birthweight
4	Birth weight
5	low birth weight
6	high birth weight
7	childhood asthma
8	wheez*
9	wheezing disorders
10	asthm*
11	Asthma in children
12	Childhood wheez*
13	1-6/or
14	7-12/or
15	13-14/and
16	limit 15 to English language

N.B: First literature search was conducted on 3rd March 2014 and was updated on 8th April 2015.

2.3.2.4 Data extraction

For the studies that were eligible to be included in the meta-analyses, the following characteristics were extracted:

a) Authors' name;
b) Year of publication;
c) Country of study;
d) Study design;
e) Sample size;
f) Study age group and gender;
g) Diagnosis (outcome) terms;
h) Birthweight categories;
i) Birthweight categorisation methods;
j) Outcome and exposure ascertainment methods; and,
k) Risk estimates.

2.3.2.5 Data standardisation

Exposure variable

Authors of the included studies used four types of exposure categorisation techniques. For comparability and not to lose data due to variation in categorisation methods, standardisation was undertaken.

a) Where authors assumed the Centre for Disease Prevention and Control (CDC, 2009) and 'recent' World Health Organisation method (WHO, 2014) that categorises birthweight as low (<2.5kg), normal (2.5-4.0kg) and high (>4.0kg) or the 'old' WHO method (Kramer, 1987) that categorises low (<2.5kg) and normal (≥2.5kg), the reported adjusted risk estimates and data on the number of cases and non-cases of each weight comparison group were combined for meta-analysis without any change;
b) Where authors adopted two or three birthweight categories with Centre for Disease Prevention and Control (CDC) or World Health Organisation method (WHO) 'normal' category as a reference and where the number of participants in each category were available, the stratum based number of cases and non-cases were aggregated before being combined with the other studies for meta-analysis of unadjusted risk estimates;
c) Where authors adopted two or three birthweight categories with the CDC or WHO normal category as a reference and provided adjusted risk estimates, the stratum based risk estimates were aggregated using recommendations from Hamling et al. (2008) before being combined with the other studies for meta-analysis of adjusted risk estimates, and
d) Where authors adopted data-driven multiple categories that could not be converted to either of the standard formats, the risk estimates were compiled in a table for descriptive analysis.

Outcome variable

Study authors used one or multiple outcome terms in their reporting. Again, for comparability among studies, where authors used a single outcome, for example, asthma or wheezing, the quoted outcome term by the author and its risk estimate were assumed for analysis. However, where authors used multiple outcome terms, the term that was highest in the hierarchy and its risk estimate were assumed for

analysis. For example, if asthma and wheezing were used together, asthma was preferred over wheezing.

2.3.2.6 Quality assessment

Papers included in the review and meta-analysis were assessed for risks of bias using the Newcastle-Ottawa quality assessment scale (Wells et al., 2000), see Appendix A , Appendix B and Appendix C for details on scoring guidelines used. Table 2.3 gives the details of studies with respective scores.

2.3.2.7 Statistical analysis

Where meta-analyses were carried out, random-effects models were preferred as an assumption was made that the studies were not functionally identical and the aim of the meta-analyses were to generalise about other populations in different parts of the world (Borenstein et al., 2009b). Estimates were pooled using the DerSimonian and Laird method (DerSimonian and Laird, 1986).

If studies presented stratum-specific estimates (e.g. by gender), then to provide correct measures of heterogeneity, the risk estimates were aggregated using fixed-effect models before being combined with the other studies for meta-analyses of adjusted risk estimates in a random-effects model. Likewise, where authors reported the number of cases and non-cases in each stratum, the total number of cases and non-cases were aggregated before being combined with the other studies for meta-analyses of unadjusted risk estimates of all studies.

To quantify between-study heterogeneity, the Cochrane Q-test (Whitehead and Whitehead, 1991) and the I^2 measure of the proportion of the total heterogeneity explained by between study variation (Higgins and Thompson, 2002) were used. Sub-group meta-analyses and sensitivity analysis of adjusted and unadjusted risk estimates were performed on nine covariates (study characteristics) in order to assess the robustness of the risk associations and levels of between-study heterogeneities.

Where summary risk estimate results showed significant variation (heterogeneity), in order to account for the sources of between-study heterogeneity, meta-regression (Lau et al., 1998; Borenstein et al., 2009d) of adjusted and unadjusted risk estimates were performed using Restricted Maximum Likelihood (REML).

In investigating evidence of publication bias and small study effects, symmetry funnel plots and bias test models (Egger et al., 1997; Sterne et al., 2001) were used.

All meta-analyses were carried out in Stata software version 12 (StataCorp, 2011). The *fixed-* and *random-effects* models were carried out using '`metan`' command. Likewise, meta-regression, bias test and funnel plot models were performed using '`metareg`','`metabias`', and '`metafunnel`' Stata commands, respectively. Five per cent significance levels and 95% confidence intervals were adopted throughout.

2.3.3 Results

2.3.3.1 Literature search

A total of 1,830 papers were recovered from EMBASE and Medline collectively. Of the total, 83 papers were read in full. Out of the 83 papers, 52 reported either the risk estimates or number of cases and non-cases of wheezing disorders in each exposure group and were included in the review (Figure 2.1). Then, 38 of the total 52 studies either used the standard birthweight categories or presented data that were convertible to the standard formats and were included in the meta-analysis (Table 2.2).

Eleven of the 52 studies used data-driven birthweight categories which were found to be inconvertible into the standard formats but can be re-grouped into two or three categories (Gold et al., 1999; Yuan et al., 2003; Sin et al., 2004; Mai et al., 2007; Garcia-Marcos et al., 2008; Davidson et al., 2010; Jeong et al., 2010; Brew et al., 2012; Lu et al., 2012; Mathew et al., 2012; Nuolivirta et al., 2012), see Appendix D . A further three studies were also not combined due to the use of asthma admission as outcome term (Kiechl-Kohlendorfer et al., 2007; Liu et al., 2014), and pooled data from different continents and presented a summary risk estimate (Mitchell et al., 2014).

The 38 studies were from Europe (18), Americas (12), Asia (5) and Oceania (3). Only 1 study was classified as a case-control while the 37 were cohort (i.e., retrospective and prospective) studies. The sample population of 31 studies ranged between 1,085 and 764,207 while 7 studies had <1,000 participants each.

2.3.3.2 Quality of studies

With a maximum score of 9 points available for each article, of the 38 included in the meta-analysis: 14 scored 7-9 (>75%), 18 scored 5-6 (50-75%), and 6 scored ≤4 (<50%) and their risks of biases can be interpreted as ‘low’, ‘moderate’ and ‘high’ respectively (Table 2.3). The quality of the studies can be categorised as high, medium and low for quality scores of 7-9/9, 5-6/9 and ≤4/9, respectively.

Figure 2.1 Birthweight and wheezing disorders literature search flow chart

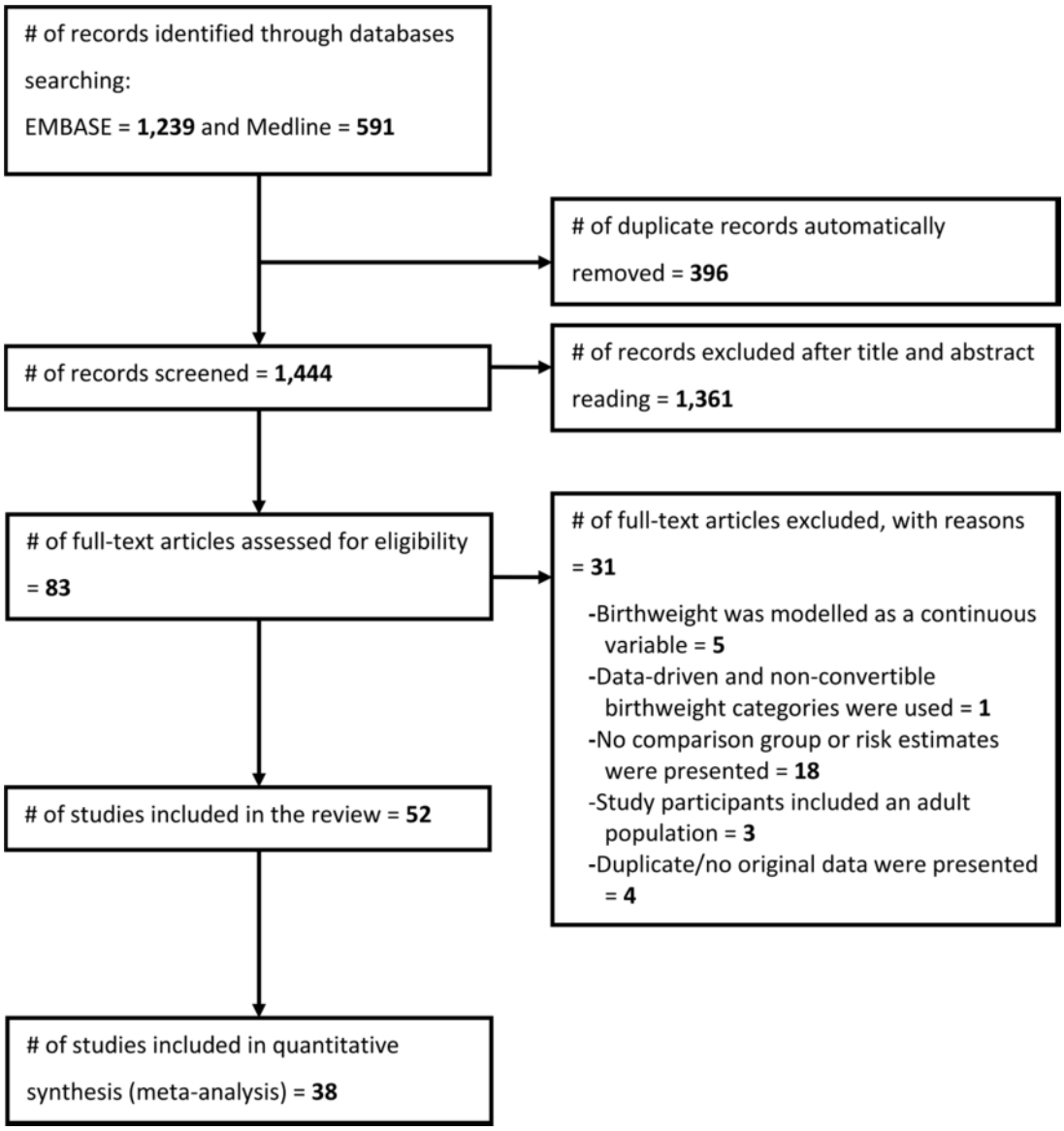


Table 2.2 Characteristics of birthweight and wheezing disorders studies included in the meta-analysis

Author , year, region	Study design	Sample size	Participants' age and gender	Outcome term	Outcome ascertainment	Exposure ascertainment	Exposure categories
Weitzman et al. (1990), USA	RC	2,927	2-5 years mixed	asthma	parent	parent	<2.5 kg and ≥2.5kg
Seidman et al. (1991), Israel*	RC	19,772	17 years boys	asthma	e-records	e-records	<2.5 kg, 2.5-4.0kg, and >4.0kg
Arshad et al. (1993), UK	PC	1,215	2 years mixed	asthma	physician	no mention	<2.5 kg and ≥2.5kg
Azizi et al. (1995), Malaysia	CC	359	1 month-5 years mixed	asthma	physician	no mention	<2.5 kg and ≥2.5kg
Lewis et al. (1995), UK*	RC	12,577	5 years mixed	wheezing	parent	e-records	<2.5 kg, 2.5-4kg, and >4.0kg
Lewis et al. (1996), UK*	RC	18,835	16 years mixed	wheezing	parent	e-records	<2.5 kg and ≥2.5kg
Schaubel et al. (1996), Canada	RC	16,207	1-4 years mixed	asthma	e-records	e-records	<2.5kg and ≥2.5kg
Sears et al. (1996), New Zealand*	PC	1,037	18 years mixed	asthma	physician	e-records	<2.5 kg, 2.5-4kg, and >4.0kg
Fergusson et al. (1997), New Zealand*	RC	888	16 years mixed	asthma	e-records	e-records	<2.5 kg, 2.5-4kg, and >4.0kg
Lilljeqvist et al. (1997), Norway*	RC	569	7-10 years mixed	asthma	parent	e-records	<2.5 kg, 2.5-4kg, and >4.0kg
Slezak et al. (1998), USA*	RC	847	3-5 years mixed	asthma	parent	no mention	≤2.5kg and >2.5kg
Wjst et al. (1998), Germany	RC	2,470	5-14 years mixed	asthma	parent	parent	<2.5 kg and ≥2.5kg
Leadbitter et al. (1999), New Zealand*	PC	735	13 years mixed	asthma	physician	e-records	<2.5 kg, 2.5-4kg, and >4.0kg
Rasanen et al. (2000) Finland*	RC	4,502	16 years mixed	asthma	parent	parent	<2.5kg and ≥2.5kg
Steffensen et al. (2000), Denmark*	PC	4,795	18 years boys	asthma	physician	e-records	<2.5 kg, 2.5-4kg, and >4.0kg
Annesi-Maesano et al. (2001), UK	RC	4065	0-18 years mixed	asthma	parent	parent	<2.5 kg and ≥2.5kg
Brooks et al. (2001), USA*	RC	8,071	3 years mixed	asthma	parent	e-records	<2.5 kg and ≥2.5kg
Ronmark et al. (2002), Sweden	RC	3,247	7-8 years mixed	asthma	parent	e-records	<2.5 kg and ≥2.5kg
Anand et al. (2003), UK	RC	256	15 years mixed	asthma	e-records	e-records	<2.5 kg and ≥2.5kg
Benicio et al. (2004), Brazil	RC	1,085	6-59 months mixed	wheezing	parent	no mention	<2.5 kg and ≥2.5kg
Bolte et al. (2004), Germany*	RC	715	5-7 years mixed	asthma	parent	e-records	<2.5 kg, 2.5-4kg, and >4.0kg
Al-Kubaisy et al. (2005), Iraq	CC	2,262	6-12 years mixed	asthma	parent	parent	<2.5 kg and ≥2.5kg

Author , year, region	Study design	Sample size	Participants' age and gender	Outcome term	Outcome ascertainment	Exposure ascertainment	Exposure categories
Bernsen et al. (2005), Netherlands*	RC	1,710	6 years mixed	asthma	e-records	e-records	<2.5kg and ≥2.5kg
Nepomnyaschy and Reichman (2006), USA	RC	1,803	3 years mixed	asthma	parent	e-records	<2.5 kg and ≥2.5kg
Remes et al. (2008), Finland*	RC	4,660	16 years mixed	asthma	parent	no mention	<2.5kg and ≥2.5kg
Ortqvist et al. (2009), Sweden*	RC	10,570	9-12 years mixed	asthma	parent	e-records	<2.5kg and ≥2.5kg
Xu et al. (2009), USA	RC	2,409	1-5 years mixed	asthma	parent	no mention	<2.5kg, 2.5–4.0kg, and >4.0 kg
Midodzi et al. (2010), Canada	PC	8,397	4-5 years mixed	asthma	physician	e-records	<2.5 kg and ≥2.5kg
Bjerg et al. (2011), Sweden	RC	2,996	11-12 years mixed	asthma	parent	no mention	<2.5 kg and ≥2.5kg
Mogensen et al. (2011), Sweden*	PC	1784	13-14 years mixed	asthma	parent	e-records	<2.5kg and ≥2.5kg
Suglia et al. (2011), USA	RC	1,815	3 years mixed	asthma	parent	parent	<2.5kg and ≥2.5kg
To et al. (2012), Canada*	RC	687,194	6 years mixed	asthma	e-records	e-records	<2.5kg and ≥2.5kg
Wang et al. (2012), Taiwan	RC	78,011	13-16 years mixed	asthma	parent	e-records	<2.5 kg and ≥2.5kg
Kallen et al. (2013), Sweden*	RC	764,207	2-11 years mixed	asthma	e-records	e-records	<2.5 kg, 2.5-4kg, and >4.0kg
Miyake and Tanaka (2013), Japan	RC	2004	3 years mixed	asthma	parent	e-records	<2.5 kg and ≥2.5kg
Yang et al. (2013), USA	RC	3,933	7 years mixed	asthma	e-records	e-records	<2.5 kg and ≥2.5kg
Granell et al. (2014), UK	PC	4,778	7 years mixed	asthma	e-records	parent	<2.5 kg and ≥2.5kg
Reis et al. (2015), Brazil	RC	1,468	1-4 years mixed	asthma	parent	parent	<2.5 kg and ≥2.5kg

PC=prospective cohort; RC=retrospective cohort; CC=case-control; * = regrouped birthweight categories; mixed=included both genders.

Table 2.3 Risk of bias assessment table using Newcastle-Ottawa quality assessment scale for birthweight and wheezing disorder studies included in the meta-analysis

Author , year, region	Study title	Selection	Comparability	outcome
Weitzman et al, 1990, USA	Racial, social, and environmental risks for childhood asthma	★★	★	★★
Seidman et al, 1991, Israel	Is low birth weight a risk factor for asthma during adolescence?	★★	★★	★★★
Arshad et al, 1993, UK	The effect of genetic and environmental factors on the prevalence of allergic disorders at the age of two years	★★	★★	★★
Azizi et al, 1995, Malaysia	Indoor Air Pollution and Asthma in Hospitalized Children in a Tropical Environment			
Lewis et al, 1995, UK	Prospective study of risk factors for early and persistent wheezing in childhood	★★	★★	★★
Lewis et al, 1996, UK	Study of the aetiology of wheezing illness at age 16 in two national British birth cohorts	★★	★★	★
Schaubel et al, 1996, Canada	Neonatal characteristics as risk factors for preschool asthma	★★★	★	★★★
Sears et al, 1996, New Zealnd	Parental and neonatal risk factors for atopy, airway hyper-responsiveness, and asthma	★★	★★	★★★
Fergusson et al, 1997, New Zealand	Perinatal factors and atopic disease in childhood	★★★	★★	★★
Lilljeqvist et al, 1997, Norway	Low birthweight, environmental tobacco smoke, and air pollution: Risk factors for childhood asthma?	★★		★★
Slezak et al, 1998, USA	Asthma prevalence and risk factors in selected Head Start sites in Chicago	★	★★	★
Wjst et al, 1998, Germany	Pulmonary function in children with initial low birth weight	★★	★★	★

Author , year, region	Study title	Selection	Comparability	outcome
Leadbitter et al, 1999, New Zealand	Relationship between foetal growth and the development of asthma and atopy in childhood	★★	★	★
Rasanen et al, 2000, Finland	Perinatal risk factors for asthma in Finnish adolescent twins	★	★★	★
Steffensen et al, 2000, Denmark	Low birth weight and preterm delivery as risk factors for asthma and atopic dermatitis in young adult males	★★★	★★	★★★
Annnesi-maesano et al, 2001,	In utero and perinatal complications preceding asthma	★★★	★★	★★★
Brooks et al, 2001, USA	Impact of low birth weight on early childhood asthma in the United States	★★	★★	★
Ronmark et al, 2002, Sweden	Incidence rates and risk factors for asthma among school children: A 2-year follow-up Report from the Obstructive Lung Disease in Northern Sweden (OLIN) studies	★★★	★	★★
Anand et al, 2003, UK	Lung function and respiratory health in adolescents of very low birth weight	★★	★	★
Benicio et al,2004, Brazil	Wheezing conditions in early childhood: prevalence and risk factors in the city of Sao Paulo, Brazil	★★	★★	★★
Bolte et al, 2004, Germany	The relation of markers of foetal growth with asthma, allergies and serum immunoglobulin E levels in children at age 5-7 years	★★	★★	★★
Al-kubaisy et al, 2005, Iraq	Risk factors for asthma among primary school children in Baghdad, Iraq	★★★		
Bernsen et al, 2005, Netherlands	Perinatal characteristics and obstetric complications as risk factors for asthma, allergy and eczema at the age of 6 years	★★★	★★	★★★
Nepomnyaschy et al, 2006, USA	Low birthweight and asthma among young urban children	★★★	★★	★★

Author , year, region	Study title	Selection	Comparability	outcome
Remes et al, 2008, Finland	High birth weight, asthma and atopy at the age of 16 years	★★	★★	★
Ortqvist et al, 2009, Sweden	Familial factors do not confound the association between birth weight and childhood asthma	★★	★★	★
Xu et al, 2009, USA	The effects of birthweight and breastfeeding on asthma among children aged 1-5 years	★★	★★	★
Midodzi et al,2010, Canada	Early Life Factors Associated with Incidence of Physician-diagnosed Asthma in Preschool Children: Results from the Canadian Early Childhood Development Cohort Study	★★★★	★★	★★
Bjerg et al, 2011, Sweden	A strong synergism of low birth weight and prenatal smoking on asthma in schoolchildren	★★	★★	★★
Mogensen et al 2011, Sweden	Association between childhood asthma and ADHD symptoms in adolescence – a prospective population-based twin study	★★	★★	★
Suglia et al, 2011, USA	Asthma and obesity in three-year-old urban children: Role of sex and home environment	★★★	★★	★★
To et al, 2012, Canada	Is large birth weight associated with asthma risk in early childhood?	★★★	★★	★★★
Wang et al ,2012, Taiwan	Joint effects of birth outcomes and childhood body mass index on respiratory symptoms	★★★	★★	★★
Kallen et al, 2013, Sweden	Association between preterm birth and intrauterine growth retardation and child asthma	★★★	★★	★★★
Miyake et al, 2013, Japan	Lack of relationship between birth conditions and allergic disorders in Japanese children aged 3 years	★★★	★★	★

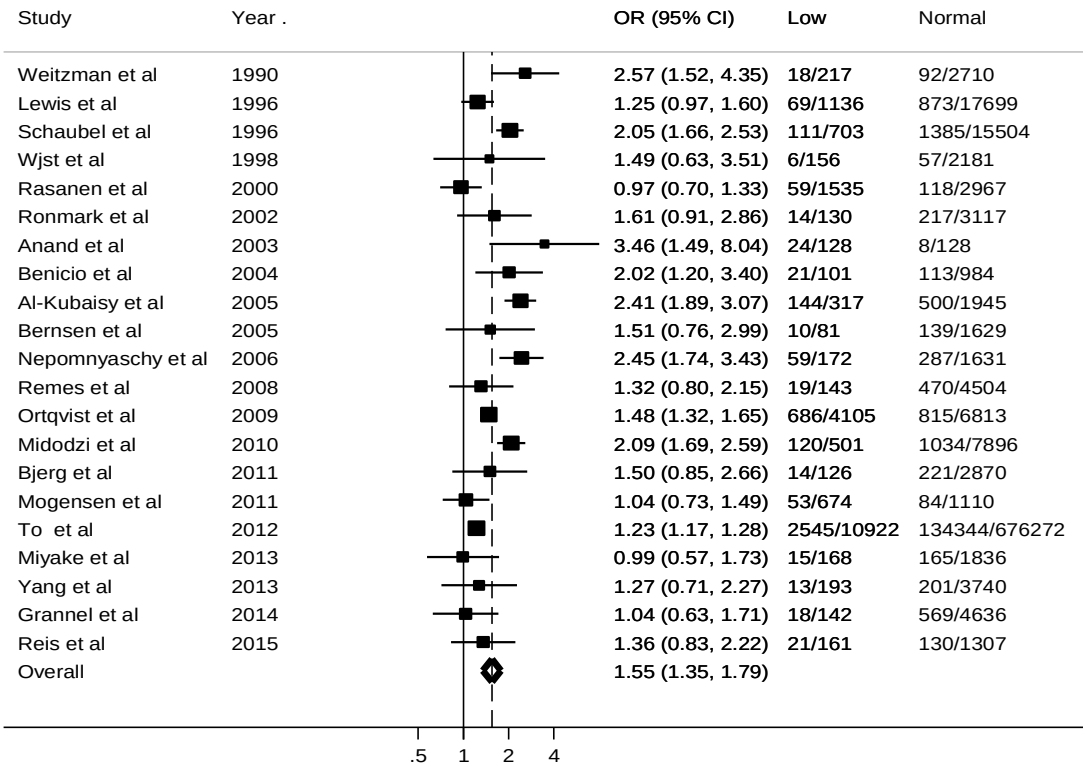
Author , year, region	Study title	Selection	Comparability	outcome
Yang et al,2013, USA	Population-based study on association between birth weight and risk of asthma: A propensity score approach	★★★	★★	★★★
Grannel et al, 2014, UK	Effects of BMI, Fat Mass, and Lean Mass on Asthma in Childhood: A Mendelian Randomization Study	★★★★	★★	★
Reis et al, 2015, Brazil	Prevalence and risk factors for wheezing in Salvador, Brazil: A population-based study	★★★	★★	★

2.3.3.3 Meta-analysis

Low birthweight

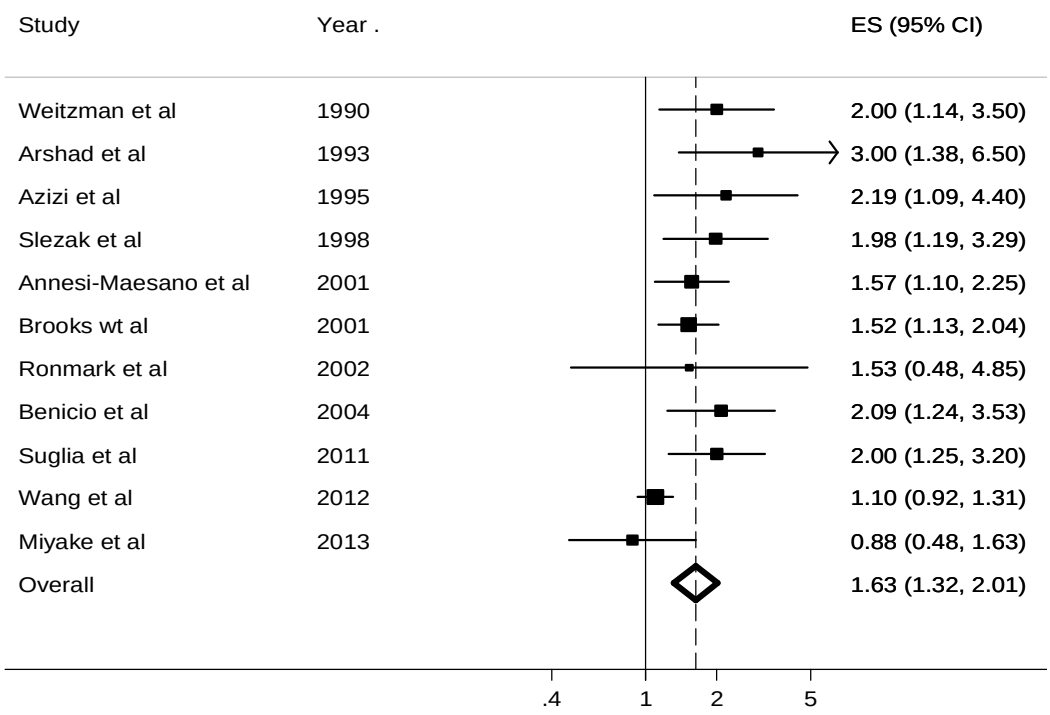
A total of 31 studies contributed data on the number of cases and non-cases of childhood wheezing disorders that included 1,425,480 children. An overall risk estimate of the studies that compared <2.5kg and ≥2.5kg of birthweight groups showed that there was a significant increased odds of wheezing disorders (OR= 1.55, 95% CI: 1.35 to 1.79, P<0.01) for <2.5kg birthweight (Figure 2.2). There was substantial heterogeneity among the studies (I^2 = 81%, 95% CI: 72% to 87%). A further meta-analysis of 11 studies that comprised 105,071 children and provided adjusted odds ratios for the same birthweight comparison groups also showed an increase of wheezing disorder risk by 63% (OR=1.63, 95% CI: 1.32 to 2.01, P<0.01) for the <2.5kg birthweight children (Figure 2.3)

Figure 2.2 Summary unadjusted odds ratios of wheezing disorders for low (<2.5kg) compared with normal (≥2.5kg) birthweight categories



Heterogeneity chi-squared = 107 (df = 20) p < 0.001, I^2 = 81% (95% CI: 72% to 87%), and the estimate of between-study variance Tau-squared = 0.06.

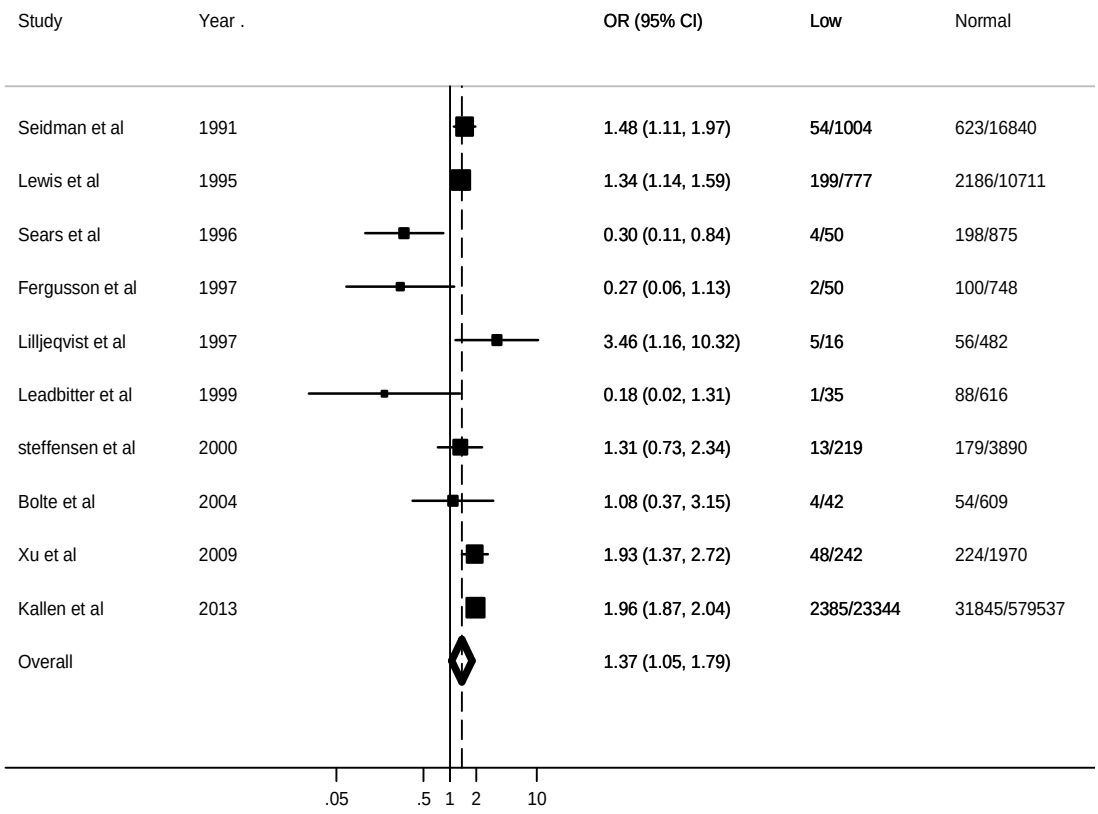
Figure 2.3 Summary adjusted odds ratios of wheezing disorders for low (<2.5kg) compared with normal (≥2.5kg) birthweight categories



Heterogeneity chi-squared = 23 (df = 10) p = 0.01, I^2 = 57% (95% CI: 16% to 78%), and the estimate of between-study variance Tau-squared = 0.06.

A summary risk estimate of 10 studies that provided data on 2.5-4.0kg and <2.5kg birthweight comparison groups showed that there is a significant increase in wheezing disorders risk for the <2.5kg birthweight children (OR=1.37, 95% CI: 1.05 to 1.79, P=0.02), and the between-study variation was very high (I^2=83%, 95% CI: 68 % to 89%), see Figure 2.4. There was not enough data to carry out meta-analysis of adjusted risk estimates for these birthweight comparison groups—only one study contributed (Xu et al., 2009).

Figure 2.4 Summary unadjusted odds ratios of wheezing disorders for low (<2.5kg) compared with normal (2.5-4.0kg) birthweight categories

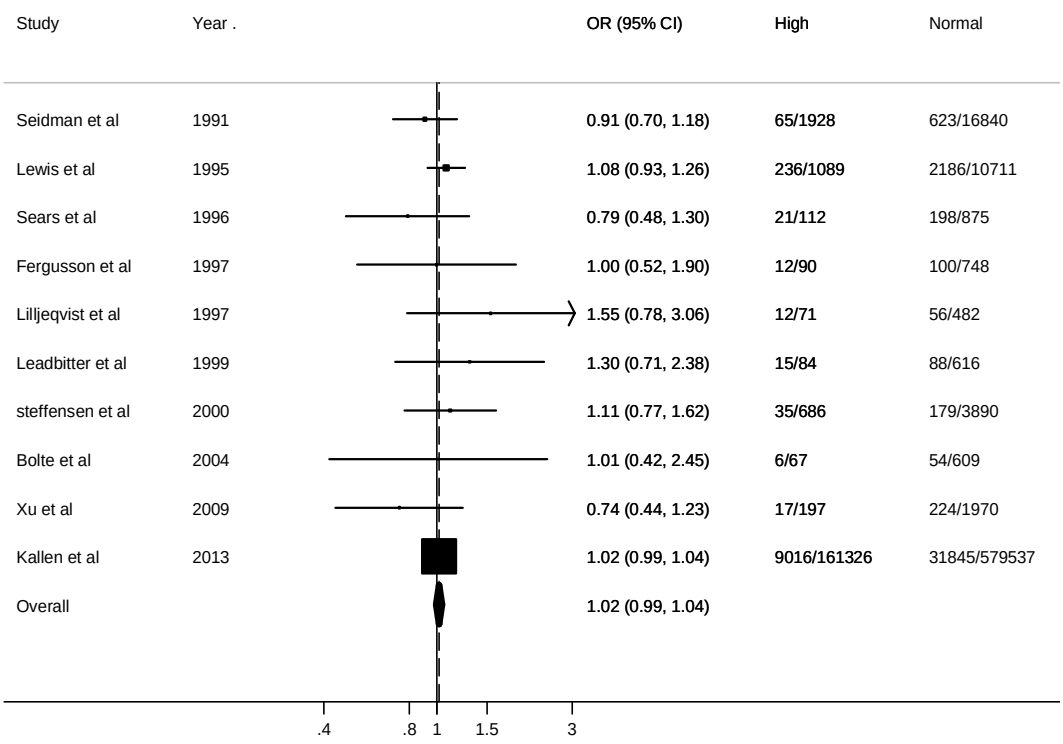


Heterogeneity chi-squared = 50 (df = 9) p < 0.001, I^2 = 83% (95% CI: 68 % to 89%), and the estimate of between-study variance Tau-squared = 0.09.

High birthweight

A total of 10 studies provided data on the number of cases and non-cases of wheezing disorders for 2.5-4.0kg and >4.0kg birthweight comparison groups (Figure 2.5). The overall odds ratio of childhood wheezing disorders for the >4.0kg birthweight was 1.02 (95% CI: 0.99 to 1.04, P=0.13), which was not significantly different from 1. There was no significant heterogeneity among the studies' odds ratio estimates (I^2 = 0%, 95% CI: 0 to 45%). When further investigated if the non-significant heterogeneity was due to the presence of Kallen et al's study (Kallen et al., 2013) that has dominated the pooled risk estimate, both the summary risk estimate and the level of heterogeneity remained stable (OR=1.03, 95% CI:0.92 to 1.15 ; Q=6 (df = 8), P = 0.63, I^2= 0%). There was not enough data to carry out meta-analysis of adjusted risk estimates for these birthweight comparison groups—only one study contributed (Xu et al., 2009).

Figure 2.5 Summary unadjusted odds ratios of wheezing disorders for high (>4.0kg) compared with normal (2.5-4.0kg) birthweight categories



Heterogeneity chi-squared = 6 (df = 9) p = 0.73, I^2 = 0% (95% CI: 0% to 45%) and the estimate of between-study variance Tau-squared = 0.00.

2.3.3.4 Subgroup analyses

Low birthweight and wheezing disorders

Subgroup meta-analyses of 21 studies that compared the low (<2.5kg) and normal (≥2.5kg) birthweight categories showed that the summary risk estimates remained significant in all subgroups of the *a priori* defined covariates, except if wheezing was used as an outcome term or diagnosis was reported by a parent or the studies were low quality (Table 2.4). When the same analysis was carried out on the studies that reported adjusted odds ratios for the same birthweight comparison groups, there was no statistically significant risk of association between low birthweight and wheezing disorders only if birthweight was extracted from e-records or the study age group were 'five years and above' or the studies were high quality (Table 2.5). The between group heterogeneities in the unadjusted odds ratios were significant except when for the outcome terms used (Table 2.4). In the adjusted odds ratios, the

between group heterogeneity was significant for outcome and exposure ascertainment, age group and study period (Table 2.5).

A further subgroup analyses of 10 studies that contributed data on the wheezing disorder cases and non-cases in the low (<2.5kg) and normal (2.5kg-4.0kg) birthweight groups were performed. The results showed inconsistent risk of association among subgroups of all the predefined study characteristics. For example, there was no significant association between low birthweight and wheezing disorders if studies used asthma as an outcome term or sample size of less than 1000 was used or studies were published before 2000 (Table 2.6). The between and within group heterogeneities were significant in most of the study characteristics.

Table 2.4 Subgroup analysis of unadjusted odds ratio of wheezing disorders for normal (≥2.5kg) compared with low birthweight (<2.5kg) categories

	Odds ratio (95% CI)	n	I^2	P_{het} [a]	P_{het} [b]
Outcome terms used					
Asthma	1.56 (1.34 to 1.82)	19	83%	<0.01	0.82
Wheezing	1.50 (0.95 to 2.39)	2	63%	0.1	
Outcome ascertainment					
E-records	1.64 (1.15 to 2.36)	5	86%	<0.01	<0.01
Parent	1.49 (1.26 to 1.77)	15	70%	<0.01	
Physician	2.09 (1.69 to 2.59)	1	-		
Exposure ascertainment					
E-records	1.52 (1.29 to 1.78)	13	83%	<0.01	0.01
Parent	1.48 (0.87 to 2.53)	4	86%	<0.01	
No mention	1.79 (1.32 to 2.42)	4	24%	0.27	
Age during diagnosis					
Five years and above	1.41 (1.21 to 1.63)	13	74%	<0.01	<0.01
Under five years	1.56 (1.12 to 2.19)	5	73%	0.02	
Mixed (0-19 years)	2.14 (1.77 to 2.57)	3	0%	0.75	
Sample size					
1000+	1.53 (1.33 to 1.76)	20	81%	<0.01	0.03
<1000	3.46 (1.49 to 8.04)	1	-		
Study period					
<2000	1.76 (1.23 to 2.51)	4	74%	<0.01	<0.01
2000+	1.51 (1.30 to 1.76)	17	81%	<0.01	
Study type					
Cohort	1.51 (1.32 to 1.72)	20	77%	<0.01	<0.01
Case-control	2.41 (1.89 to 3.07)	1	-		
Study quality score [c]					
High (7-9/9)	1.63 (1.22 to 2.18)	7	90%	<0.01	<0.01
Medium (5-6/9)	1.41 (1.24 to 1.61)	11	23%	0.22	
Low (≤4/9)	1.90 (0.90 to 3.98)	3	91%	<0.01	

[a] P for heterogeneity within subgroups.

[b] P for heterogeneity between subgroups.

[c] Sensitivity analysis.

Table 2.5 Subgroup analysis of adjusted odds ratios of wheezing disorders for normal (≥2.5kg) compared with low (<2.5kg) birthweight categories

	Odds ratio (95% CI)	n	I^2	P_{het} [a]	P_{het} [b]
Outcome terms used					
Asthma	1.59 (1.28 to 1.98)	10	57%	0.01	0.13
Wheezing	2.09 (1.24 to 3.53)	1	-		
Outcome ascertainment					
Parent	1.53 (1.24 to 1.89)	9	55%	0.02	0.02
Physician	2.52 (1.50 to 4.23)	2	0%	0.06	
Exposure ascertainment					
E-records	1.21 (0.96 to 1.52)	4	33%	0.21	<0.01
Parent	1.71 (1.29 to 2.89)	2	0%	0.42	
No mention	2.14 (1.65 to 2.79)	5	0%	0.93	
Age during diagnosis					
Five years and above	1.11 (0.93 to 1.32)	2	0%	0.58	<0.01
Under five years	1.63 (1.11 to 2.40)	4	57%	0.07	
Mixed (0-19 years)	1.85 (1.49 to 2.31)	5	0%	0.85	
Sex					
Boys	1.52 (1.13 to 2.04)	1	-	-	0.59
Both	1.67 (1.30 to 2.14)	10	61%	<0.01	
Sample size					
1000+	1.56 (1.24 to 1.97)	9	59%	<0.01	0.06
<1000	2.01 (1.36 to 3.09)	2	0%	0.82	
Study period					
<2000	2.16 (1.59 to 2.93)	4	0%	0.83	<0.01
2000+	1.44 (1.15 to 1.81)	7	56%	0.34	
Study type					
Cohort	1.60 (1.29 to 1.99)	10	58%	0.01	0.21
Case-control	2.19(1.09 to 4.4)	1	-		
Study quality score [c]					
High (7-9/9)	1.42 (0.80 to 2.54)	2	81%	0.02	<0.01
Medium (5-6/9)	1.70 (1.41 to 2.03)	9	11%	0.34	

[a] P for heterogeneity within subgroups.

[b] P for heterogeneity between subgroups.

[c] Sensitivity analysis.

Table 2.6 Subgroup analysis unadjusted odds ratios of wheezing disorders for normal (2.5-4.0kg) compared with low (<2.5kg) birthweight categories

	Odds ratio (95% CI)	n	I^2	P_{het} [a]	P_{het} [b]
Outcome terms used					
Asthma	1.33 (0.95 to 1.85)	9	75%	<0.001	<0.01
Wheezing	1.34 (1.14 to 1.59)	1	-		
Outcome ascertainment					
E-records	1.50 (0.98 to 2.30)	3	81%	<0.01	<0.01
Parent	1.61 (1.16 to 2.24)	4	51%	0.1	
Physician	0.49 (0.13 to1.89)	3	78%	0.01	
Exposure ascertainment					
E-records	1.27 (0.93 to 1.72)	9	84%	<0.001	0.9
No mention	1.93 (1.37 to 2.72)	1	-		
Age during diagnosis					
Five years and above	1.10 (0.76 to 1.59)	8	66%	<0.01	<0.01
Mixed (0-19 years)	1.96 (1.87 to 2. 04)	2	0%	0.9	
Gender					
Mixed	1.32 (0.94 to 1.85)	8	84%	<0.001	0.04
Boys	1.44(1.12 to 1.87)	2	0%	0.71	
Sample size					
1000+	1.62 (1.29 to 2.02)	5	82%	<0.001	0.01
<1000	0.61 (0.20 to 1.91)	5	75%	0.03	
Study period					
<2000	1.00 (0.62 to 1.63)	6	76%	<0.01	<0.01
2000+	1.95(1.85 to 2.05)	4	0.6%	0.39	
Study quality score [c]					
High (7-9/9)	1.14 (0.75 to 1.74)	6	81%	<0.001	0.01
Medium (5-6/9)	1.56 (1.10 to 2.21)	2	70%	0.06	
Low (≤4/9)	0.86 (0.03 to 23.90)	2	88%	<0.01	

[a] P for heterogeneity within subgroups.

[b] P for heterogeneity between subgroups.

[c] Sensitivity analysis.

High birthweight and wheezing disorders

Subgroup meta-analyses of 10 studies that contributed data on the cases and non-cases of wheezing disorders in the high (>4.0kg) and normal (2.5-4.0kg) birthweight categories showed that the risk of association was not significant across all categories of the predefined study characteristics and the study quality (Table 2.7). Both the within and between group heterogeneities were insignificant.

Table 2.7 Subgroup analysis of unadjusted odds ratios of wheezing disorders for normal (2.5-4.0kg) compared with high birthweight (>4.0kg) categories

	Odds ratio (95% CI)	n	I^2	P_{het} [a]	P_{het} [b]
Outcome terms used					
Asthma	1.02 (0.99 to 1.04)	9	0%	0.69	0.45
Wheezing	1.08 (0.93 to1.26)	1	0%	0.73	
Outcome ascertainment					
E-records	1.02 (0.99 to1.04)	3	0%	0.69	0.82
Parent	1.06 (0.89 to1.25)	4	5%	0.36	
Physician	1.04 (0.80 to1.36)	3	0%	0.40	
Exposure ascertainment					
E-records	1.02 (1.00 to1.04)	9	0%	0.80	0.22
No mention	0.74 (0.44 to1.23)	1			
Age during diagnosis					
Five years and above	1.04 (0.93 to1.17)	8	0%	0.73	0.66
Mixed (0-19 years)	0.96 (0.76 to1.22)	2	34%	0.22	
Gender					
Boys	0.97(0.79 to1.20)	2	0%	0.38	0.66
Mixed	1.02 (1.00 to1.04)	8	0%	0.64	
Sample size					
1000+	1.02 (0.99 to1.04)	5	0%	0.55	0.78
<1000	1.02(1.00 to1.04)	5	0%	0.55	
Study period					
<2000	1.04 (0.92 to1.17)	6	0%	0.51	0.75
2000+	1.02 (0.99 to1.04)	4	0%	0.63	
Study quality score [c]					
High (7-9/9)	1.02 (0.99 to 1.04)	3	0	0.54	0.54
Medium (5-6/9)	1.01 (0.89 to 1.15)	4	0	0.42	
Low (≤4/9)	1.26 (0.87 to 1.82)	3	0	0.65	

[a] P for heterogeneity within subgroups.

[b] P for heterogeneity between subgroups.

[c] Sensitivity analysis.

2.3.3.5 Meta-regression analysis

Birthweight and wheezing disorders

Investigating the sources of between-study heterogeneities of the unadjusted low birthweight odds ratios showed that 64% (adjusted R-squared=64%, P=0.04) of the variance was explained by the *a priori* selected covariates in the <2.5kg and ≥2.5kg birthweight comparisons. However, none of the variance was explained by the *a priori* selected covariates in the adjusted <2.5kg and ≥2.5kg and the unadjusted <2.5kg and 2.5-4.0kg birthweight comparisons (Table 2.8). No further investigation of between-study heterogeneity (i.e. meta-regression analysis) was carried out for the high (>4.0kg) and normal (2.5-4.0kg) birthweight categories' odds ratios as there was no statistically significant variation among the studies (Table 2.8).

Table 2.8 Meta-regression analysis of odds ratios of wheezing disorders for low compared with normal birthweight categories

Study characteristics	Coefficient (95% CI)	P-value
Unadjusted odds ratios for normal (≥2.5kg) versus low (<2.5kg)		
Outcome terms used (ref=asthma)	-0.14 (-0.58 to 0.31)	0.53
Outcome ascertainment (ref=physician)	-0.03 (-0.27 to 0.22)	0.83
Exposure ascertainment (ref=e-records)	0.01 (-0.19 to 0.21)	0.91
Age during diagnosis (ref=five and above)	0.24 (0.05 to 0.44)	0.02
Sample size (ref=less than 1000)	-1.00 (-2.02 to 0.02)	0.05
Study period (ref=before 2000)	-0.19 (-0.53 to 0.14)	0.24
Study type (ref=cohort)	0.65 (0.16 to 1.14)	0.01
Overall (adjusted R-squared) = 64%)		0.04
Adjusted odds ratios for normal (≥2.5kg) versus low (<2.5kg)		
Outcome terms used (ref=asthma)	-0.54 (-2.62 to 1.54)	0.38
Outcome ascertainment (ref=physician)	-0.08 (-1.23 to 1.07)	0.78
Exposure ascertainment (ref=e-records)	0.82 (-0.83 to 2.48)	0.17
Age during diagnosis (ref=five and above)	-0.24 (-1.19 to 0.72)	0.40
Sex (ref=mixed)	0.55 (-0.60 to 1.71)	0.18
Sample size (ref=less than 1000)	0.01 (-1.65 to 1.67)	0.98
Study period (ref=before 2000)	0.58 (-1.54 to 2.70)	0.36
Study type (ref=cohort)	-0.07 (-3.04 to 2.91)	0.93
Overall (adjusted R-squared) = 0%)		0.28
Unadjusted odds ratios for normal (2.5-4.0kg) versus low (<2.5kg)*		
Outcome ascertainment (ref=physician)	0.86 (-0.42 to 2.15)	0.12
Exposure ascertainment (ref=e-records)	-0.44 (-2.00 to 1.12)	0.44
Age during diagnosis (ref=five and above)	0.52 (-1.37 to 2.41)	0.44
sex (ref=mixed)	1.21 (-2.08 to 4.5)	0.33
Sample size (ref=less than 1000)	-0.01 (-2.70, 2.68)	0.99
Study period (ref=before 2000)	0.20 (-1.93, 2.31)	0.79
Overall (adjusted R-squared= 0%)		0.42

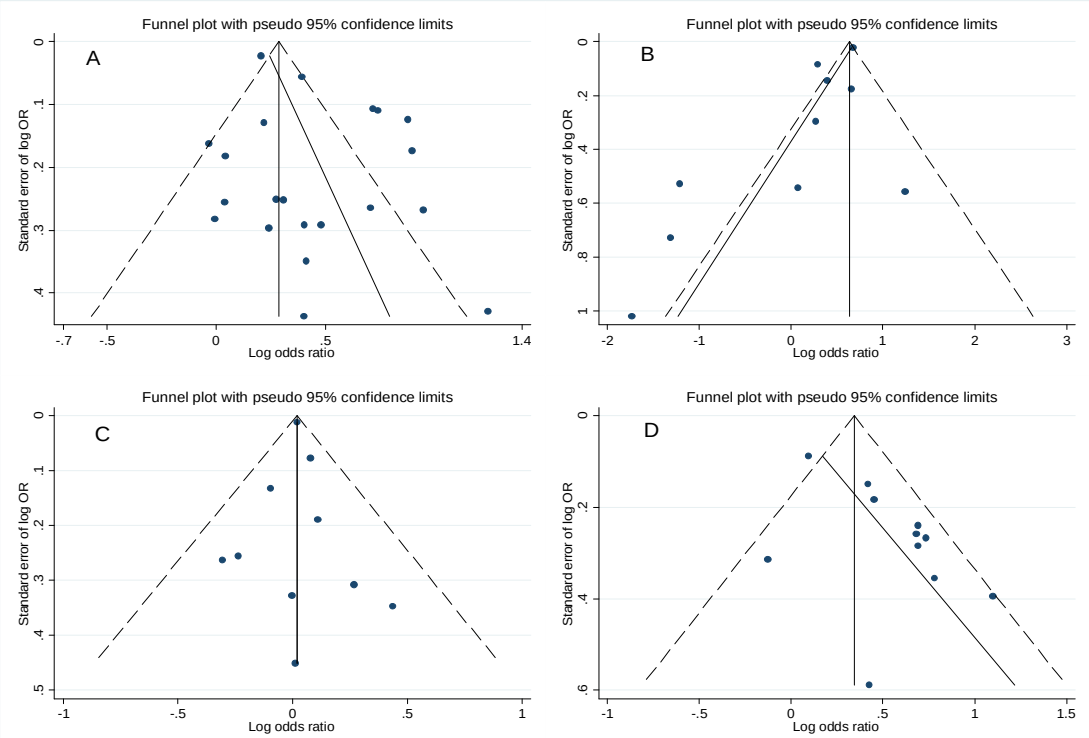
*: outcome terms was dropped due to collinearity.

2.3.3.6 Investigating biases (small study effects)

Birthweight and wheezing disorders

The funnel plots showed no evidence of asymmetry for the high (>4.0kg versus 2.5-4.0kg) birthweight unadjusted odds ratios (Figure 2.6C). However there was some evidence of asymmetry for the low birthweight (2.5kg versus ≥2.5kg and <2.5kg versus 2.5-4.0kg) unadjusted odds ratio estimates (Figure 2.6A and Figure 2.6B) and low birthweight (2.5kg versus ≥2.5kg) adjusted odds ratio estimates (Figure 2.6D). This was also reflected in Egger's tests (Table 2.9), with no evidence of small-study effects for Figure 2.6C (p=0.99), but some evidence for Figure 2.6A (p=0.04), Figure 2.6B (p=0.02) and Figure 2.6D (p=0.02).

Figure 2.6 Egger's funnel plots of birthweight and wheezing disorder studies



(A) <2.5kg (low) versus ≥2.5kg (normal) birthweight ;(B) <2.5kg (low) versus 2.5-4.0kg (normal) birthweight ; (C) >4.0kg (high) versus 2.5-4.0kg (normal) birthweight and (D) <2.5kg (low) versus ≥2.5kg (normal) birthweight risk estimate funnel plots. Unadjusted odds ratio in A, B, and C, and adjusted risk estimates in D. In all funnel plots, the middle solid line is the summary odds ratio estimate and the two diagonal dotted lines are the 95% confidence limits around the summary odds ratio, and the slant solid lines in figures A, B and D are the fitted regression lines for Egger's small-study effect test. Note that the fitted regression line in C is exactly aligned to and obscured by the middle solid line.

Table 2.9 Egger's test of bias for small study effects in birthweight and wheezing disorders studies

Parameter	Coefficient (95% CI)	P-value
Unadjusted odds ratios for normal (≥2.5kg) versus low (<2.5kg)		
Slope	0.21 (0.10 to 0.32)	<0.01
Bias	1.32 (0.07 to 2.58)	0.04
Adjusted odds ratios for normal (≥2.5kg) versus low (<2.5kg)		
Slope	-0.01 (-0.32 to 0.30)	0.92
Bias	2.09 (0.51 to 3.67)	0.02
Unadjusted odds ratios for normal (2.5-4.0kg) versus low (<2.5kg)		
Slope	0.70 (0.60 to 0.80)	<0.01
Bias	-1.90 (-3.40 to -0.40)	0.02
Unadjusted odds ratios for normal (2.5-4.0kg) versus high (>4.0kg)		
Slope	0.12 (-0.01 to 0.05)	0.17
Bias	0.002 (-0.73 to 0.73)	0.99

2.3.4 Discussion

2.3.4.1 Key findings

This meta-analysis is the most recent comprehensive analysis as it includes 77 studies published until March 2015. The results showed that low birthweight (defined as <2.5kg) children have an increased risk of wheezing disorders when compared to the normal birthweight children (defined as ≥2.5) based on unadjusted (OR= 1.55, 95% CI: 1.35 to 1.79) and adjusted (OR=1.63, 95% CI: 1.32 to 2.01) estimates. If the low birthweight are compared with those of 2.5-4.0kg birthweight children, there is a 37% increase of wheezing disorders risk (unadjusted OR=1.37, 95% CI: 1.05 to 1.79) although it must be considered that there was a significant between-study heterogeneity and some evidence of small study effects or publication bias. However, there is a weak evidence to suggest that high birthweight (defined as >4.0kg) children have increased odds of wheezing disorders (unadjusted OR=1.02, 95% CI: 0.99 to 1.04, P=0.13), when compared to the normal birthweight (defined as 2.5-4.0kg).

2.3.4.2 Results in context of previous reviews and meta-analyses

The unadjusted pooled risk estimates for low birthweight are moderately higher than those of a recent meta-analysis by Mu et al. (2014) that reported unadjusted ORs of 1.28 (95% CI: 1.09 to 1.50) and 1.34 (95% CI: 1.13 to 1.60) for studies that used two (i.e. <2.5 versus ≥2.5kg) and three (i.e. <2.5 versus 2.5-4.0kg) birthweight categories respectively. However, some studies included in their meta-analysis were not consistent had adult population, and the fact that more studies were included in this study than theirs may have possibly influenced the difference in robustness of the summary risk estimates. Xu et al. (2014) also reported a RR of 1.15 (95% CI: 1.08 to 1.22) for low birthweight children. However, birthweight categories and risk reporting methods were not consistent across the studies included their meta-analysis so their results may not be comparable with this meta-analysis's findings.

The unadjusted summary risk of association between high birthweight and wheezing disorders was not statistically significant in contrast to a previous meta-analysis by Flaherman and Rutherford (2006) that reported a 20% increase of asthma risk (RR= 1.2; 95% CI: 1.1 to 1.3). However, it must be noted that the studies included in the previous meta-analysis had used different cut-off points and measurement types for

high birthweight, and risk estimation methods (relative risk and odds ratio). Combining studies with different cut-off points may under or overestimate the summary risk estimate. Likewise, combining relative risks and odd ratios may have a similar effect in common diseases (McNutt et al., 2003; Viera, 2008).

The studies that were not included in this meta-analysis reported inconsistent risk of association for the low birthweight categories (Appendix D), although a recent ISAAC Phase III study that used similar birthweight categories has reported an odds ratio of 1.20 (95%: 1.12 to 1.30) for low birthweight (Mitchell et al., 2014). However, the majority of the studies reported no risk of association for the high birthweight group, agreeing with the findings of this meta-analysis.

Based on the pooled odds ratio results, the adjusted and unadjusted summary odds ratios for two and three birthweight categories were similar. This may strongly suggest that low birthweight is an independent risk factor for childhood wheezing disorders although it must be noted that there are also some evidence of bias in the funnel plots and Egger's tests of bias (Egger et al., 1997; Sterne et al., 2001) which may indicate that there was potential publication bias for studies that showed no significant risk of association (Sterne and Habord, 2004).

Based on the heterogeneity measures (Q-test and I^2), there was a considerable level of between-study variation in the low birthweight unadjusted risk estimates although this could be due to high precision or high sample size of studies included in this meta-analysis (Rücker et al., 2008) as illustrated in the forest plot (Figure 2.2). The studies were mostly precise and had narrow confidence intervals. However, there was no significant heterogeneity among the unadjusted risk estimates of high birthweight and asthma and this could be due to having less precise risk estimates with wider confidence intervals as demonstrated by the forest plot (Figure 2.5).

2.3.4.3 Strengths and weaknesses

This work has limitations and results should be interpreted cautiously. First, in the low birthweight and overweight summary risk estimates, there was a significant and substantial level of between-study variation that was not explained by the *a priori* selected covariates. Second, there is also some evidence of funnel plot asymmetry which may indicate a potential small study effect such as potential publication bias (Egger et al., 1997). Third, as in any systematic review and meta-analysis, a

possibility of potentially relevant studies being missed cannot be ruled out. Fourth, the results are based on epidemiologic observational studies and are solely dependent on the quality of the primary studies included.

The strength of this work is that it was possible to produce consistent risk estimates due to the use of harmonised data. Combining adjusted risk estimates was a primary choice among previous authors. This technique may, however, under or overestimate the association between exposure and outcome variables due to exclusion of studies that used non-standard weight categories or combining all irrespective of the type of exposure categorisation method used. In order to improve validity of the summary risk estimates, data harmonisation techniques were implemented and more studies were included than if previous authors' techniques were used. Most importantly, it was possible to produce more consistent summary risk estimates of birthweight (i.e., low and high birthweight) on wheezing disorders than if results of studies were to be combined irrespective of cut-off points as used by previous authors. The other strength of this work is also that it was possible to extract and analyse both adjusted and unadjusted risk estimates, which can be used as an internal validation to each other.

In conclusion, the results show that there is strong evidence that suggests low birth (<2.5kg) is a risk factor for wheezing disorders during childhood and adolescence. However, there is weak evidence for an increase of asthma or wheezing disorders risk for high birthweight children.

2.4 BMI and wheezing disorders

2.4.1 Critique of past systematic reviews and meta-analyses

Previous observational epidemiologic studies suggest that overweight/obesity and childhood asthma are associated. However, an inconsistency in the results remains. A meta-analysis of four observational epidemiologic studies reported a 50% (RR=1.50, 95% CI: 1.2 to 1.8) increased risk of childhood asthma for overweight (Flaherman and Rutherford, 2006). However, the included studies used a variety of risk estimate definitions: three used odds ratios (Castro-Rodriguez et al., 2001; Chinn and Rona, 2001; Xu et al., 2002) and another used relative risk (Gilliland et al., 2003).

Results from a recent meta-analysis of 6 cohort studies by Chen et al. (2013) reported relative risks of 1.19 (95% CI: 1.03 to 1.37) and 2.02 (95% CI: 1.16 to 3.5) for childhood asthma in those who were overweight and obese respectively. In a meta-analysis of 6 prospective studies, it was also reported that there is a 35% (RR=1.35, 95% CI: 1.15 to 1.58) and 50% (RR=1.50, 95% CI: 1.22 to 1.83) increase in risk of childhood asthma for overweight and obesity respectively (Egan et al., 2013). However, the age of study populations, Body Mass Index (BMI) categorisations, and risk estimate definitions were not consistent across the studies included in the two meta-analyses. For example, in the meta-analysis by Egan et al. (2013), one study used data-driven quintile BMI categories, (Gold et al., 2003) whilst the other two studies used only high risk children, (Zhang et al., 2010; Ho et al., 2011), two used relative risk, (Gilliland et al., 2003; Gold et al., 2003) one used hazard ratios (Mannino et al., 2006), and the other three used odds ratio (Mamun et al., 2007; Zhang et al., 2010; Ho et al., 2011) as risk estimate definitions. Likewise, in the meta-analysis by Chen et al. (2013), one study included adult population (Burgess et al., 2007), and another used bronchitis as the outcome variable instead of asthma or wheezing symptoms (Lee et al., 2013).

Combining studies that include child and adult populations, use non-standard and inconsistent BMI categories and a variety of risk estimate definitions in a meta-analysis may bias the summary risk estimates. For example, suppose that two studies used 30^{th} centile, three used 10^{th} centile and four other used 5^{th} centile as cut-off points for underweight. Then, it becomes difficult to combine these 9 studies in a

meta-analysis as the cut-off points used are not equivalent. The last group used a standard BMI cut-off point for underweight (5th centile) and the other two groups used cut-off points of convenience where some individuals grouped as underweight in these studies have normal BMI according to the standard BMI categorisation methods. Similarly, although the estimates from odds ratios and relative risks are similar when the disease is rare (<10%), they diverge as the prevalence increases (McNutt et al., 2003; Viera, 2008), potentially biasing the summary risk estimates derived from combined odds ratios and relative risks.

2.4.2 Sythematic literature review methods

2.4.2.1 Literature Search Strategy

The reviews were carried out in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009). Online searches were carried out using the EMBASE and MEDLINE databases. Table 2.10 gives the details of terms and phrases used for the literature search.

2.4.2.2 Inclusion criteria

Eligible papers were those published as an article, in English, and reported original research about the effects of BMI on wheezing disorders in children 0-19 years of age. No lower limit for a time of publication was set. However, the literature search covered the period to March 2015. Case-control, cohort, and cross-sectional studies were included.

2.4.2.3 Exclusion criteria

Papers were excluded if:

a) BMI was modelled as a continuous variable as an assumption was made that the risk of outcome (wheezing disorder) is higher in the lower and higher ends of BMI bands. Thus, 'standard' categorical variable BMI was considered to be more appropriate;
b) Authors claimed BMI was included in their analyses but no comparison group or risk estimates were presented in text or table of the papers;
c) Authors used data-driven multiple categories of BMI that cannot be converted into the 'standard' categories;
d) Studies included an adult population with no separate data available for children and adolescents; and,
e) Authors used data-driven multiple BMI categories and if the number of categories presented were generally too few (<4) to allow combination with studies through estimating nonlinear dose-response curves (Orsini et al., 2012).

Table 2.10 Terms and phrases used during literature search

	Body mass index
1	BMI
2	Current weight
3	Child weight
4	Child BMI
5	Child obesity
6	Obes* adj2 children
7	underweight
8	overweight
9	Child* obesity
10	Child* growth
11	High BMI
12	High weight
13	Low BMI
14	Asthm*
15	Wheez*
16	wheezing
17	Wheezing disorders
18	Asthma in children
19	Childhood asthma
20	Childhood wheez*
21	1-13/or
22	14-20/or
23	21-22/and
24	Limit 23 to English language

N.B: First literature searches were conducted on 4th June 2014. Updated was carried out on 8th April 2015.

2.4.2.4 Data extraction

For the studies that were eligible to be included in the meta-analyses, the following characteristics were extracted:

a) Authors' name;
b) Year of publication;
c) Country of study;
d) Study design;
e) Sample size;
f) Study age group and gender;

g) Diagnosis (outcome) terms;
h) Body mass index (exposure) categories;
i) Body mass index categorisation methods;
j) Outcome and exposure ascertainment methods; and,
k) Risk estimates.

2.4.2.5 Data standardisation

Exposure variable

Data on exposure variable varied according to the cut-off points of BMI categories adopted by authors:

a) The CDC: $<5^{th}$ centile, $\geq 5^{th}$ and $<85^{th}$ centiles, $\geq 85^{th}$ and $<95^{th}$ centiles, and $\geq 95^{th}$ centile for underweight, normal, overweight, and obese categories respectively (CDC, 2014);
b) The International Obesity Task Force (IOTF): Age and sex specific cut-off points that are extrapolated from the adult BMI cut-offs of $18.5kg/m^2$, $25kg/m^2$, and $30kg/m^2$ for underweight, overweight, and obesity respectively (Cole et al., 2000; Cole et al., 2007);
c) The WHO: 85^{th}-95^{th} ($1SD^{+}$) and $\geq 95^{th}$ ($2SD^{+}$) centiles for overweight and obese, respectively (NOO, 2011); and,
d) Data-driven multiple BMI categories.

For comparability and not to lose data due to variation in categorisation methods, data were standardised as follows:

a) Where authors used one of the standard category methods (CDC, IOTF or WHO), the reported adjusted risk estimates and data on the number of cases and non-cases of each BMI comparison group were combined for meta-analysis without any change;
b) Where authors adopted data-driven BMI categories with the CDC, IOTF or WHO normal category as a reference and where the number of participants in each category was available, the stratum based number of cases and non-cases were aggregated before being combined with the other studies for meta-analysis of unadjusted risk estimates; and,

c) Where authors adopted two or three birthweight categories with CDC, IOTF or WHO normal category as a reference and provided adjusted risk estimates, the stratum based risk estimates were aggregated using from Hamling et al. (2008) method before being combined with the other studies for meta-analysis of adjusted risk estimates.

Outcome variable

Study authors used one or multiple outcome terms. Thus, for comparability among studies, where authors used a single outcome, for example, asthma or wheezing, the quoted outcome term by the author and its risk estimate were assumed for analysis. However, where authors used multiple outcome terms, the term that was highest in the hierarchy and its risk estimate were assumed for analysis. For example, if asthma and wheezing were used together, asthma was preferred over wheezing.

2.4.2.6 Quality assessment

Papers included in the review and meta-analysis were assessed for risks of bias using the Newcastle-Ottawa quality assessment scale (Wells et al., 2000), see Appendix A , Appendix B and Appendix C for details on scoring guidelines used.

2.4.2.7 Statistical analysis

Random effects models were adopted in pooling estimates using the DerSimonian and Laird method (DerSimonian and Laird, 1986). If studies presented stratum-specific estimates (e.g. by gender), then to provide correct measures of heterogeneity, the risk estimates were aggregated using fixed-effect models before being combined with the other studies for meta-analyses of adjusted risk estimates in a random-effects model. Likewise, if the number of cases and non-cases in each stratum were reported, the total number of cases and non-cases were aggregated before being combined with the other studies for meta-analyses.

The Cochrane Q-test (Whitehead and Whitehead, 1991) and the I^2 (Higgins and Thompson, 2002) were used in estimating between-studies heterogeneity. Sub-group meta-analyses and sensitivity analyses were performed in order to assess the robustness of the risk associations and levels of between-study heterogeneities within a covariate. Then, meta-regression (Lau et al., 1998; Borenstein et al., 2009d)

of risk estimates were performed if a covariate showed significant heterogeneity among its levels.

Publication bias and small study effects were investigated using symmetry funnel plots and bias test models (Egger et al., 1997; Sterne et al., 2001). All meta-analyses were carried out in Stata software version 12 (StataCorp, 2011).

2.4.3 Results

2.4.3.1 Literature search

A total of 4,013 papers were retrieved from EMBASE and Medline. Of these, 91 were read in full. Out of the 91 papers, 45 were included in the review (Figure 2.7). A total of 39 studies that reported either the risk estimates or number of cases and non-cases of wheezing disorders in each exposure group were included in the meta-analysis (Table 2.11). Results from six studies were not combined with the other studies for meta-analysis as they used slightly different centile cut-off points (von Kries et al., 2001; Mai et al., 2007; Okabe et al., 2011; Okabe et al., 2012; Mitchell et al., 2013; Willeboordse et al., 2013).

The 39 studies were from Europe (12), Americas (18), Asia (7) and Oceania (2). The studies were cross-sectional (25), case-control (2) and cohort (12). Only 5 of 39 studies involved a sample population of <1000 each.

2.4.3.2 Quality of studies

Out of the total 39 studies included in the meta-analysis: twenty-five scored 7-9, thirteen scored 5-6, and their risks of biases can be interpreted as ‘low’ and ‘moderate’ respectively (Table 2.12).

Figure 2.7 Body mass index and wheezing disorders literature search flow chart

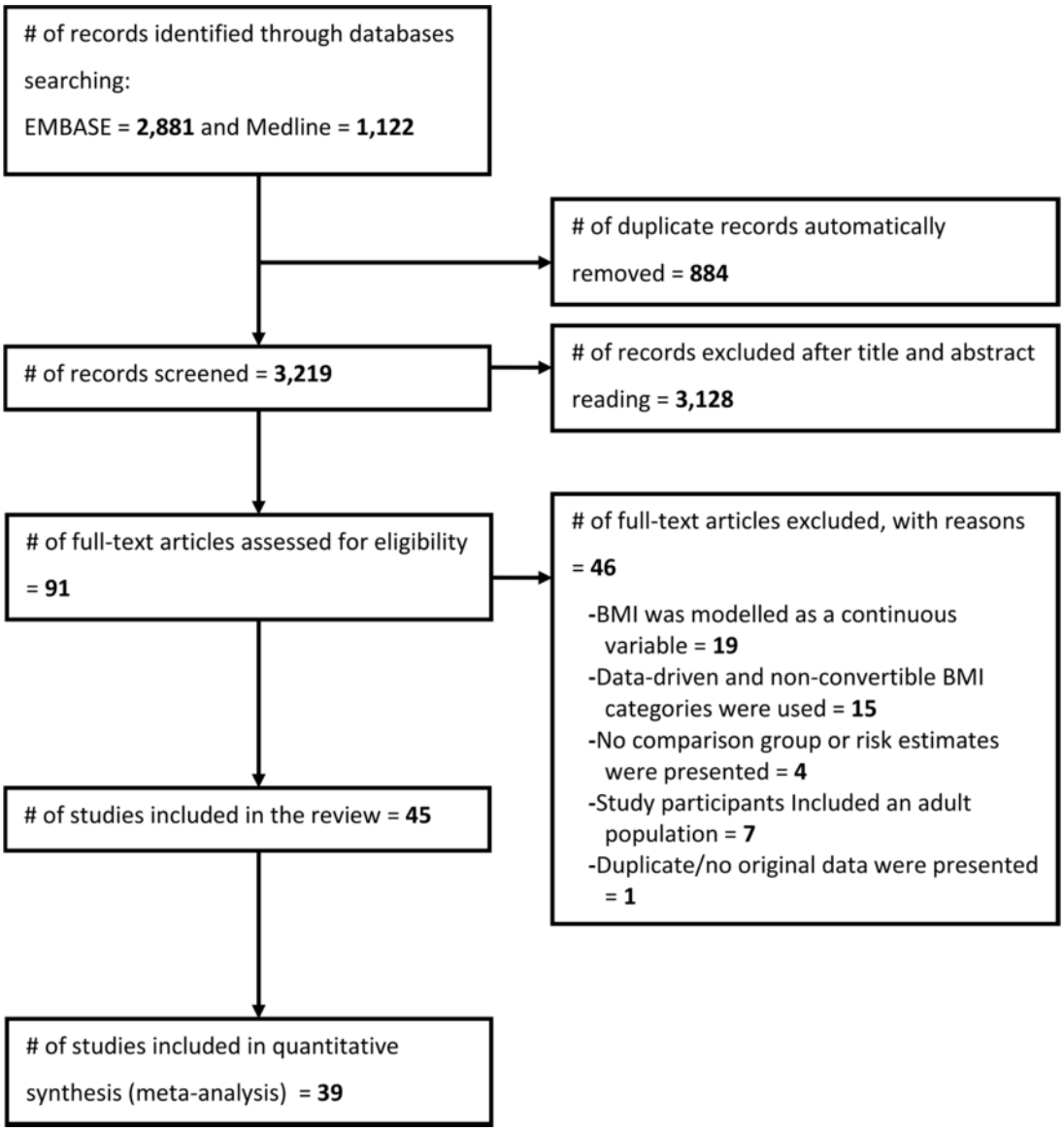


Table 2.11 Characteristics of BMI and wheezing disorders studies included in the meta-analysis

Author , year, region	Study design	Sample size	Participants age and gender	outcome term	Outcome ascertainment	Exposure ascertainment	Exposure categories
Gennuso et al. (1998), USA	CC	171	4-16 years mixed	asthma	e-records	e-records	overweight: 85th-95th and obesity: ≥95th centile
Chinn and Rona (2001), UK	PC	4,743	9 years mixed	asthma	parent	trained	IOTF cut-offs for underweight, overweight, and obese.
Rodríguez et al. (2002), USA	CS	12,388	2 months-16 years mixed	asthma	parent	no mention	overweight: ≥85th centile
Gilliland et al. (2003), USA	PC	3,792	7-18 years mixed	asthma	child	trained	overweight: 85th-95th and obese:≥95th centile
Bibi et al. (2004), Israel	CS	5,984	8 years mixed	asthma	parent	trained	non-obese≤ 95th and obese:≥95th centile
Cassol et al. (2005), Brazil	CS	4,010	13-14 years mixed	asthma	child	trained	underweight: <5th overweight: 85th-95th and obese:≥95th centile
Saha et al. (2005), USA	RC	2,544	5-18 years mixed	asthma	e-records	e-records	overweight: 85th-95th and obese:≥95th centile
Wickens et al. (2005), New Zealand	CS	1,287	11-12 years mixed	asthma	parent	trained	IOTF cut-offs for underweight, overweight, and obese.
Kwon et al. (2006), USA	CS	853	2-11 years mixed	asthma	parent	no mention	underweight: <5th , overweight: 85th-95th and obese:≥95th centile
Shamssain (2006), UK	CS	7,000	5-16 years mixed	asthma	parent	no mention	IOTF cut-offs for underweight, overweight, and obese.
Van De Ven et al. (2006), Netherlands	CS	10,087	12-14 years mixed	asthma	parent	parent	IOTF cut-offs for underweight, overweight, and obese.
Davis et al. (2007), USA	CS	471,969	12-18 years mixed	asthma	child	child	underweight: <5th , overweight: 85th-95th and obese:≥95th centile

Author , year, region	Study design	Sample size	Participants age and gender	outcome term	Outcome ascertainment	Exposure ascertainment	Exposure categories
Tollefsen et al. (2007), Norway	PC	1,477	17-19 mixed	wheezing	child	trained	overweight: ≥85th centile
Tsai et al. (2007), Taiwan	CS	2,218	11-12 years mixed	asthma	child	e-records	underweight: <5th , overweight: 85th-95th and obese:≥95th centile
Vargas et al. (2007), USA	CC	2,053	3-5 years mixed	asthma	parents	e-records	overweight: 85th-95th and obese:≥95th centile
Garcia-Marcos et al. (2008), Spain	CS	874	6-8 years mixed	asthma	parent	trained	overweight: >85th centile
Jacobson et al. (2008), USA	CS	517	1-5 years mixed	asthma	parent	trained	overweight: 85th-95th and obese:≥95th centiles
Kusunoki et al. (2008), Japan	CS	45,520	7-15 years mixed	asthma	parent	e-records	overweight: 85th-95th and obese:≥95th centiles
Ahmad et al. (2009), USA	CS	63,981	7-17 years mixed	asthma	parent	parent	CDC: obese≥95th centiles
He et al. (2009), China	CS	2,179	8-13 years mixed	asthma	parent	no mention	IOTF cut-offs for underweight, overweight, and obese.
Kuschnir and da Cunha (2009), Brazil	CS	2,858	13-14 years mixed	asthma	trained	trained	underweight: <5th and overweight:≥85th centiles
Scholtens et al. (2009), Netherlands	PC	3,756	8 years mixed	asthma	parent	parent	overweight: ≥85th centile
Tai et al. (2009), Australia	CS	1,509	4-5 years mixed	asthma	trained	trained	IOTF cut-offs for overweight, and obese.
Tsai and Tsai (2009), Taiwan	CS	1,329	10-12 years mixed	asthma	child	child	underweight: <5th , overweight: 85th-95th and obese:≥95th centiles
Vazquez-Nava et al. (2010), Mexico	RC	1,160	4-5 years mixed	asthma	parent	trained	underweight: <5th , overweight: 85th-95th and obese:≥95th centiles

Author , year, region	Study design	Sample size	Participants age and gender	outcome term	Outcome ascertainment	Exposure ascertainment	Exposure categories
Visness et al. (2010), USA	CS	16,074	2-19 years mixed	asthma	parent and child	trained	overweight: 85^{th}-95^{th} and obese:$\geq 95^{th}$ centiles
Cibella et al. (2011), Italy	CS	708	10-16 years mixed	asthma	child	trained	overweight: $\geq 85^{th}$ centile
Matos et al. (2011), Brazil	PC	1,129	4-12 years mixed	asthma	parent	trained	overweight: 1SD+
Suglia et al. (2011),USA	PC	1,815	3 years mixed	asthma	parent	trained	underweight: $<5^{th}$, overweight: 85^{th}-95^{th} and obese:$\geq 95^{th}$ centiles
Yao et al. (2011),Taiwan	CS	5,351	4-18 years mixed	asthma	parent	trained	IOTF cut-offs for overweight, and obese.
Black et al. (2012) USA*	RC	681,122	6-19 years mixed	asthma	e-records	e-records	underweight: $<5^{th}$, overweight: 85^{th}-95^{th}, and obese:$\geq 95^{th}$ centiles
Magnusson et al. (2012), Sweden	PC	2,075	8 years mixed	asthma	parent	e-records	overweight: $\geq 85^{th}$ centile
Noal et al. (2012), Brazil	PC	4,441	11-15 years mixed	wheezing	parent	trained	overweight: 1SD+ and obese: 2SD+
Guibas et al. (2013), Greece	CS	1,626	2-5 years mixed	asthma	parent	trained	IOTF cut-offs for overweight, and obese.
Guibas et al. (2013), Greece	CS	2,015	9-13 years mixed	asthma	parent	trained	IOTF cut-offs for overweight, and obese.
Silva et al. (2013), Brazil	CS	1,500	6-12 years mixed	wheezing	parent	trained	overweight: $\geq 85^{th}$ centile
Yiallouros et al. (2013), Cyprus*	CS	10,981	7-17 years mixed	asthma	parent and child	trained	overweight:>1SD+
Granell et al. (2014), UK	PC	4,835	7-9 years mixed	asthma	parent	trained	IOTF cut-offs for overweight and obese.
Wang et al. (2014) China	CS	30,056	2-14 years mixed	asthma	parent	trained	underweight: $<5^{th}$, overweight: 85^{th}-95^{th} and obese:$\geq 95^{th}$ centiles

CC= Case-control, CS=Cross-sectional, PC=Prospective Cohort, RC=Retrospective Cohort, mixed=included both genders; *=regrouped

Table 2.12 Risk of bias assessment table using Newcastle-Ottawa quality assessment scale for BMI and wheezing disorder studies included in the meta-analysis

Author , year, region	Study title	Selection	Comparability	Outcome
Gennuso et al, 1998, USA	The relationship between asthma and obesity in urban minority children and adolescents	★★★★	★★	★★
Chinn et al, 2001, UK	Can the increase in body mass index explain the rising trend in asthma in children?	★★★★	★	★★
Rodríguez et al, 2002, USA	Identification of population subgroups of children and adolescents with high asthma prevalence: Findings from the third national health and nutrition examination survey	★★★	★★	★★
Gilliland et al, 2003, USA	Obesity and the risk of newly diagnosed asthma in school-age children	★★★★	★★	★
Bibi et al, 2004, Israel	The relationship between asthma and obesity in children: Is it real or a case of over diagnosis?	★★★	★	★★
Cassol et al, 2005, Brazil	Prevalence and severity of asthma among adolescents and their relationship with the body mass index	★★★	★	★★
Saha et al, 2005, USA	Individual and neighbourhood-level factors in predicting asthma	★★★★	★★	★★★
Wickens et al, 2005, New Zealnd	Obesity and asthma in 11-12 year old New Zealand children in 1989 and 2000	★★	★★	★★
Kwon et al, 2006, USA	Childhood asthma and extreme values of body mass index: The Harlem Children's Zone Asthma Initiative	★★	★★	★★
Shamssain et al, 2006, UK	The association between overweight and respiratory symptoms in schoolchildren	★★	★	★★

Author , year, region	Study title	Selection	Comparability	Outcome
Van De Ven et al, 2006, Netherlands	Atopic diseases and related risk factors among Dutch adolescents	★★	★★	★★
Davis et al, 2007, USA	An association between asthma and BMI in adolescents: results from the California Healthy Kids Survey	★★	★★	★★
Tollefsen et al, 2007, Norway	Female gender is associated with higher incidence and more stable respiratory symptoms during adolescence	★★★★	★★	★★
Tsai et al, 2007, Taiwan	Associations of BMI, TV-watching time, and physical activity on respiratory symptoms and asthma in 5th grade schoolchildren in Taipei, Taiwan	★★	★★	★★
Vargas et al, 2007, USA	Relationship of body mass index with asthma indicators in Head Start children	★★★	★★	★★
Garcia-Marcos et al, 2008, Spain	Percent body fat, skin-fold thickness or body mass index for defining obesity or overweight, as a risk factor for asthma in schoolchildren: which one to use in epidemiological studies?	★★★	★★	★
Jacobson et al, 2008, USA	Asthma, body mass, gender, and Hispanic national origin among 517 preschool children in New York City	★★★	★★	★
Kusunoki et al, 2008, Japan	Obesity and the prevalence of allergic diseases in schoolchildren	★★★	★★	★★
Ahmad et al, 2009, USA	Association between Obesity and asthma in US children and adolescents	★★	★★	★
He et al, 2009, China	Respiratory health in overweight and obese Chinese children	★★	★★	★★
Kuschnir 2009, Brazil	Association of overweight with asthma prevalence in adolescents in Rio de Janeiro, Brazil	★★★	★★	★★

Author , year, region	Study title	Selection	Comparability	Outcome
Scholtens et al, 2009, Netherlands	Overweight and changes in weight status during childhood in relation to asthma symptoms at 8 years of age	★★	★★	★★
Tai et al, 2009, Australia	Association between asthma symptoms and obesity in preschool (4-5 year old) children	★★★	★	★★
Tsai et al, 2009, Taiwan	The association of BMI and sedentary time with respiratory symptoms and asthma in 5th grade schoolchildren in Kaohsiung, Taiwan	★★	★★	★★
Vazquez-Nava et al, 2010, Mexico	Association between obesity and asthma in preschool Mexican children	★★	★	★★★
Visness et al, USA, 2010	Association of Childhood Obesity With Atopic and Nonatopic Asthma: Results From the National Health and Nutrition Examination Survey 1999–2006	★★★	★★	★★
Cibella et al, 2011, Italy	A cross-sectional study assessing the relationship between BMI, asthma, atopy, and eNO among schoolchildren	★★	★★	★
Matos et al, 2011, Brazil	Overweight, asthma symptoms, atopy and pulmonary function in children of 4-12 years of age: findings from the SCAALA cohort in Salvador, Bahia, Brazil	★★★★	★★	★
Suglia et al, 2011, USA	Asthma and obesity in three-year-old urban children: Role of sex and home environment	★★★	★★	★★
Yao et al, Taiwan, 2011	Associations of age, gender, and BMI with prevalence of allergic diseases in children: PATCH study	★★★	★★	★★
Black et al, 2012, USA	Higher prevalence of obesity among children with asthma	★★★★	★★	★★★
Magnusson et al, 2012, Sweden	Early childhood overweight and asthma and allergic sensitization at 8 years of age	★★★	★★	★

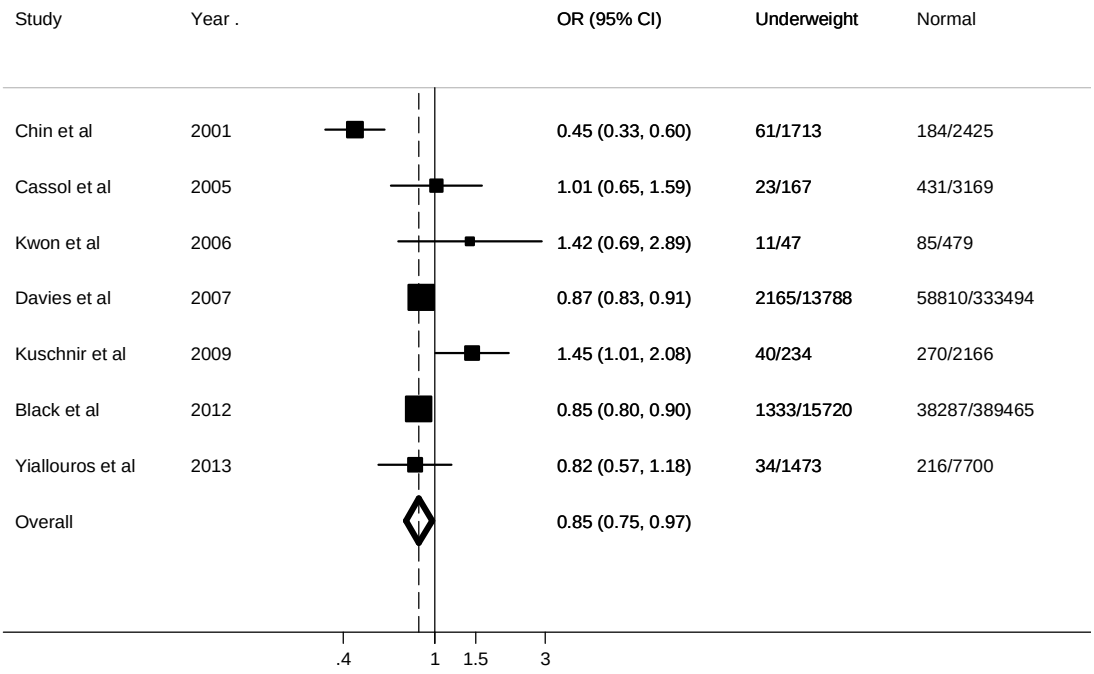
Author , year, region	Study title	Selection	Comparability	Outcome
Noal et al, 2012, Brazil	Is obesity a risk factor for wheezing among adolescents? A prospective study in southern Brazil	★★★★	★★	★★
Guibas et al, 2013, Greece	The obesity-asthma link in different ages and the role of Body Mass Index in its investigation: Findings from the Genesis and Healthy Growth Studies	★★★	★★	★★
Silva et al, 2013, Brazil	Prevalence of Wheezing and its Association with Body Mass Index and Abdominal Obesity in Children	★★★	★★	★★
Yiallourous, 2013, Cyprus	Associations of body fat percent and body mass index with childhood asthma by age and gender	★★★	★★	★★
Grannel et al, 2014, UK	Effects of BMI, Fat Mass, and Lean Mass on Asthma in Childhood: A Mendelian Randomization Study	★★★★	★★	★
Wang et al, 2014, China	Gender-specific differences in associations of overweight and obesity with asthma and asthma-related symptoms in 30,056 children: result from 25 districts of North-eastern China	★★★	★★	★★

2.4.3.3 Meta-analysis

Underweight

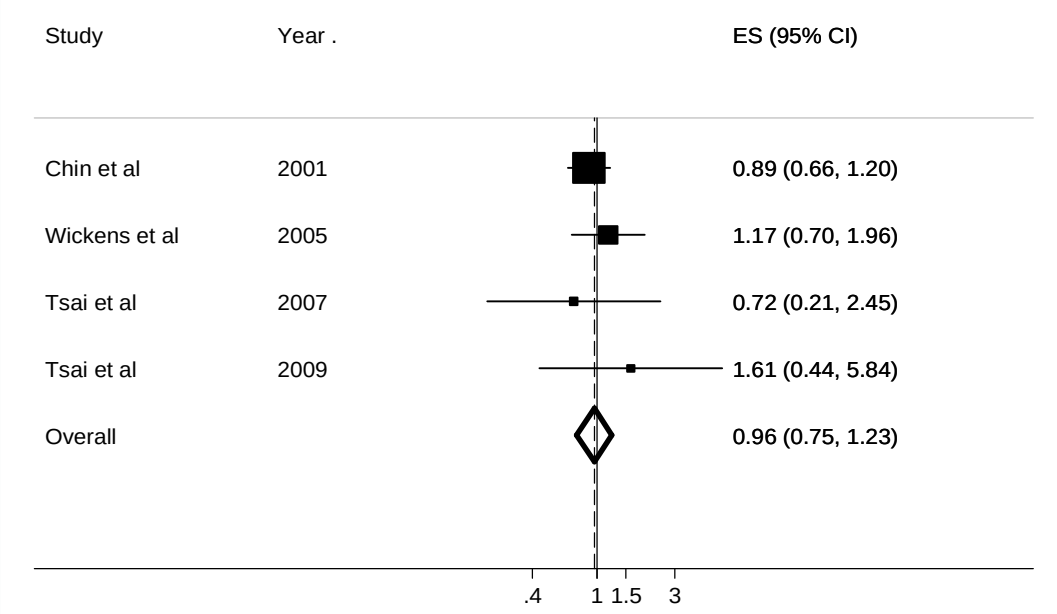
A total of 7 studies presented data on the number of cases and non-cases of wheezing disorders in underweight and normal BMI groups comprising a total of 772,040 children (Figure 2.8). The summary risk estimate of the studies showed that there was a significant decrease odds of wheezing disorders (OR= 0.85, 95% CI: 0.75 to 0.97; P=0.02) for the underweight children (Figure 2.8). However, there was considerable heterogeneity among the studies (Q=29, df =6, P<0.01; I^2=79%, 95% CI: 58% to 89%). When the same analysis was performed on four studies (Chinn and Rona, 2001; Wickens et al., 2005; Tsai et al., 2007; Tsai and Tsai, 2009) that provided adjusted risk estimates, the overall was odds ratio was 0.96 (95% CI: 0.75 to 1.23; P=0.75) with no heterogeneity among studies (Q=2, d.f=3, P=0.65; I^2=0%), see Figure 2.9.

Figure 2.8 Summary unadjusted odds ratios of wheezing disorders for underweight compared with normal BMI categories



Heterogeneity chi-squared = 29 (df = 6) p < 0.001, I^2 =79% (95% CI: 58% to 89%), and the estimate of between-study variance Tau-squared = 0.01.

Figure 2.9 Summary adjusted odds ratios of wheezing disorders for underweight compared with normal BMI categories

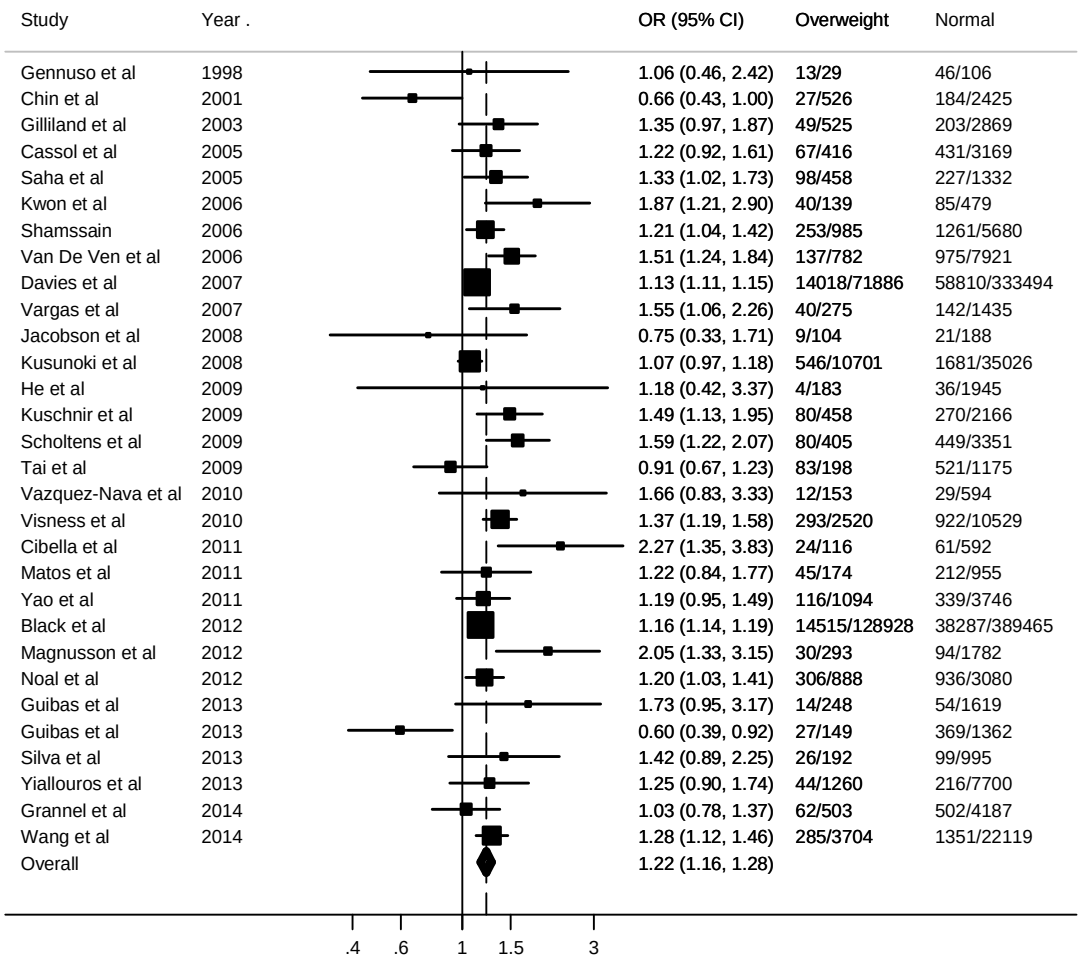


Heterogeneity chi-squared = 2 (df = 3) p = 0.65, I^2 =0%, and the estimate of between-study variance Tau-squared < 0.001.

Overweight

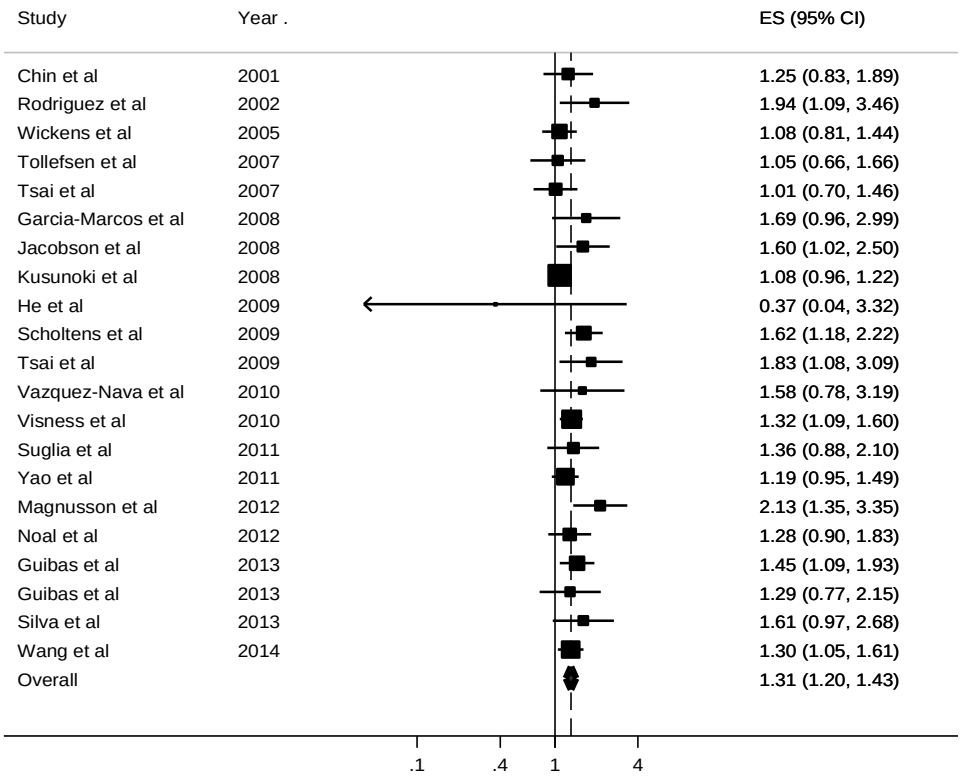
A total of 30 studies presented data on the number of cases and non-cases of wheezing disorders in the overweight and normal BMI groups that included a total of 1,079,732 children (Figure 2.10). The summary of the odds ratio showed that there was a significant increased risk of wheezing disorders (OR= 1.22, 95% CI: 1.16 to 1.28; p<0.001), see Figure 2.10. There was a considerable heterogeneity among the studies (Q=78, df = 29; I^2 = 63%, 95% CI: 45% to 75%). When further analysis was carried out on 21 studies that presented adjusted risk estimates of overweight on childhood wheezing disorders, the summary risk estimate was slightly accentuated (OR=1.31, 95% CI: 1.20 to 1.43, P<0.01) whereas the between-study heterogeneity substantially decreased (Q=27, df=20, P=0.12; I^2=27%, 95% CI: 0.0 to 57%), see Figure 2.11.

Figure 2.10 Summary unadjusted odds ratios of wheezing disorders for overweight compared with normal BMI categories



Heterogeneity chi-squared = 78 (df = 29) $p < 0.001$, I^2 = 63% (95% CI: 45% to 75%), and the estimate of between-study variance Tau-squared = 0.004.

Figure 2.11 Summary adjusted odds ratios of wheezing disorders for overweight compared with normal BMI categories

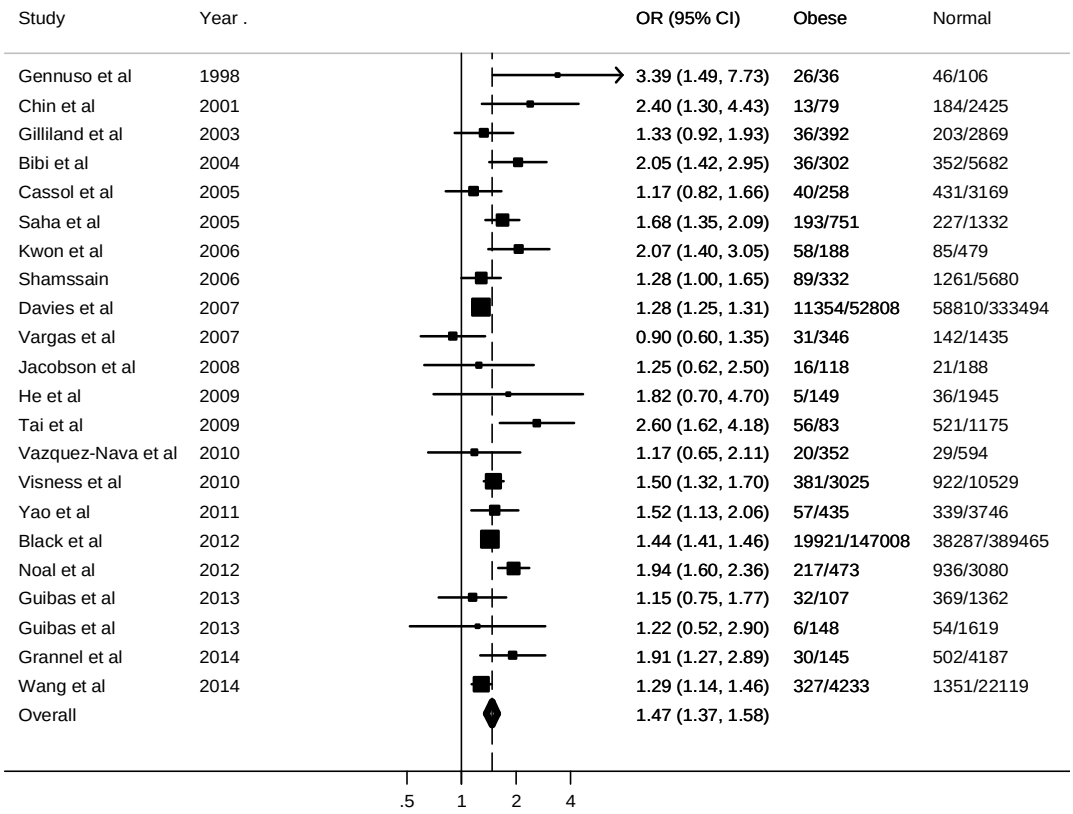


Heterogeneity chi-squared = 27 (df = 20) p = 0.13, I^2 = 26% (95% CI: 0% to 57%), and the estimate of between-study variance Tau-squared = 0.01.

Obesity

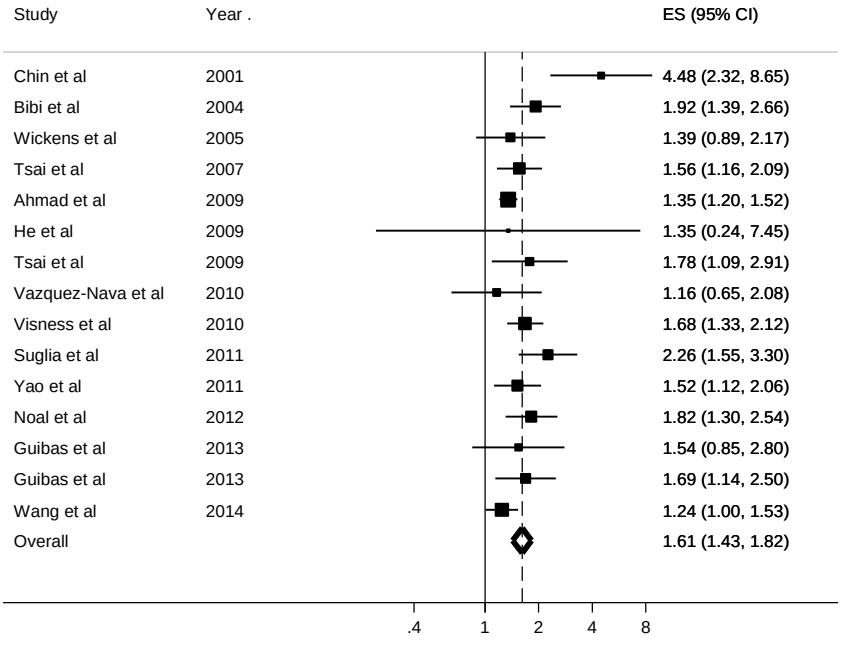
A total of 22 studies presented data on the number of cases and non-cases of wheezing disorders in the normal and obese groups that comprised 1,007,418 children (Figure 2.12). The overall risk estimate showed that there was a significant increase in the risk of wheezing disorders for obesity (OR=1.46, 95% CI: 1.36 to 1.57), see Figure 2.12. There was a substantial heterogeneity among the studies (Q= 113, df = 21; I^2 = 82%, 95% CI: 73% to 87%). However, when the analysis was repeated on the adjusted risk estimate of obesity on wheezing disorders available from 16 studies (Figure 2.13), the heterogeneity was attenuated (Q=28, d.f=14, P=0.02; I^2=49%, 95% CI: 8% to 72%) whilst the summary risk estimate slightly increased (OR=1.61, 95% CI: 1.43 to 1.82).

Figure 2.12 Summary unadjusted odds ratios of wheezing disorders for obese compared with normal BMI categories



Heterogeneity chi-squared = 113 (df = 21), $p < 0.001$, I^2 = 82% (95% CI: 73% to 87%), and the estimate of between-study variance Tau-squared = 0.009.

Figure 2.13 Summary adjusted odds ratios of wheezing disorders for obese BMI compared with normal BMI categories



Heterogeneity chi-squared = 28 (df = 14), p = 0.02, I^2 = 49% (95% CI: 8% to 72%), and the estimate of between-study variance Tau-squared = 0.02.

2.4.3.4 Subgroup analyses

Underweight and wheezing disorders

Subgroup meta-analyses of underweight risk estimates on childhood wheezing disorders from 7 studies showed that the strength of the risk estimates remained stable across each subgroup of the predefined covariates or study characteristics. However, the risk of wheezing disorders significantly reduced if BMI was categorised using CDC or IOTF, age group was 'Five years and above', sample size was >1000, three BMI categories were used during analysis. No subgroup analysis was conducted for adjusted odds ratios as the studies contributed were few.

The heterogeneities within each subgroup of the covariates were significant; and, except for the covariate 'exposure categorisation method', the heterogeneities between subgroups of the covariates were not significant (Table 2.13).

Table 2.13 Subgroup analysis of unadjusted odds ratios of wheezing disorders for underweight compared with normal BMI categories

	Odds ratio (95% CI)	n	I^2	P_{het} [a]	P_{het} [b]
Outcome ascertainment					
E-records/trained	1.08(0.64 to 1.8)	2	88%	<0.01	0.06
Child	0.87 (0.83 to 0.91)	1	-	-	
Parent	0.81 (0.50 to1.31)	4	81%	<0.01	
Exposure ascertainment					
E-records/trained	0.84 (0.62 to 1.15)	5	85%	<0.001	0.27
Child	0.87 (0.83 to 0.91)	1	-	-	
No mention	1.42 (0.69 to 2.89)	1	-	-	
Exposure categorisation method					
CDC	0.90 (0.82 to 0.97)	5	62%	0.03	<0.001
IOTF	0.45 (0.33 to 0.61)	1	-	-	
WHO	0.82 (0.57 to 1.18)	1	-	-	
Age during diagnosis					
Five years and above	0.84 (0.74 to 0.96)	6	82%	<0.001	0.17
Mixed (0-19 years)	1.42 (0.69 to 2.90)	1	-	-	
Sample size					
<1000	1.42 (0.69 to 2.89)	1	-	-	0.17
1000+	0.84 (0.74 to 0.96)	6	82%	<0.001	
Study Design					
cohort	0.62 (0.34 to 1.17)	2	94%	<0.001	0.12
Cross-sectional	0.85 (0.75 to 0.97)	5	59%	0.05	
Number of BMI categories					
Three	1.09 (0.62 to 1.91)	2	79%	0.03	0.07
Four	0.81 (0.71 to 0.93	5	81%	<0.01	
Study quality score[c]					
High (7-9/9)	0.89 (0.66 to 1.19)	6	83%	<0.01	0.44
Medium (5-6/9)	0.87 (0.83 to 0.91)	1	-	-	

[a] P for heterogeneity within subgroups.

[b] P for heterogeneity between subgroups.

[c] Sensitivity analysis.

Overweight and wheezing disorders

When subgroup meta-analysis of the 30 studies that presented wheezing disorder risks of overweight by the predefined study characteristics was carried out, except for the IOTF categorisation method and papers published before 2000 that were not statistically significant, the strength and direction of the summary risk estimates in each subgroup remained stable. The within subgroup heterogeneity was not significant for the wheezing outcome term, e-records outcome ascertainment, parental exposure ascertainment, WHO BMI categorisation method, sample size less than 1000, and case-control study design while it was significant for the rest of the subgroups. Except for outcome and exposure ascertainment, age group during diagnosis, sample size, and study design subgroups, there was no significant heterogeneity between subgroups of the other covariates (Table 2.14).

Subgroup analyses of 21 studies that presented adjusted odds ratio of wheezing disorders for overweight, except if outcome was ascertained by a child, WHO BMI categorisation was used, 'under five' age group, or four BMI categories were used during analysis, there was a significant increase of the risk across all subgroups of the *a priori* selected study characteristics. The between subgroup heterogeneities were all in insignificant (Table 2.15).

Table 2.14 Subgroup analysis of unadjusted odds ratios of wheezing disorders for overweight compared with normal BMI categories

	Odds ratio (95% CI)	n	I^2	P_{het} [a]	P_{het} [b]
Outcome terms used					
Asthma	1.22 (1.16 to 1.28)	28	65%	<0.01	0.44
wheezing	1.23 (1.05 to 1.42)	2	0%	0.5	
Outcome ascertainment					
E-records/trained	1.19 (1.05 to 1.36)	5	40%	0.16	0.01
Child	1.30 (1.05 to 1.60)	4	64%	0.04	
Parent	1.25 (1.14 to 1.38)	21	64%	<0.01	
Exposure ascertainment					
E-records/trained	1.21 (1.13 to 1.30)	25	58%	<0.01	<0.01
child	1.13 (1.11 to 1.15)	1	-	-	
Parent	1.54 (1.32 to 1.80)	2	0%	0.77	
No mention	1.38 (1.01 to 1.88)	2	42%	0.18	
Exposure categorisation method					
CDC	1.24 (1.17 to 1.30)	18	66%	<0.01	0.76
IOTF	1.06 (0.88 to 1.29)	9	72%	<0.01	
WHO	1.21 (1.06 to 1.39)	3	0%	0.98	
Age during diagnosis					
Five years and above	1.21 (1.15 to 1.27)	18	65%	<0.01	0.03
Mixed (0-19 years)	1.21 (1.07 to 1.37)	12	55%	0.01	
Sample size					
<1000	1.57 (1.10 to 2.23)	5	37%	0.17	0.01
1000+	1.21 (1.15 to 1.27)	25	63%	<0.01	
Study period					
<2000	1.06 (0.46 to 2.42)	1	-	-	0.84
2000+	1.22 (1.17 to 1.28)	29	64%	<0.01	
Study design					
Cohort	1.23 (1.08 to 1.41)	9	67%	<0.01	
Case-control	1.45 (1.03 to 2.05)	2	0%	0.41	0.17
Cross-sectional	1.24 (1.15 to 1.34)	19	66%	<0.01	
Number of BMI Categories					
Two	1.48 (1.19 to 1.84)	7	78%	0.07	
Three	1.23 (1.13 to 1.33)	18	35%	<0.01	0.01
Four	1.15 (1.09 to 1.21)	5	74%	<0.01	
Study quality score [c]					
High (7-9/9)	1.21 (1.13 to 1.30)	20	55%	<0.01	0.07
Medium (5-6/9)	1.34 (1.15 to 1.56)	10	73%	<0.01	

[a] P for heterogeneity within subgroups.
[b] P for heterogeneity between subgroups.
[c] Sensitivity analysis.

Table 2.15 Subgroup analysis of adjusted odds ratios of wheezing disorders for overweight compared with normal BMI categories

	Odds ratio (95% CI)	n	I^2	P_{het} [a]	P_{het} [b]
Outcome terms used					
Asthma	1.32 (1.20 to 1.45)	18	33%	0.08	0.90
wheezing	1.28 (1.00 to 1.63)	3	0%	0.47	
Outcome ascertainment					
Child	1.20 (0.85 to 1.70)	3	45%	0.16	0.58
Parent	1.32 (1.21 to 1.45)	18	27%	0.14	
Exposure ascertainment					
E-records/trained	1.27(1.17 to 1.38)	18	19%	0.22	0.13
Parent	1.62(1.18 to 2.22)	1	-	-	
No mention	1.20 (0.28 to 5.24)	2	51%	0.15	
Exposure categorisation method					
CDC	1.38 (1.22 to 1.56)	14	45%	0.03	0.92
IOTF	1.23 (1.07 to 1.40)	6	0%	0.65	
WHO	1.28 (0.90 to 1.83)	1	-	-	
Age during diagnosis					
Five years and above	1.30 (1.12 to 1.52)	11	52%	0.02	0.35
Mixed (0-19 years)	1.33 (1.20 to 1.47)	9	0%	0.83	
Under five	1.36 (0.88 to 2.10)	1	-	-	
Sample size					
<1000	1.59 (1.09 to 2.32)	2	0%	0.98	0.21
1000+	1.30 (1.18 to 1.42)	19	30%	0.11	
Study design					
Cohort	1.43 (1.22 to 1.68)	7	4%	0.40	
Case-control	1.27 (1.15 to 1.39)	14	27%	0.17	0.07
Number of BMI Categories					
Two	1.48 (1.16 to 1.89)	7	67%	0.01	
Three	1.32 (1.19 to 1.45)	10	0%	0.94	0.42
Four	1.18 (0.96 to 1.47)	4	21%	0.28	
Study quality score [c]					
High (7-9/9)	1.20 (1.12 to 1.29)	14	0%	0.51	<0.01
Medium (5-6/9)	1.71 (1.42 to 2.05)	7	0%	0.80	

[a] P for heterogeneity within subgroups.
[b] P for heterogeneity between subgroups.
[c] Sensitivity analysis.

Obesity and wheezing disorders

Subgroup meta-analysis of 22 studies that presented the number of wheezing disorders cases and non-cases in the obese and normal BMI groups showed that except for the case-control study design subgroup, significant increase of risk was observed in all subgroups of the *a priori* selected study characteristics (Table 2.16). The same analysis on 15 studies that presented adjusted odds ratios also showed that except in the unknown BMI measurement source (no mention) and sample size of <1000 groups, there was a significant increase of wheezing disorders risk across all subgroups of the predefined study characteristics (Table 2.17).

In the unadjusted odds ratios analyses, the within subgroup heterogeneities were significant except for the outcome ascertainment through a child, exposure ascertainment not mentioned, IOTF BMI categorisation method, and sample size less than 1000 subgroups; and, the between subgroup heterogeneities were significant except for the sample size covariate (Table 2.16). In addition, except for the type of study design, no significant heterogeneity between subgroups was observed in the adjusted OR analyses results (Table 2.17)

Table 2.16 Subgroup analysis of unadjusted odds ratios of wheezing disorders for obese compared with normal BMI categories

	Odds ratio (95% CI)	n	I^2	P_{het} [a]	P_{het} [b]
Outcome terms used					
Asthma	1.44 (1.34 to 1.55)	21	80%	<0.01	0.01
Wheezing	1.94 (1.60 to 2.36)	1	-	-	
Outcome ascertainment					
E-records/trained	1.81 (1.39 to 2.36)	4	75%	0.01	<0.01
Child	1.28 (1.25 to 1.31)	3	0%	0.86	
Parent	1.51 (1.33 to 1.71)	15	52%	0.01	
Exposure ascertainment					
E-records/trained	1.52 (1.38 to 1.66)	18	59%	<0.01	<0.01
Child	1.28 (1.25 to 1.31)	1	0%	-	
No mention	1.60 (1.11 to 2.31)	3	53%	0.12	
Exposure categorisation method					
CDC	1.41 (1.30 to 1.53)	13	86%	<0.01	<0.01
IOTF	1.61 (1.31 to 1.99)	8	42%	0.10	
WHO	1.94 (1.60 to 2.36)	1	-	-	
Age during diagnosis					
Five years and above	1.43 (1.31 to 1.57)	11	69%	<0.01	0.05
Mixed (0-19 years)	1.53 (1.30 to 1.80)	11	87%	<0.01	
Sample size					
<1000	1.75 (1.15 to 2.65)	4	49%	0.12	0.07
1000+	1.46 (1.35 to 1.57)	18	84%	<0.01	
Study period					
<2000	3.39 (1.49 to 7.73)	1	-	-	0.03
2000+	1.47 (1.37 to 1.58)	21	82%	<0.01	
Study design					
Cohort	1.62 (1.35 to 1.95)	6	64%	0.02	
Case-control	1.65 (0.45 to 6.1)	2	88%	0.01	<0.01
Cross-sectional	1.46 (1.36 to 1.57)	14	62%	<0.01	
Number of BMI categories					
Two	2.05 (1.42 to 2.95)	1	0%	-	
Three	1.50 (1.33 to 1.69)	16	56%	<0.01	0.01
Four	1.40 (1.26 to 1.56)	5	94%	<0.01	
Study quality score [c]					
High (7-9/9)	1.47 (1.30 to 1.67)	16	57%	<0.01	0.02
Low (≤4/9)	1.40 (1.26 to 1.56)	6	94%	<0.01	

[a] P for heterogeneity within subgroups.

[b] P for heterogeneity between subgroups.

[c] Sensitivity analysis.

Table 2.17 Subgroup analysis of adjusted odds ratio of wheezing disorders for obese compared with normal BMI categories

	Odds ratio (95% CI)	n	I^2	P_{het} [a]	P_{het} [b]
Outcome terms used					
Asthma	1.60 (1.41 to 1.82)	14	50%	0.02	0.25
Wheezing	1.82 (1.30 to 2.54)	1	-	-	
Outcome ascertainment					
Child	1.61 (1.26 to 2.08)	2	0%	0.65	0.57
Parent	1.62 (1.41 to 1.86)	13	55%	0.01	
Exposure ascertainment					
E-records/trained	1.67 (1.46 to 1.91)	13	46%	0.04	0.07
Parent	1.35 (1.20 to 1.52)	1	-	-	
No mention	1.35 (0.24 to 7.45)	1	-	-	
Exposure categorisation method					
CDC	1.55 (1.34 to 1.78)	8	52%	0.04	0.22
IOTF	1.76 (1.31 to 2.36)	6	49%	0.08	
WHO	1.82 (1.30 to 2.54)	1	-	-	
Age during diagnosis					
Five years and above	1.69 (1.38 to 2.07)	8	58%	0.02	0.10
Under five	2.26 (1.55 to 3.30)	1	-	-	
Mixed (0-19 years)	1.49(1.29 to 1.71)	6	17%	0.30	
Sample size					
<1000	1.16 (0.65 to 2.08)	1	-	-	0.38
1000+	1.63 (1.44 to 1.85)	14	51%	0.01	
Study design					
Cohort	2.06 (1.38 to 3.18)	4	69%	0.02	<0.01
Cross-sectional	1.45 (1.34 to 1.57)	11	0%	0.48	
Number of BMI categories					
Two	1.56 (1.11 to 2.18)	2	75%	0.05	
Three	1.57 (1.37 to 1.81)	9	25%	0.22	0.18
Four	1.90 (1.28 to 2.84)	4	68%	0.02	
Study quality score [c]					
High (7-9/9)	1.69 (1.46 to 1.96)	11	52%	0.02	0.02
Medium (5-6/9)	1.36 (1.22 to 1.52)	4	0%	0.69	

[a] P for heterogeneity within subgroups.

[b] P for heterogeneity between subgroups.

[c] Sensitivity analysis.

2.4.3.5 Meta-regression analysis

Body mass index and wheezing disorders

Investigating the sources of between-study heterogeneity in the unadjusted odds ratios of wheezing disorders for the overweight BMI category showed that none of the between-study heterogeneity was explained by the *a priori* selected covariates (adjusted R-squared=0%, P=0.40), see Table 2.18. When the same analysis was carried out for the obese group, no significant proportion of the between-study heterogeneity in the adjusted and unadjusted odds ratios (P=0.57 and P=0.19, respectively) were explained by the *a priori* selected covariates or study characteristics (Table 2.19). No meta-regression analysis was carried out for the unadjusted underweight risk estimates on childhood wheezing disorders due to not having enough observations for the models to converge.

Table 2.18 Meta-regression analysis of unadjusted odds ratios of wheezing disorders for overweight compared with normal BMI categories

Study characteristic	Coefficient (95% CI)	P-value
Outcome terms used (ref=asthma)	0.12 (-0.28 to 0.52)	0.56
Outcome ascertainment (ref= e-records/trained)	-0.03 (-0.18 to 0.13)	0.74
Exposure ascertainment (ref=e-records/trained)	0.05 (-0.08 to 0.16)	0.34
Exposure categorisation method (ref=CDC)	-0.11 (-0.27 to 0.05)	0.18
Age during diagnosis (ref=five and above)	-0.02 (-0.26 to 0.22)	0.88
Sample size (ref=less than 1000)	-0.27 (-0.70 to 0.14)	0.19
Study period (ref=before 2000)	-0.49 (-1.60 to 0.64)	0.38
Study type (ref=cohort)	-0.03 (-0.14 to 0.08)	0.58
Number of weight categories (ref=two)	-014 (-0.30 to 0.12)	0.08
Overall (adjusted R-squared= 0%)		0.40

Table 2.19 Meta-regression analysis adjusted and unadjusted odds ratios of wheezing disorders for obese compared with normal BMI categories

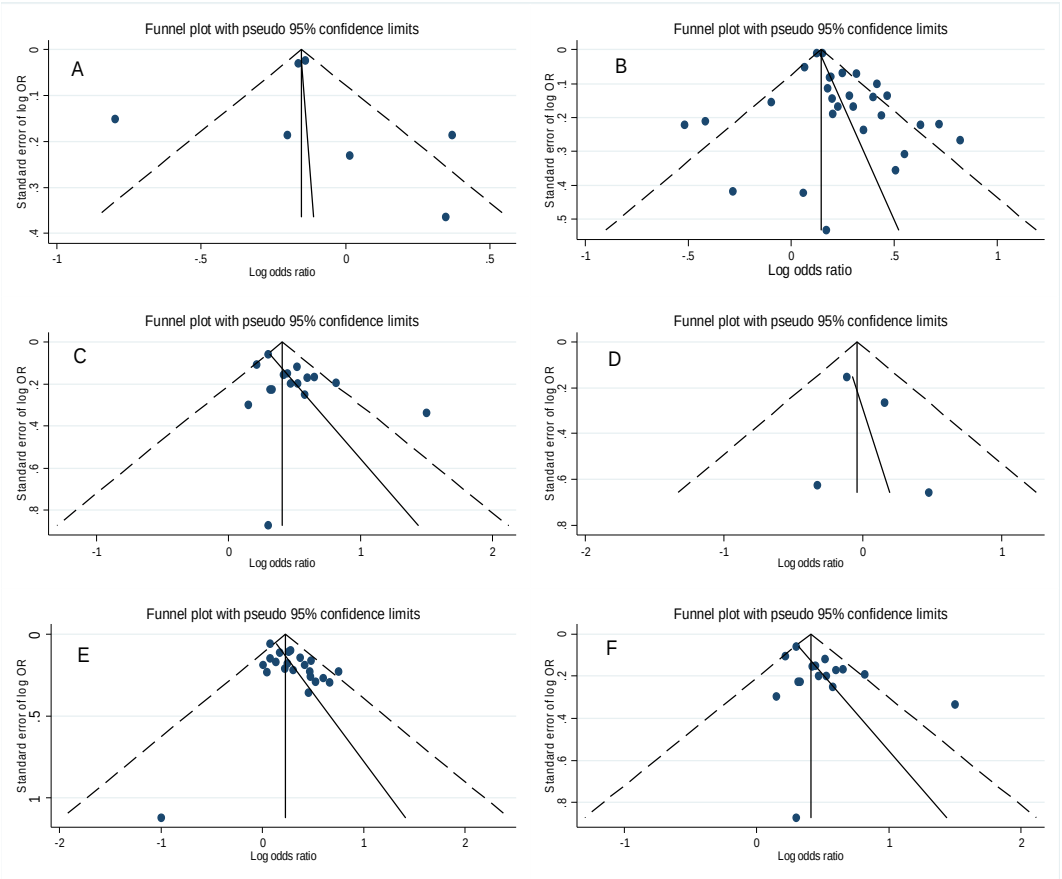
Study characteristics	Coefficient (95% CI)	P-value
Unadjusted odds ratios		
Outcome terms used (ref=asthma)	0.18 (-0.66 to 1.02)	0.65
Outcome ascertainment (ref= e-records/trained)	-0.14 (-0.37 to 0.09)	0.21
Exposure ascertainment (ref=e-records/trained)	0.004 (-0.19 to 0.20)	0.96
Exposure categorisation method (ref=CDC)	0.16 (-0.20 to 0.52)	0.35
Age during diagnosis (ref=five and above)	-0.02 (-0.42 to 0.38)	0.91
Sample size (ref=less than 1000)	-0.28 (-0.90 to 0.35)	0.35
Study period (ref=before 2000)	0.46 (-0.82 to 1.73)	0.45
Study type (ref=cohort)	0.04 (-0.16 to 0.24)	0.65
Number of weight categories (ref=two)	-0.11 (-0.41 to 0.19)	43
Overall (adjusted R-squared= 0%)		0.57
Adjusted odds ratios		
Outcome terms used (ref=asthma)	-0.62 (-1.62 to 0.38)	0.18
Outcome ascertainment (ref= e-records/trained)	-0.06 (-0.81 to 0.69)	0.85
Exposure ascertainment (ref=e-records/trained)	-0.17 (-0.41 to 0.07)	0.13
Exposure categorisation method (ref=CDC)	0.05 (-0.38 to 0.48)	0.78
Age during diagnosis (ref=five and above)	-0.21 (-0.53 to 0.12)	0.17
Sample size (ref=less than 1000)	0.96(0.06to 1.87)	0.04
Study type (ref=cohort)	-0.39 (-0.66 to -0.12)	0.01
Number of BMI categories (ref=two)	-0.11 (-0.52 to 0.29)	0.51
Overall (adjusted R-squared= 10%)		0.19

2.4.3.6 Investigating biases (small study effects)

Body mass index and wheezing disorders

The funnel plot and bias test results of adjusted and unadjusted risks of wheezing disorders for different BMI were not consistent. The unadjusted risk estimates of overweight (P=0.05), showed some evidence of asymmetry but not in the underweight and obese categories (P=0.92 and P=0.31, respectively), see Figure 2.14 and Table 2.20. However, the same analysis for the adjusted risk estimates showed some evidence of asymmetry for the overweight (P=0.02) and obese (0.04) but not in the underweight (P=0.57), see Figure 2.14 and Table 2.20.

Figure 2.14 Egger's funnel plots of BMI and childhood wheezing disorder studies



Underweight versus normal BMI (A and D); overweight versus normal BMI (B and E) and obesity versus normal BMI (C and F) odds ratio funnel plots. Unadjusted odds ratios in A, B and C, and adjusted odds ratios in D, E and F. In all funnel plots, the middle solid line is the summary odds ratio estimate and the two diagonal dotted lines are the 95% confidence limits around the summary odds ratio, and the slant solid lines in all figures are the fitted regression lines for Egger's small-study effect test.

Table 2.20 Egger's test of bias for small study effects in BMI and wheezing disorders studies

Parameter	Coefficient (95% CI)	P-value
Unadjusted odds ratios for normal versus underweight		
Slope	-0.16 (-0.31 to -0.001)	0.05
Bias	0.13 (-3.10 to 3.35)	0.92
Adjusted odds ratios for normal versus underweight		
Slope	-0.15 (-1.01 to 0.70)	0.52
Bias	0.52 (-2.86 to 3.91)	0.57
Unadjusted odds ratios for normal versus overweight		
Slope	0.13 (0.10 to 0.16)	<0.01
Bias	0.72 (0.01 to 1.43)	0.05
Adjusted odds ratios for normal versus overweight		
Slope	0.07 (-0.07 to 0.22)	0.31
Bias	1.14 (0.18 to 2.18)	0.02
Unadjusted odds ratios for normal versus obese		
Slope	0.31 (0.27 to 0.35)	<0.01
Bias	0.65 (-0.64 to 1.94)	0.30
Adjusted odds ratios for normal versus obese		
Slope	0.23 (0.04 to 0.43)	0.02
Bias	1.42 (-0.73 to 0.73)	0.04

2.4.4 Discussion

2.4.4.1 Key findings

The results of unadjusted and adjusted risk estimates for underweight children are inconclusive, that is, a significant (unadjusted OR= 0.85, 95% CI: 0.75 to 0.97) and an insignificant (adjusted OR=0.96, 95% CI: 0.75 to 1.23; P=0.75) decreased risk of wheezing disorders for the unadjusted and adjusted summary estimates respectively. However, overweight (unadjusted OR= 1.22, 95% CI: 1.16 to 1.28 and adjusted OR=1.31, 95% CI: 1.20 to 1.43) and obese (unadjusted OR =1.46, 95% CI: 1.36 to 1.57 and adjusted OR =1.46, 95% CI: 1.36 to 1.57) children have an increased risk of wheezing disorders when compared with the normal BMI children although there was a significant between-study heterogeneity and some evidence of small study effects or publication bias.

2.4.4.2 Results in context of previous reviews and meta-analyses

If meta-analysis was restricted to cohort studies as per Chen et al. (2013) and Egan et al. (2013) , the summary relative risk estimates for overweight and obesity are 1.21 (95% CI: 1.08 to 1.36) and 1.42 (1.31 to 1.54), respectively. However, the summary relative risk estimates for only cohort studies may not be comparable to that of Egan et al. (2013) as the risk estimate definition was not consistent across the studies included in their meta-analysis. The overweight summary relative risk estimates for only cohort studies and that reported by Flaherman and Rutherford (2006) meta-analysis may also not be comparable for the same reasons.

One notable difference between this and the three previous meta-analyses results is that the summary risk estimates of this meta-analysis have narrower confidence intervals and are higher than those previously reported. This is likely to be due to the larger number of participants in this meta-analysis because of data harmonisation, consistent definition of the risk estimates and BMI categorisation methods were used.

Based on the subgroup meta-analyses of the unadjusted risk estimate results, it can be noted that the summary ORs estimates tended to attenuate as the number of BMI categories used by study authors increased. For example, the summary associated risk of overweight on wheezing disorders for authors that used two BMI categories was twice and three times of those which used three and four BMI categories

respectively (Table 2.14). A similar pattern was also observed in the obesity risk estimates according to the number of BMI categories used by authors (Table 2.16).

The subgroup meta-analyses by study design also showed that the summary risk estimates of the cohort and cross-sectional studies are very similar, both for the overweight and obese BMI categories. This may indicate that cross-sectional studies can be as credible as cohort studies although the findings need to be validated by other meta-analyses in other fields or with more data included. Cross-sectional studies are easier and cheaper to conduct than case-control and cohort studies, and this can have implication for cost saving and efficiency.

Based on the heterogeneity measures (Q-test and I^2), it can be noted that there was a considerable level of between-study variation in the underweight, overweight and obesity unadjusted risk estimates although this could also be due to the studies included in this meta-analysis had high sample size (Rücker et al., 2008). As illustrated in the forest plots, there were a few studies with large samples and high precision of risk estimates that can have dominating effects for the between-study heterogeneities (Figure 2.10 and Figure 2.11). However, except for the obese group, the same pattern was not observed in the adjusted risk estimates: the between-study heterogeneities were low in underweight and overweight risk estimates.

2.4.4.3 Strengths and weaknesses

This work has limitations and results should be interpreted cautiously. First, in the low birthweight and overweight summary risk estimates, there was a significant and substantial level of between-study variation that was not explained by the *a priori* selected covariates. Second, there is also some evidence of funnel plot asymmetry which may indicate a potential small study effect such as potential publication bias (Egger et al., 1997). Third, as in any systematic review and meta-analysis, a possibility of potentially relevant studies being missed cannot be ruled out. Fourth, the results are based on epidemiologic observational studies and are solely dependent on the quality of the primary studies included.

The strength of this work is that it was possible to produce consistent risk estimates due to the use of harmonised data. Combining adjusted risk estimates was a primary choice among previous authors. This technique may, however, under or overestimate the association between exposure and outcome variables due to

exclusion of studies that used non-standard weight categories or combining all irrespective of the type of exposure categorisation method used. In order to improve validity of the summary risk estimates, data harmonisation techniques were implemented and more studies were included than if previous authors' techniques were used. The other strength of this work is also that it was possible to extract and analyse both adjusted and unadjusted risk estimates, which can be used as an internal validation to each other.

In conclusion, the results suggest that underweight may be associated with reduced odds of childhood wheezing disorders. Overweight and obese children have increased odds of wheezing disorders. However, although the findings assert that overweight/obesity and childhood wheezing disorders are associated, the causality or their temporal relationship deserves further investigation.

2.5 Growth patterns and wheezing disorders

2.5.1 Critique of past epidemiologic studies

The effect of early childhood growth on wheezing disorders has not been widely studied. Results from a handful of previous studies are inconsistent with some suggesting fast growth predisposes to wheezing disorders (Mamun et al., 2007; Scholtens et al., 2009; Pike et al., 2010; Zhang et al., 2010; Flexeder et al., 2012; Magnusson et al., 2012; van der Gugten et al., 2012; Anderson et al., 2013; Rzehak et al., 2013; Sonnenschein-van der Voort et al., 2014b; Magnus et al., 2015) and others reporting reduced risk of wheezing disorders (Mai et al., 2005; Sonnenschein-van der Voort et al., 2012; Sonnenschein-van der Voort et al., 2014b; De Korte-De Boer et al., 2015).

In addition, all of these studies, with the exception of one (Rzehak et al., 2013), assumed homogenous growth among children, either used statistical techniques that can now be improved on or a non-standard growth data analysis that makes comparison and replication of results very difficult. For example, three (Mamun et al., 2007; van der Gugten et al., 2012; Magnus et al., 2015) used data-driven standardised scores (SDS), four (Scholtens et al., 2009; Sonnenschein-van der Voort et al., 2012; Sonnenschein-van der Voort et al., 2014b; De Korte-De Boer et al., 2015) used country specific SDS and another one (Flexeder et al., 2012) used non-standardised weight measurements.

2.5.2 Sythematic literature review methods

2.5.2.1 Literature Search Strategy

The reviews were carried out in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009). Online searches were carried out using the EMBASE and MEDLINE databases. Table 2.21 gives the details of terms and phrases used for the literature search.

2.5.2.2 Inclusion criteria

Eligible papers were those published as an article, in English, and reported original research about the effects of childhood growth patterns on wheezing disorders in children 0-19 years of age. The literature search included publications until March 2015.

Table 2.21 Terms and phrases used during literature search

	Childhood growth
1	Growth
2	Growth trajectory
3	Child growth
4	Childhood growth
5	Growth pattern*
6	Weight change
7	BMI change
8	Weight increase
9	Weight gain
10	BMI increase
11	BMI gain
12	Slow growth
13	Fast growth
14	Rapid growth
15	Wheezing disorders
16	Wheez*
17	Childhood asthma
18	Asthm*
19	1-14/or
20	15-18/or
21	19-20/and
22	Limit 21 to English language

2.5.2.3 Data extraction

For the studies that were eligible to be included in the meta-analyses, the following characteristics were extracted:

a) Authors' name;
b) Year of publication;
c) Country of study;
d) Sample size;
e) Study age group and gender;
f) Diagnosis (outcome) terms;
g) Exposure terms used;
h) Outcome and exposure ascertainment methods; and,
i) Risk estimates.

2.5.2.4 Data standardisation

Exposure variable

No data standardisation was carried out for studies that investigated the effects of childhood growth on wheezing disorders as there was no loss of information due to growth data categorisation.

Outcome variables

Study authors used one or multiple outcome terms in their reporting. Again, for comparability among studies, where authors used a single outcome, for example, asthma or wheezing, the quoted outcome term by the author and its risk estimate were assumed for analysis. However, where authors used multiple outcome terms, the term that was highest in the hierarchy and its risk estimate were assumed for analysis. For example, if asthma and wheezing were used together, asthma was preferred over wheezing.

2.5.2.5 Quality assessment

Papers included in the review and meta-analysis were assessed for risks of bias using the Newcastle-Ottawa quality assessment scale (Wells et al., 2000), see Appendix A for details on scoring guideline used.

2.5.2.6 Statistical analysis

Studies were too few and diverse to be combined in a meta-analysis. Thus, no statistical analysis was planned to produce a quantitative summary of wheezing disorders risk estimates in relation to childhood growth patterns.

2.5.3 Results

2.5.3.1 Literature search

The search yielded 2,115 studies with 1,741 of whom screened for eligibility. Eighteen studies were read in full resulting 15 studies to be included in the review (Figure 2.15).

The included studies were from Europe (13), America (1) and Oceania (1), see Table 2.22. Results of these studies were not combined to form a summary estimate through a meta-analysis. Hence, only a descriptive summary has been presented (Table 2.24). One study was not included in the descriptive table because the age range when a second weight measurement occurred was not recorded in the text or table (Sonnenschein-van der Voort et al., 2014a).

2.5.3.2 Quality of studies

Out of 14 studies, eight scored 7/9, five scored 6/9 and another one study scored 4/9 (Table 2.23)

Figure 2.15 Childhood growth patterns and wheezing disorders literature search flow chart

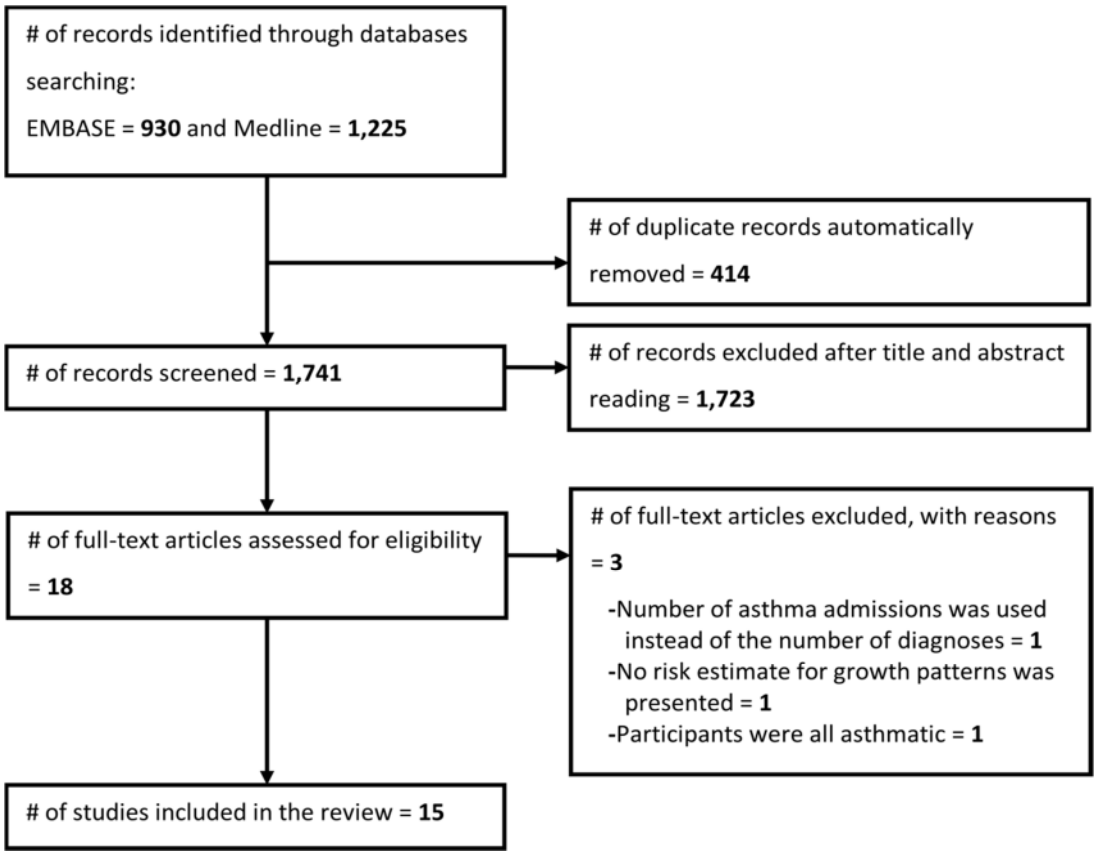


Table 2.22 Characteristics of childhood growth patterns and wheezing disorder studies included in the review

Author , year, region	Sample size	Participants' age and gender	Outcome term	Outcome ascertainment	Exposure ascertainment	Exposure terms used
Mai et al. (2005), Sweden	74	12 years mixed	asthma	parent	no mention	1 weight SDS increase between 0 and 6 months
Mamun et al. (2007), Australia	3,759	14 years mixed	asthma	parent	trained	1 BMI SDS increase between 5 and 14 years
Scholtens et al. (2009), Netherlands	3,756	8 years mixed	asthma	parent	parent	Changes in BMI status between 1 and 7 years
Pike et al. (2010), UK	1,548	3 years mixed	wheezing	parent	trained	1 weight SDS increase between 0 and 12 months
Zhang et al. (2010), USA	285	8 years mixed	asthma	e-records	e-records	0.67 weight SDS increase between 0 and 6 months
Flexeder et al. (2012), Germany	9,086	10 years mixed	asthma	parent	trained	Pick weight velocities at 4, 6 or 10 years
Magnusson et al. (2012), Sweden	2,075	12 years mixed	asthma	parent	trained	BMI changes between 1 and 7 years
Sonnenschein-van der Voort et al. (2012), Netherlands	5,125	1-4 years mixed	wheezing	parent	e-records	0.67 SDS increase between 3 and 12 months
van der Gugten et al. (2012), Netherlands	1,431	1 years mixed	wheezing	parent	no mention	0.67 SDS increase during first 3 months
Anderson et al. (2013), Belarus	12,171	6 years mixed	wheezing	parent	trained	Weight velocities between 0 and 60 months of age.
Rzehak et al. (2013), Multi-centre	12,050	6 years mixed	asthma	parent/trained	parent/trained	Rapid growth before the age of 2 years
Sonnenschein-van der Voort et al. (2014b), UK	9,723	17 years mixed	asthma	parent	trained	Weight SDS changes during 0-3 and 3-12 months; 1-3, 3-7 and 7-10 years
De Korte-De Boer et al. (2015), Netherlands	566	3 years mixed	wheezing	parent	e-records	Weight SDS changes during 1-7, 7-14, 24-36 months of age.
Magnus et al. (2015), Norway	24,827	7 years mixed	asthma	parent	parent	1 weight SDS change during the first 36 months

Table 2.23 Risk of bias assessment table using Newcastle-Ottawa quality assessment scale for childhood growth patterns and wheezing disorder studies included in the review

Author , year, region	Study title	Selection	Comparability	Outcome
Mai et al, 2005, Sweden	Early rapid weight gain and current overweight in relation to asthma in adolescents born with very low birth weight.	★★		★★
Mamun et al, 2007, Australia	Increasing body mass index from age 5 to 14 years predicts asthma among adolescents: evidence from a birth cohort study	★★★	★★	★★
Scholtens et al, 2009, Netherlands	Overweight and changes in weight status during childhood in relation to asthma symptoms at 8 years of age	★★	★★	★★
Pike et al, 2010, UK	Patterns of foetal and infant growth are related to atopy and wheezing disorders at age 3 years	★★	★★	★★
Zhang et al, 2010, USA	Early childhood weight status in relation to asthma development in high-risk children	★★	★★	★★★
Flexeder et al, 2012, Germany	Growth velocity during infancy and onset of asthma in school-aged children	★★★	★★	★★
Magnusson et al, 2012, Sweden	Early childhood overweight and asthma and allergic sensitization at 8 years of age	★★★	★	★★
Sonnenschein-van der Voort et al, 2012, Netherlands	Foetal and Infant Growth and Asthma Symptoms in Preschool Children	★★★	★★	★★
van der Gugten et al, 2012, Netherlands	Rapid early weight gain is associated with wheeze and reduced lung function in childhood	★★★	★	★★
Anderson et al, 2013, Belarus	Associations of postnatal growth with asthma and atopy: The PROBIT Study	★★★	★★	★★
Rzehak et al, 2013, Multicentre	Body mass index trajectory classes and incident asthma in childhood: Results from 8 European Birth Cohorts—a Global Allergy and Asthma European Network initiative	★★★	★	★★

Author , year, region	Study title	Selection	Comparability	Outcome
Sonnenschein-van der Voort et al, 2014, UK	Influence of childhood growth on asthma and lung function in adolescence	★★★★	★★	★
De Korte-De Boer et al, 2015, Netherlands	Early life growth and the development of preschool wheeze, independent from overweight: The LucKi Birth Cohort Study	★★★★	★	★★
Magnus et al, 2015, Norway	Peak weight and height velocity to age 36 months and asthma development: the Norwegian mother and child cohort study	★★★	★★	★

2.5.3.3 Summary of childhood growth patterns and wheezing disorders studies

The measurements and stages of growth investigated were diverse, and those studies that used the same growth measurements were too few to be combined. Thus, no meta-analysis was carried out for growth and wheezing disorders studies' risk estimates.

Based on the diverse growth measurements and stages, the studies reported an inconsistent risk of association between growth patterns and wheezing disorders. For example, three studies reported an insignificant risk reduction (Mai et al., 2005; Sonnenschein-van der Voort et al., 2012; Sonnenschein-van der Voort et al., 2014b), five studies reported an insignificant increase (Scholtens et al., 2009; Pike et al., 2010; Zhang et al., 2010; Magnusson et al., 2012; Anderson et al., 2013), and eight studies reported a significant increase in the risk of wheezing disorders (Mamun et al., 2007; Scholtens et al., 2009; Flexeder et al., 2012; Magnusson et al., 2012; van der Gugten et al., 2012; Rzehak et al., 2013; Sonnenschein-van der Voort et al., 2014b; De Korte-De Boer et al., 2015; Magnus et al., 2015) for fast growth (Table 2.24).

Table 2.24 Summary of studies that investigated the association between childhood growth patterns and childhood wheezing disorders

Study	Growth extent and period	Unadjusted OR (95% CI)	Adjusted OR (95% CI)
Mai et al, 2005	1weight SDS increase during the first 6 months		0.49 (0.23 to 1.02)
	1weight SDS increase during the first 18 months		0.63 (0.31 to 1.26)
Mamun et al, 2007	1 BMI SDS change between 5 and 14 years	1.14 (1.03 to 1.27)	1.13 (1.02 to 1.25)
Scholtens et al, 2009	**between 1- 2 and 6-7 years**		
	early overweight	0.99 (0.73 to 1.34)	1.02 (0.72 to 1.43)
	late overweight	1.73 (1.26 to 2.39)	1.77 (1.21 to 2.58)
	persistent overweight	1.35 (0.88 to 2.06)	1.40 (0.86 to 2.28)
	between 3- 5 and 6-7 years		
	early overweight	1.13 (0.79 to 1.64)	1.14 (0.72 to 1.80)
	late overweight	1.65 (1.15 to 2.38)	1.71 (1.10 to 2.66)
	persistent overweight	1.56 (1.10 to 2.23)	1.57 (1.06 to 2.34)
Pike et al, 2010	1 weight SDS increase between 0 and 6 months	1.08 (1.03 to 1.12)	1.05 (1.01 to 1.09)
	1 weight SDS increase between 6 and 12 months	1.04 (0.98 to 1.09)	1.06 (1.00 to 1.12)
Zhang et al, 2010	0.67 weight SDS increase between 0 and 6 months	1.16 (0.59 to 2.33)	1.11 (0.50 to 2.51)
Flexeder et al, 2012	velocity of 13kg per year during 1-10 years		1.22 (1.02 to 1.47)*
Magnusson et al, 2012	**BMI changes between 1-1.5 and 7 years**		
	early overweight	1.09 (0.64 to 1.87)	1.21 (0.69 to 2.11)
	late overweight	2.34 (1.37 to 3.97)	2.51 (1.45 to 4.35)
	persistent overweight	1.87 (0,96 to 3.56)	1.89 (0.92 to 3.87)
	BMI changes between 4 and 7 years		
	early overweight	1.17 (0.56 to 2.43)	1.25 (0.58 to 2.67)
	late overweight	1.99 (1.06 to 3.73)	2.09 (1.09 to 3.98)
	persistent overweight	2.39 (1.38 to 4.13)	2.49 (1.38 to 4.49)
Sonnenschein-van der Voort et al, 2012	0.67 weight SDS increase during 3and 6 months		0.97 (0.88 to 1.06)
	0.67 weight SDS increase during 6 and 12 months		0.95 (0.86 to 1.04)

Study	Growth extent and period	Unadjusted OR (95% CI)	Adjusted OR (95% CI)
van der Gugten et al, 2012	1 SDS increase during the first 3 months	1.26 (1.11 to 1.45)	1.16 (1.01 to 1.34)*
Anderson et al, 2013	**1 weight SDS increase per month during 0 and 60 months**		
	between 0 and 3 months		1.12 (0.95 to 1.32)
	between 3 and 12 months		1.15 (0.96 to 1.36)
	between 12 and 34 months		1.00 (0.84 to 1.18)
	between 34 and 60 months		1.01 (0.80 to 1.28)
Rzehak et al, 2013	1.32 BMI SDS increase per year during 0-2 years	1.22 (1.08 to 1.39)	1.27 (1.06 to 1.51)*
	2.5 BMI SDS increase per year during 0-6 years	1.42 (0.90 to 2.27)	1.24 (0.62 to 2.47)
Sonnenschein-van der Voort et al, 2014	1 weight SDS increase during 0-10 years		
	between 0 and 3 months		1.18 (1.01 to 1.37)
	between 3 and 12 months		0.89 (0.75 to 1.06)
	between 1 and 3 years		1.03 (0.87 to 1.23)
	between 3 and 7 years		1.04 (0.86 to 1.26)
	between 7 and 10 years		0.92 (0.70 to 1.21)
De Korte-De Boer et al, 2015	1 weight SDS increase per year between 0-3 years	0.89 (0.74 to 1.05)	0.88 (0.73 to 1.05)
	1 BMI SDS increase per year between 0-3 years	1.16 (0.99 to 1.36)	1.16 (0.99 to 1.37
Magnus et al, 2015	1 weight SDS increases during the first 36 months	1.14 (1.08 to 1.20)	1.13 (1.07 to 1.20)*

*= relative risk

2.5.4 Discussion

Until the time of writing up of this thesis, no meta-analysis of previous studies of childhood growth and wheezing disorders was carried out. Results from previous epidemiologic studies remain inconsistent. For example, Mai et al reported that there was an insignificant wheezing disorder risk reduction for 1 weight SDS increase between birth and six months (Mai et al., 2005). Three other studies also reported an insignificant risk increases for 1 weight SDS increase between birth and three months (Pike et al., 2010; Anderson et al., 2013; Sonnenschein-van der Voort et al., 2014b). However, van der Gugten et al. (2012) reported a significant increase of wheezing disorders risk for 1 weight SDS increase during the first three months. Moreover, three other studies that investigated weight SDS increases during longer follow up periods have reported a significant increase in the risk of wheezing disorders (Mamun et al., 2007; Rzehak et al., 2013; Magnus et al., 2015), although another study has also reported an insignificant reduction of the risk (De Korte-De Boer et al., 2015).

In conclusion, data on childhood growth patterns and childhood disorders is sparse and results from majority of the studies remain inconclusive (i.e. the risk estimates included the value of 1). The lack of the use of standardised anthropometric measurement has also made difficult for results of the studies to be combined for meta-analyses. Thus, the association between childhood growth patterns and wheezing disorders needs further investigation.

CHAPTER 3

METHODS AND MATERIALS

3.1 Chapter overview

This chapter presents a detailed description of the methods and materials used, and the steps followed during analyses of childhood anthropometric measurements and wheezing disorders using the Born in Bradford (BiB) cohort data.

In section 3.2, a review of common methodological issues in research, namely: *causality, confounding, missing data* and *longitudinal data analysis methods* is presented. Novel approaches such as identification of confounders using *Direct Acyclic Graphs* (DAGs), *maximum likelihood* and *multiple imputation* for missing data estimation, and statistical techniques for longitudinal data analyses are discussed.

In sections 3.3-3.9, the study design, ethics statement, data collection and standardisation, variables for missing data estimation and analyses models are described in detail. The use of DAGs for selection of confounding variables and the steps followed are also presented.

Section 3.10 details the statistical methods and software used in the series of analyses, namely: the incidence and burden of childhood allergic conditions and the effects of birthweight, weight at the age of 3 years and childhood growth patterns on wheezing disorders. Latent growth models formulation, models estimation and fit evaluations techniques are described. Model implementation of non-linear growth modelling techniques (i.e. *polynomials*, *free-loading*, and *piecewise*), and best model selection procedures using Mplus software are described in detail.

3.2 Methodological issues in research

3.2.1 Association and causality

The primary aim of epidemiological research is to establish a valid association between two variables, known as *exposure* and *outcome*. Information on a set of variables is collected or becomes available for analysis. Then, the putative association between the two variables is investigated using regression models. The model variables are broadly categorised into three: *exposure*, *outcome* and potential *confounders*. However, more often, no clear definition of *confounding* variables is assumed and best model selection is based on stepwise selection methods where a variable is retained if its coefficient is statistically significant (Hernán et al., 2002).

Although a potential association (i.e. statistical relationship) between the *exposure* and *outcome* variables can be established this way, a causal relationship between the variables can not be implied (Hernán et al., 2002; McNamee, 2003). More importantly, an indiscriminate inclusion of variables into analysis models may bias risk estimates (Hernán et al., 2002; Tu et al., 2005; Schisterman et al., 2009). For example, in a simulation study by Tu et al. (2005), an inclusion of a variable that is in the causal pathway between the *exposure* and *outcome* variables was observed to cause a spurious association when there is no relationship between exposure and outcome variables; and, a reversing and exaggerating effect on the estimate when there is a negative and positive genuine relationship between the exposure and outcome variables, respectively. In order to avoid risk estimate bias and be able to infer causality, one has to adopt appropriate definition and tools to identify confounding variables.

3.2.1.1 Confounding and confounders

Confounding variable is a variable that satisfies three key aspects (McNamee, 2003):

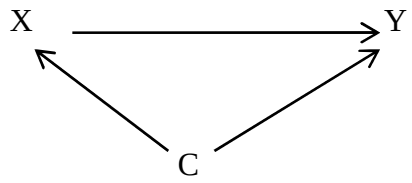
a) be a cause or a proxy cause of a disease;
b) be correlated with the main exposure variable; and,
c) not affected by the exposure variable.

Confounding occurs when an association between an exposure and outcome variables is due to a causal effect of the exposure on the outcome and sharing of a common cause variable (Hernán et al., 2002). For example, in Figure 3.1, X and Y

are statistically associated because: X causes Y, and X and Y share a variable C. In other words, variable C is confounding the causal association between X and Y so it is a confounder.

The graph in Figure 3.1 represents a causal assumption about the relationship between X, Y and C (Greenland et al., 1999; Hernán et al., 2002). A graph becomes *acyclic* if there is no directed path in the graph that forms a closed loop, hence, *Direct Acyclic Graph (DAG)* (Greenland et al., 1999). The causal relationship and the temporal precedence assumptions are based on a *priori* knowledge or speculative hypothesis about the subject matter, not on the data (Hernán et al., 2002). For example, in Figure 3.1, it is assumed that X has temporal precedence over Y, and C has temporal precedence over X and Y

Figure 3.1 Graphical reperesentation of exposure (X), outcome (Y) and a confounder (C) variable



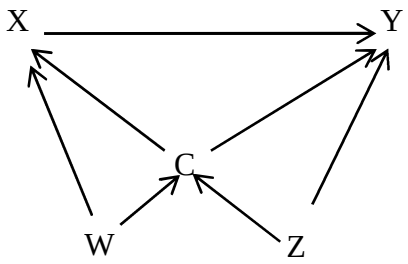
3.2.1.2 Direct Acyclic Graphs (DAGs)

Direct Acyclic Graphs (DAGs) are valuable tools in visualising the causal structure of a substantive model and identifying confounding and confounders appropriately (Greenland et al., 1999; Hernán et al., 2002).

Terminologies in DAG (Greenland et al., 1999): each of the variables are called *nodes* or *vertices* (e.g. C, X, Y, and Z in Figure 3.2). A line or arrow that connects any of two nodes is called an *arc* or *edge*. A node where an arc exits is a *parent* or *ancestor* or a *cause,* and a node that an arc enters is a *child* or *descendant.* For example in Figure 3.2, C, W and Z are *ancestors* of a *descendant* or *child* X. A *path* is the sequence of arcs connecting two or more nodes, e.g., Z—C—X is a *path* between Z and X. Two variables are called *adjacent* if they are directly connected by an *arc*, e.g, X and Y are adjacent. A *backdoor* path is a path, other than the direct path, from the exposure to the outcome. For example, all paths from X to Y except the direct path are backdoor paths. A *collider* is a variable where a path enters and

exits through arrowheads, e.g. C in Figure 3.2 is a *collider*. A path is *blocked* if it has one or more *colliders*; otherwise it is *unblocked*.

Figure 3.2 Direct Acyclic Graph

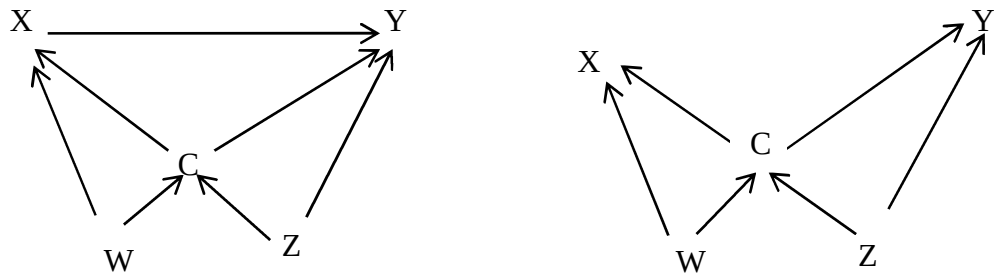


Identification of confounding and confounders

Potential confounders are qualified using the following algorithm (Greenland et al., 1999):

a) All single-headed arrows that exit from exposure variable (cause) are deleted. e.g.: the *arrow* from X to Y in Figure 3.3a is deleted to form Figure 3.3b.
b) The presence of any unblocked path from exposure to outcome (disease) in the new graph is checked, that is, if exposure and disease remain associated after the exposure effect is removed is examined. In Figure 3.3b, once the path from X to Y is removed, it can be noted that the two variables share common ancestors C, W and Z. Thus, C, W and Z are *confounders*; hence, there is a *confounding*.

Figure 3.3 Graphical representation of confounders assessment process



a) b)

Implementing the above two algorithms would result in a *sufficient* set of variables. That is, all *backdoor* paths from X to Y will be *blocked* by including the set of variables, known as, *confounders*. For example, variables C, W and Z in Figure 3.3 are sufficiently enough to block the backdoor path from X to Y. However, including all three variables in a model as confounders may not be ideal choice for two main reasons. First, having fewer variables in a model increases the degree of freedom and model efficiency (Shrier and Platt, 2008). Second, cost of data collection is directly proportional to the number of variables. Thus, selecting a minimum subset of variables for adjustment is always preferable (Greenland et al., 1999; Shrier and Platt, 2008).

The model in Figure 3.3 is not a complex model; selection of *minimally sufficient* sets of confounders can be achieved by adding a linking line between two variables who share a *child* (i.e. W and Z). Thus, it becomes clear that either C and W or C and Z would form the *minimally sufficient* sets of confounding variables.

For complex models, minimally sufficient sets confounders can be identified by a six step algorithm recommended by Shrier and Platt (2008). However, identifying of confounders by hand can be very difficult task and errors can arise. Thus, it is much faster and safer to generate automated results of minimally sufficient sets of confounders using DAGitty software (Textor et al., 2011) than doing the task by hand.

3.2.2 Missing data in research

Missing data are inevitable in research but *complete cases analyses* were more commonly used to address the problem in the past (Schafer and Graham, 2002; Wood et al., 2004). However, analyses that are restricted to individuals who have no missing data in any of the analysis variables have two disadvantages (Royston et al., 2009; Sterne et al., 2009; White et al., 2011). First, results can be biased if the remaining data become unrepresentative of the sample population due to an exclusion of those individuals with missing data in any of the analysis variables. Second, the exclusion of individuals with missing data in variables may cause a substantial reduction of sample size which can lead to a loss of study power and estimate precision. In order to address these concerns, researchers used either *ad hoc (also* known as *traditional methods)* or *principled* (also known as *modern methods*) missing data estimation methods in the past.

Ad hoc methods such as *mean substitution* (missing values are replaced by the average of the complete data) and *regression imputation* (missing values are replaced by conditional means), *hot deck imputation* (replacing missing values by a random draw from the observed values) were widely implemented as alternatives to case deletion or complete case analysis (Schafer and Graham, 2002; Duncan et al., 2006c). However, although these methods can estimate the missing values, they often lead to biased parameter estimates (Schafer and Graham, 2002). Thus, *multiple imputation* (Rubin, 1987) and *maximum likelihood* (Dempster et al., 1977) methods have been recommended.

3.2.2.1 Multiple imputation

Multiple imputation (MI) is a statistical technique for handling of missing data where the missing values are replaced by a set of simulated values from the distribution of observed data (Schafer and Graham, 2002 ; White et al., 2011). Model parameters estimation using MI is carried out in three stages: *imputation of data sets, statistical analysis of individual sets and combining of results* (White et al., 2011).

Imputation of datasets

At this stage, missing data are filled with *m* independent simulated sets of values from the posterior distribution of the missing data conditional on the observed data

resulting in several completed data sets (White et al., 2011). Imputation can be carried out using a joint modelling or fully conditional specification, also known as multiple imputations by chained equations (MICE), the latter being easier and flexible to implement than the former (van Buuren, 2007).

MICE is carried out through a sequence of regression models (White et al., 2011). For example, suppose that variables x_1, x_2 and x_3 have incomplete and x_4 complete data. First, x_1 is regressed on x_2 , x_3 and x_4 restricted to individuals with information on x_1. Missing values in x_1 are replaced by simulated draws from the corresponding posterior predictive distribution of x_1 assuming a vague prior distribution for the parameters in the regression model. Then, x_2 is regressed on x_1 , x_3 and x_4 restricted to individuals with information on x_2 (note that x_1 is complete at this stage). Missing values in x_2 are replaced as in x_1. The process is repeated for x_3 to complete a cycle. The procedure is repeated for several cycles to produce a single dataset and the whole procedure is repeated *m* times to produce *m* imputed datasets (White et al., 2011).

In MI, the number of data sets to be imputed (*m*) depends mainly on the fraction of missing information (FMI) and White et al recommend that *m* should be at least equal to the percentage of incomplete cases (White et al., 2011). Thus, $m \geq 1\text{-}p$; where *p* is the proportion of individuals with complete observation in all of the variables.

Statistical analysis

This stage is straight forward, each completed dataset is analysed separately by fitting a regression model. In the end, *m* sets of parameter estimates are obtained from separate analyses of *m* datasets.

Combining of results

At this stage, the *m* sets of parameter estimates are combined using Rubin's rules (Rubin, 1987) incorporating both the within-imputation variability and between-imputation variability (White et al., 2011). Consistency of results produced from the *m* imputed datasets can be checked using Monte Carlo errors from a jackknife procedure (Royston et al., 2009); if the errors of the beta coefficient (or risk estimate), test statistic (t-value) and the p-value are less than 10% of the standard

error, less than 0.1, and less than 0.2 respectively, the consistency is considered to be adequate (White et al., 2011).

3.2.2.2 Maximum likelihood

Maximum likelihood (ML), unlike MI, does not create datasets but rather estimate the parameters directly by maximizing the complete data log likelihood function (Dempster et al., 1977; Enders, 2001a; Schafer and Graham, 2002). There are three types of ML techniques: *Expectation Maximisation, Full Information Maximum Likelihood* and *multi-group approach* (Enders, 2001b).

The *Expectation Maximisation* (EM) algorithm uses a two-step iterative procedure to estimate parameters (Enders, 2001b). In the E step, missing values are replaced with the conditional expectation of the missing data given the observed data and an initial estimate of the covariance matrix. In the M step, ML estimates of the mean vector and covariance matrix are obtained using the sufficient statistics calculated at the previous E step. Then, the resulting parameter estimates in the M step are used to derive the new estimates of missing values at the next E step, and the process begins again. The algorithm repeatedly cycles until convergence is achieved (Enders, 2001b).

Multi-group approach: the sample is divided into G-subgroups based on their pattern of missing data. That is, observations within each of the G-subgroups have the same set of variables present and missing. A likelihood function is computed for each of the G groups, and the group-wise likelihood functions are accumulated across the entire sample and maximized (Enders, 2001b).

Full Information Maximum Likelihood (FIML): originally developed for factor analysis, is similar to the multi-group approach although the likelihood function is calculated at the individual, rather than the group level (Enders, 2001b). However, the key identifying feature of FIML is that it uses the raw data as input and hence can use all the available information in the data as opposed to the other ML methods that use observed covariance matrix which contains less information than the raw data (Enders and Bandalos, 2001; Schafer and Graham, 2002).

3.2.2.3 Missing data mechanisms

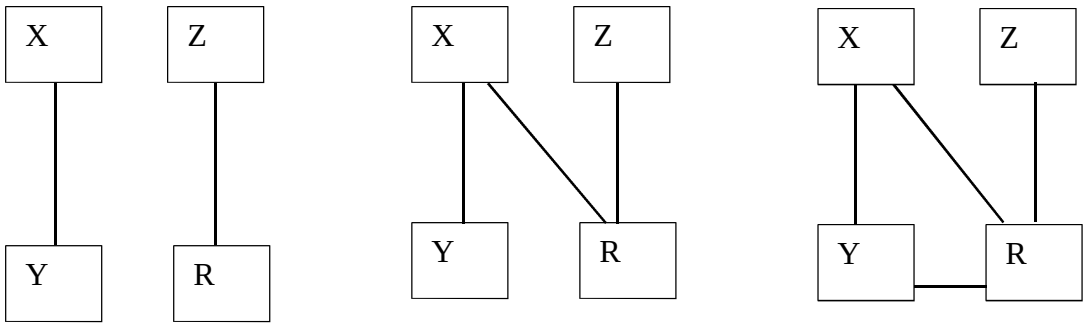
Modern missing data estimation methods (i.e. MI and ML) are not always the best option. Implementing the techniques consumes resources and time, and sometimes there may be no benefit of using them. To avoid unnecessary waste of resources, analysts are advised to look into the *properties* or *mechanisms of the missing data* prior to executing missing data estimation methods (Collins et al., 2001; Schafer and Graham, 2002; Sterne et al., 2009). There are three types (mechanisms) of missing data: *missing completely at random*, *missing at random*, and *missing not at random* (Collins et al., 2001; Schafer and Graham, 2002; Sterne et al., 2009).

Suppose that we have haemoglobin level as an outcome or dependent variable Y partly missing, cause of missing variable Z, a missing indicator or missingness variable R with values of 1 if Y is missing and 0 if Y is not missing, gender as an explanatory variable X completely observed . Therefore the three missing data mechanisms can be elucidated as follows (Collins et al., 2001; Schafer and Graham, 2002; Sterne et al., 2009). In *missing completely at random* (MCAR), the reason for missingness (Z) on the outcome variable (Y) is entirely unrelated to the outcome variable itself. For example, some blood haemoglobin levels (Y) are missing due to contamination of samples (Z); see Figure 3.4a. In *missing at random* (MAR), the cause of missingness (Z) may be related to the outcome variable (Y) but only indirectly through another variable. For instance, missing haemoglobin levels (Y) may be higher than the non-missing levels but only because women missed the appointment provided that there is no relationship between missingness and blood haemoglobin level within men and women groups; see Figure 3.4b. In the case of *missing not at random* (MNAR), the outcome variable (Y) itself is associated with the missingness (R). For example, people with lower blood haemoglobin level tended to miss appointments (R) because they felt very tired (Z); see Figure 3.4c.

Assessing data for mechanisms of missingness prior to performing analysis has implication on the decisions to be taken. If the data missingness is MCAR, complete cases analysis and modern techniques (MI or ML) provide identical results (Wood et al., 2004) so complete cases analysis may be preferred over MI or ML. However, if the mechanism of missingness is either MAR or MNAR, analyses based on the modern missing data estimation methods provide unbiased and more efficient

parameter estimates than analyses based on *ad hoc* or case deletion (i.e. complete cases) methods.

Figure 3.4 Graphical representation of missing data mechanisms; (a) MCAR (b) MAR and (c) MNAR



a) b) c)

3.2.3 Analysis of weight change and childhood growth patterns in relation to disease outcomes

Childhood weight change and growth patterns have been reported as predictors of health during childhood and adult life. For example, higher growth rate during childhood and adulthood has been related to hypertension (Hardy et al., 2004; Eriksson et al., 2007; Halldorsson et al., 2011), chronic heart disease (Baker et al., 2007; Owen et al., 2009) diabetes (Eriksson et al., 2003), and asthma (Rzehak et al., 2013). However, except for the method by Rzehak et al. (2013) multiple regression approaches are prone to collinearity problems caused by the repeated weight measurements that can lead to biased coefficient estimates (Duncan and Duncan, 2004; Tu et al., 2013).

3.2.3.1 Longitudenal continuous data analysis methods

Based on literature, the statistical techniques used were: generalised estimating equations (Ballinger, 2004; Hwang and Takane, 2005), multilevel linear models (Bryk and Raudenbush, 1987; Howe et al., 2013) and latent growth models (Muthen, 2001; Duncan and Duncan, 2004; Muthén, 2004; Duncan et al., 2006b). Generalised estimating equation (GEE) models were developed by Liang and Zeger for longitudinal panel data analysis (Liang and Zeger, 1986) and recently have been proposed for longitudinal continuous data (Hwang and Takane, 2005) although their

use have so far been limited. GEE models are superior to the multiple regression models as the effects of within-subject correlation are corrected by multiplying of the variances against the matrix of correlation coefficients (Ballinger, 2004).

Multilevel linear models (MLMs), also known as Hierarchical linear models, address the issues of correlation between repeated measurements, status and change over time (Bryk and Raudenbush, 1987) like GEE models. In fact, MLMs and GEEs can be seen as equivalent except that the random growth factors (i.e. intercept and slope) are assumed to be Gaussian (normally distributed) in MLMs whereas no such assumption needs to be made in GEE models (Ballinger, 2004). However, GEE and MLMs are only capable of deriving a single overall mean growth trajectory, that is, the growth trajectories of individuals over time are assumed to be homogeneous.

Latent growth models (LGMs) account the within-subject level correlations and capture individual differences of growth trajectories over time through the continuous random factors, that is, intercept and slope (Muthén, 2004; Duncan et al., 2006c), like GEEs and MLMs. However, LGMs more flexible than GEE and MLMs as the homogeneity assumption can be relaxed and whether the population under study is made-up of one or sub-groups of population can be tested.

In summary, the difference between GEEs, MLMs, and LGMs depends on the assumption made about *growth trajectories*. For example, suppose that we want to know the growth trajectories of children in a school. In both GEEs and MLMs, measurement occasions are nested within individuals, who are nested within the school. The two models will provide only one growth trajectory. In other words, these models only assume a homogeneous growth among the children and cannot differentiate between more than one distinct growth trajectories within the school. However, ignoring the potential variability among individuals can provide biased estimates (Muthen, 2001; Muthén, 2004; Duncan et al., 2006b). The LGMs on the other hand, assume that the population under study can be made up of unknown sub-groups or growth trajectory groups (Duncan et al., 2006c; Gilthorpe et al., 2014). This assumption then allows us to test if there is more than one distinct growth trajectory and parameters be estimated accordingly.

3.2.3.2 Approaches for the analysis of growth trajectories in relation to disease outcomes

Analysis of growth trajectories in relation to covariates (e.g. disease outcome) remains to be challenging. Based on literature, researchers have used a variety of approaches to identify the potential growth trajectory groups in a population and investigate their relationship with covariates. Broadly, the approaches can be grouped into *one-step* and *three-step* methods (Vermunt, 2010; Asparouhov and Muthén, 2014).

The one-step method

The *one-step* approach is a type of analysis where a researcher uses a set of covariates as predictors of growth trajectories (Vermunt, 2010; Asparouhov and Muthén, 2014). In this approach, the identification of growth trajectories and their relationship with covariates are carried out simultaneously. Here, the main interest is to understand which of the covariates predict the growth trajectories and the model implementation is straight forward. However, this modelling approach has attracted criticisms for the following reasons (Vermunt, 2010). First, convergence of models may become difficult when the number of potential covariates is large. Second, it introduces additional model building problems, such as whether one should decide about the number of classes in a model with or without covariates. Third, the simultaneous approach does not fit with the logic of most applied researchers, who view building a classification model is a step that comes before introducing covariates.

The three-step method

Unlike in the *one-step* method, identification of growth trajectories and investigation of covariates relationship with the growth trajectories are carried out in separate steps (Asparouhov and Muthén, 2014). In the first step, the growth trajectories are estimated using only the repeated weight measurements overtime. In the second step, the most likely class variable is created using the latent class posterior distribution obtained during the first step. In the third step, the most likely class is regressed on covariates, taking into account the misclassification or classification uncertainty.

In practice, the three-step modelling approach is also implemented in a *classify-analyse* fashion where the growth class memberships of individuals are estimated using the first two steps of the *three-step* approach. The class membership data are then saved and merged with the original data. In step three, a covariate (distal outcome variable, also known as disease outcome variable) can be regressed on the categorical latent growth class variable. The drawback of *classify-analyse* approach is that the classification uncertainty is ignored and parameter estimates can be biased as a consequence (Lanza et al., 2013). However, the approach works well if the classification quality of the classification models (*entropy*) is greater or equal to 0.80 (Clark, 2010).

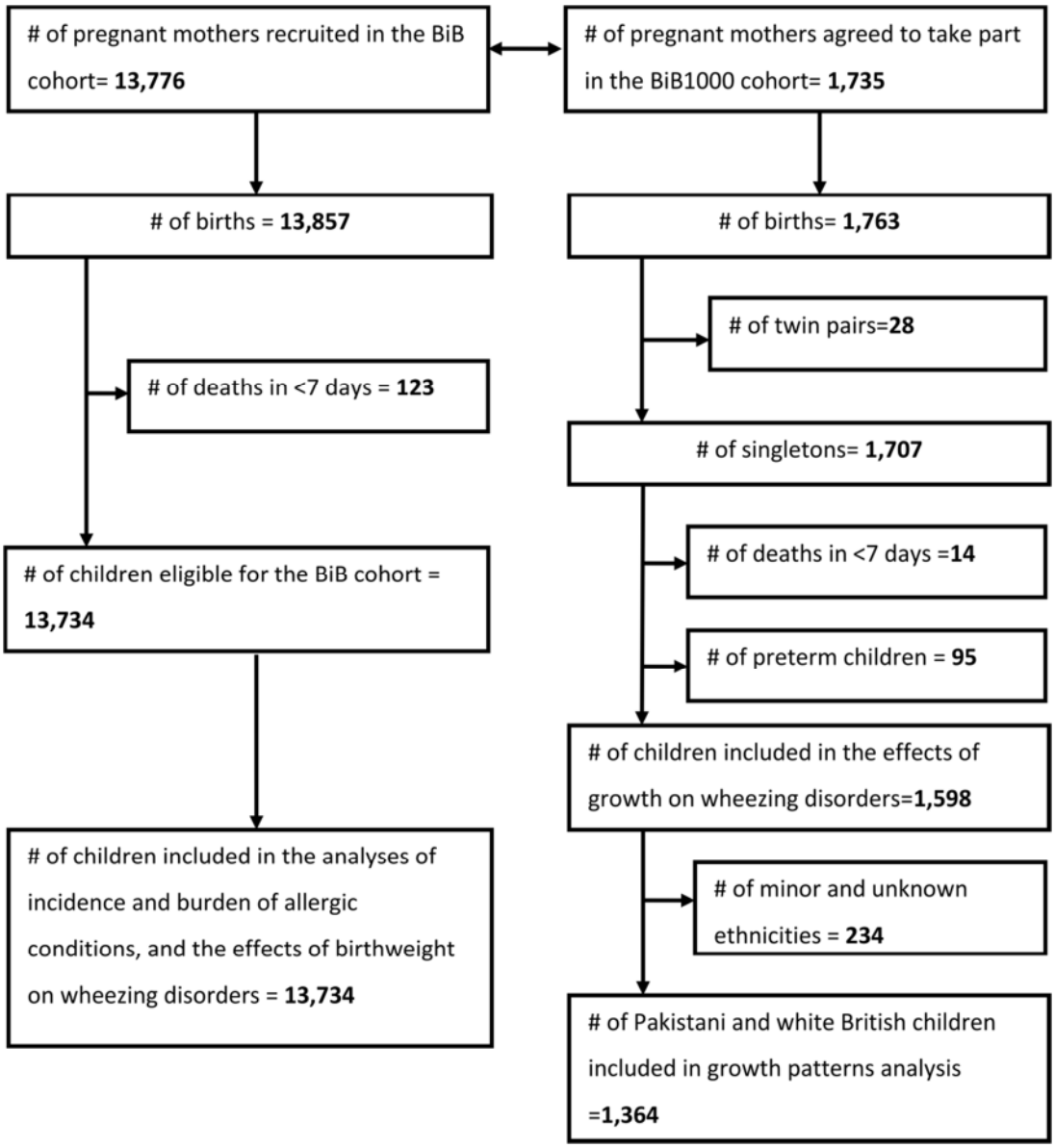
3.3 Study design and participants

The BiB study is a prospective cohort, mainly bi-ethnic, which examines the impact of environmental, genetic and social factors on health of the population of Bradford (Wright et al., 2013). The project's main aims focus on following up over 13,500 children from childhood up to adulthood life. Further information about the BiB project is available from the project website (http://www.borninbradford.nhs.uk/about-the-project/)

Participants were pregnant mothers who attended Bradford Royal Infirmary antenatal clinic and wished to give birth at the hospital. Recruitment of participants started in March 2007 and ended in December 2010. A total of 13,776 pregnant mothers were recruited resulting in 13,857 births. Of the total births, 123 died before 7 days. Thus, data from 13,734 children were used for analyses of incidence and burden of allergic diseases, and the effects of birthweight on wheezing disorders (Figure 3.5).

At the time of initial recruitment, pregnant mothers were also asked to join a sub-cohort known as BiB1000 for regular follow up. A total of 1,735 mothers agreed to participate resulting in 1,763 babies; 1707 singletons and 28 twins. With the exclusion of 14 babies due to death before 7 days and another 94 due to being preterm, a total of 1,598 children were included in the investigation of effects of weight at the age of 3 years and childhood growth on wheezing disorders. With a further exclusion of 234 babies who were either minority or unknown (missing value for) ethnicity, a total of 1,364 Pakistani and white British children were included in growth patterns analysis of the two ethnicities (Figure 3.5).

Figure 3.5 Study participants' recruitment process flow chart for the analyses carried out using the Born in Bradford cohort



3.4 Ethics statement

Ethics approval for the Born in Bradford project was granted by Bradford Research Ethics Committee (Ref 07/H1302/112.).

3.5 Data collection

3.5.1 Incidence and burden of wheezing disorders, eczema and rhinitis

Three data sources were used:

a) Baseline questionnaire for information on ethnicity of the mother (i.e., proxy for child's ethnicity);
b) The hospital maternity records for sex, date and outcome of birth; and,
c) Information on wheezing disorders, eczema, and rhinitis drug prescription from the SystmOne (http://www.tpp-uk.com/products/systmone) primary care data.

Baseline questionnaire and hospital maternity data were directly available from the data and research team of the BiB project (http://www.borninbradford.nhs.uk/our-team/) upon request. However, collection of drug prescription data was conducted in two steps. First, a list of drug family names and chapters were extracted from the British National Formulary (BNF) for Children Handbook version 2015 and were given to the data manager of BiB project. Based on the list of drug names and chapters the BiB data manager then conducted a SystmOne data query and extracted the data, which were then available upon request. The list of drug family names and BNF chapters used are available in Appendix E .

3.5.2 Effects of birthweight and weight at the age of 3 years on wheezing disorders

Four data sources were used:

a) The hospital maternity records for information on birth weight, gestational age, gender of a child, number of births, birth outcome;
b) The BiB1000 cohort records for weight at the age of 3 years (36 months);
c) Baseline questionnaire for information on the mother's ethnicity, smoking and socioeconomic status; and,
d) SystmOne (http://www.tpp-uk.com/products/systmone) primary care data for information on wheezing disorders.

All data were available from the BiB project data and research team upon request.

3.5.3 Describing growth patterns of white British and Pakistani children

The following data sources were used:

a) The hospital maternity records for information on birth weight, gestational age, gender of a child, number of births, birth outcome;
b) The community health records for weights at 1 and 3 months;
c) The BiB1000 cohort records for weights at 6, 12, 18, 24 and 36 months; and
d) Baseline questionnaire for information on the mother's ethnicity, smoking and socioeconomic status.

All data were available from the BiB project data and research team upon request.

3.5.4 The effect of childhood growth patterns on childhood wheezing disorders

Five data sources were used:

a) The hospital maternity records for information on birth weight, gestational age, gender of a child, and number of live births, birth outcome;
b) BiB1000 cohort records for weight at 6, 12, 18, 24 and 36 months of age;
c) Community health records for weight at 1 and 3 months of age;
d) Baseline questionnaire for information on the mother's ethnicity, smoking and socioeconomic status ; and,
e) SystmOne primary care data for information on wheezing disorders.

Baseline questionnaire, BiB1000 cohort, community health, and hospital maternity data were available from the BiB project data and research team upon request. Wheezing disorders drug prescription data were collected as in section 3.5.1. However, data on disease diagnosis were collected as follows: a list of wheezing disorder terms and ids were extracted using NHS Clinical Terminology Browser (http://www.hscic.gov.uk/standards) and were given to the BiB data manager. A SystmOne primary care data query and extraction were conducted by the BiB data manager, which were then available upon request. See Appendix E and Appendix F for the list of drugs and diseases terms used.

3.6 Data standardisation

Prior to conducting growth patterns analyses, repeated weight measurements of 1,598 BiB1000 children were converted into standardised weight scores (SD score, SDS). Age-specific and sex-specific weight SDS were derived based on WHO growth standards (WHO, 2006) in LMSgrowth Microsoft excel add-in software (Pan and Cole, 2012). However, the WHO growth standards population that were used to derive the SDS scores was made up of singleton term births. Therefore, multiple and preterm births were excluded from the growth patterns analyses.

The WHO growth standards included children from Brazil, Ghana, India, Norway, Oman and USA (WHO, 2006). Therefore, it can be said that it is a representative of all children in the world, and the BiB cohort in specific.

To facilitate interpretability of the growth trajectories, the weight SDS were then converted into percentiles using a one-sided normal standard distribution. For example, weight SDS of -1.64, 0, 1.04, and 1.64 are equivalent to the 5^{th}, 50^{th}, 85^{th} and 95^{th} centiles, respectively.

Note that the use weight SDS was preferred over raw weight due to two main reasons. First they are equivalent to BMIs so they become standard comparison tool for growth of children over time (WHO, 2006). Second, they are convertible to percentiles which then can be plotted onto the child growth charts in order to assess growth patterns over time (Pan and Cole, 2012).

Birthweights of 13,734 BiB children were classified into three; based on the CDC (CDC, 2009) and WHO (WHO, 2014) guidelines where <2.5kg=low, 2.5-4.0kg=normal and >4.0kg=high.

Weight at the age of 3 years for 1,598 BiB1000 children was first converted into age-specific and sex-specific weight SDS. Then, weight SDSs were categorised into underweight ($<5^{th}$ centile), normal ($\geq5^{th}$ and $<85^{th}$ centiles), overweight ($\geq 85^{th}$ and $<95^{th}$ centiles), and obese ($\geq95^{th}$ centile) categories based on CDC classification (CDC, 2014). Weight SDS of -1.64, 1.04, and 1.64 were used cut-off points for the 5^{th}, 85^{th} and 95^{th} centiles, respectively.

3.7 Outcome definition and ascertainment

3.7.1 Incidence and burden of wheezing disorders, eczema and rhinitis

Incidence of allergic conditions was confirmed through questionnaires and clinician-diagnosis data in the past. However, these types of data are prone to recall and reporting bias respectively which are very likely to underestimate the true impact level of a disease. In order to overcome the potential for the misdiagnoses, a treatment based algorithm was used to allocate a diagnosis of allergic conditions where eczema, rhinitis and wheezing disorder cases were ascertained by the existence of at least two respective drug prescriptions a minimum of one week and maximum of 12 months apart. Appendix E lists the drugs used to confirm the diagnosis of eczema, rhinitis and wheezing disorders.

3.7.2 Effects of birthweight, weight at the age of 3 years and growth patterns on childhood wheezing disorders

Four disease definitions were drawn up based on diagnostic codes and prescribed medication details entered by general practitioners onto the primary care database.

a) *Asthma diagnosis*: confirmed by the presence of asthma codes in the record;
b) *Wheezing symptoms*: confirmed by the presence of wheezing diagnosis codes in the record;
c) *Wheezing disorder diagnosis:* it is wheezing disorder based on diagnosis and was confirmed by the presence of asthma or wheezing diagnosis codes in the record; and,
d) *Wheezing disorder treatment:* it is wheezing disorder based on treatment so was confirmed by the existence of at least two drug prescriptions indicated for the treatment of asthma a minimum of 1 week and maximum of 12 months apart.

In the process, the following assumptions were made. First, if a drug was prescribed only once in 12 months, there is a high possibility that the reason for the consultation was another illness other than wheezing disorders. Second, if two drugs were prescribed in a week, there is a high possibility that the reason for the two consultations was the same issue.

Appendix E and Appendix F list the drugs and disease names used to confirm diagnosis.

3.8 Variables for analysis

3.8.1 Incidence and burden of wheezing disorders, eczema and rhinitis

Wheezing disorders, eczema and rhinitis were the outcome variables for the analyses of incidence and prevalence of allergic conditions. Sex, ethnicity and birth year were used as stratifying variables.

3.8.2 Effects of birthweight and weight at 3 years on childhood wheezing disorders

3.8.2.1 Exposure variable

Birthweight was classified based on the CDC (CDC, 2009) and WHO (WHO, 2014) guidelines where <2.5kg=low, 2.5-4.0kg=normal and >4.0kg=high. Weight at the age of 3 years was categorised into underweight (<5th centile), normal (≥5th and <85th centiles), overweight (≥ 85th and <95th centiles), and obese (≥95th centile) categories based on CDC classification (CDC, 2014). Hence, the exposure variables were categorical forms of birthweight and weight at the age of 3 years.

3.8.2.2 Outcome variables

The outcome variable was wheezing disorder. Four disease definitions of wheezing disorders were used: *asthma diagnosis*, *wheezing symptoms*, *wheezing disorders diagnosis* and *wheezing disorders treatment;* see section 3.7 for details on how these disease definitions were drawn up.

3.8.2.3 Confounding variables

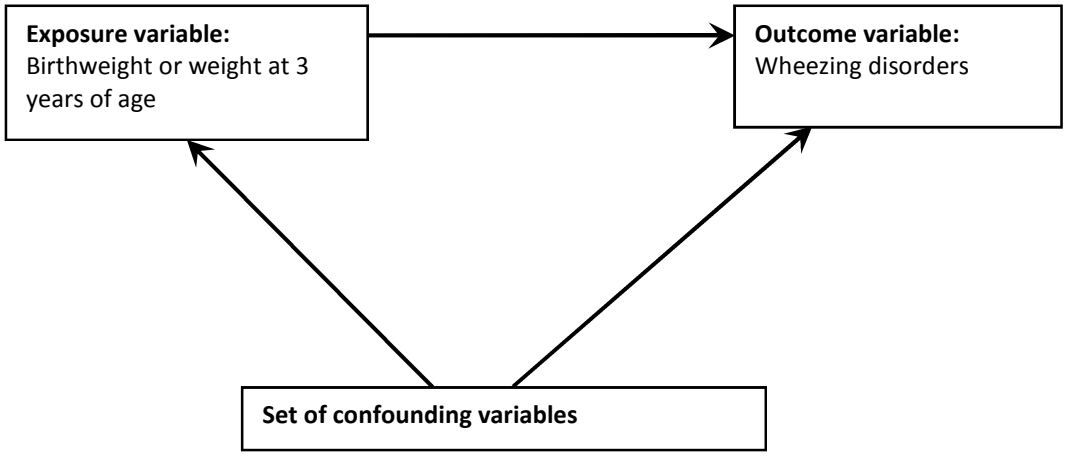
Selection of confounding variables was carried out in four steps. First, based on *a priori* understanding and speculative hypothesis, and a *second-person* opinion, a list of variables that believed to be linked with the exposure and/or outcome variables was constructed. For birthweight (exposure) and wheezing disorders (outcome) models; ethnicity, family asthma, gender, gestational age, household environment, maternal smoking, maternal BMI, maternal age, maternal feeding habits, number of live births, outdoor playing time and parity were listed. For weight at the age of 3 years (exposure) and wheezing disorders (outcome) models, birthweight, breast feeding, ethnicity, family asthma, gender, gestational age, household environment, maternal smoking, maternal BMI, maternal age, maternal feeding habits, number of live births, outdoor playing time and parity were identified.

Second, causal diagrams were drawn to represent the relationship between *main variables* and the *other covariates*. Note that '*main variables'* refers to the *exposure* and *outcome* variables. Figure 3.6 illustrates the schematic view of confounding variables when investigating the effects of birthweight and weight at the age of 3 years on wheezing disorders.

Third, based on DAGs principle (Greenland et al., 1999; Shrier and Platt, 2008), covariates were retained if they are not affected by but have a direct effect on the exposure and outcome variables (McNamee, 2003).

Fourth, selection of a *minimally sufficient* sets of confounding variables using a six-step algorithm recommended by Shrier and Platt (2008) was attempted initially. However, owing to the models complexities, DAGitty software (Textor et al., 2011) was used to generate automated results.

Figure 3.6 Diagrammatic view of the relationship between confounding and main variables for models that investigated the association of birthweight and weight at the age of 3 years with wheezing disorders



The three rectangular boxes, populated by the exposure, outcome and confounding variables, represent the nodes or vertex of the DAGs. The single headed arrows that connect the variables are arcs and indicate that the two variables are causally related. Note that the confounding variables have two single-headed arrows directed to the exposure and outcome variables.

3.8.3 Describing the growth patterns of white British and Pakistani children

The standardised scores of repeated weight measurements, that is, weight SDS, were the exposure variables and the growth patterns identified by the LGCMs and GMMs were the outcome variables. The time scores used were: 0, 1, 3, 6, 12, 18, 24 and 36.

The growth class analyses models were adjusted for the following variables that were known to affect birthweight and childhood growth: mother's ethnicity (Saxena et al., 2004), maternal smoking during pregnancy and parity (Ong et al., 2002), and maternal level of education, that is, as a proxy for SES (Luo et al., 2006)

3.8.4 Effects of growth patterns on childhood wheezing disorders

3.8.4.1 Exposure variables

Two types of exposure variables were assumed for the analyses of the effect of childhood growth patterns on wheezing disorders.

Velocity

As a starting point, LGCMs were carried out in order to drive growth patterns of the BiB1000 children between birth and 3 years of age. A piecewise model with two knots or joint points (i.e. at 3 and 12 months) was fitted to the BiB1000 children's growth data resulting in 3 *velocities;* see section 3.10.2.2 for details about piecewise models. The velocities (i.e. slopes) were between birth and 3 months, 3 and 12 months, and 12 and 36 months. These velocities are continuous and were used as exposure variables. The time scores used were 0, 1, 3, 6, 12, 18, 24 and 36.

Growth classes

Growth mixture models were carried out in order to derive growth classes based on the children's growth trajectory similarities from birth until the age of 3 years. Each of the growth classes or group of children would have distinct trajectory characteristics. For example, a class can be composed of children who grow consistently, that is, their weight percentiles remain the same starting from birth until the age of 3 years. A class can also be made up of children who show low growth percentiles during the first 6 months and very high growth percentiles between 2-3 years, or vice versa. Collectively, these *growth classes* make up a single

categorical variable, that is, the exposure variable. The same time scores were used as in the LGCMs and GMMs above.

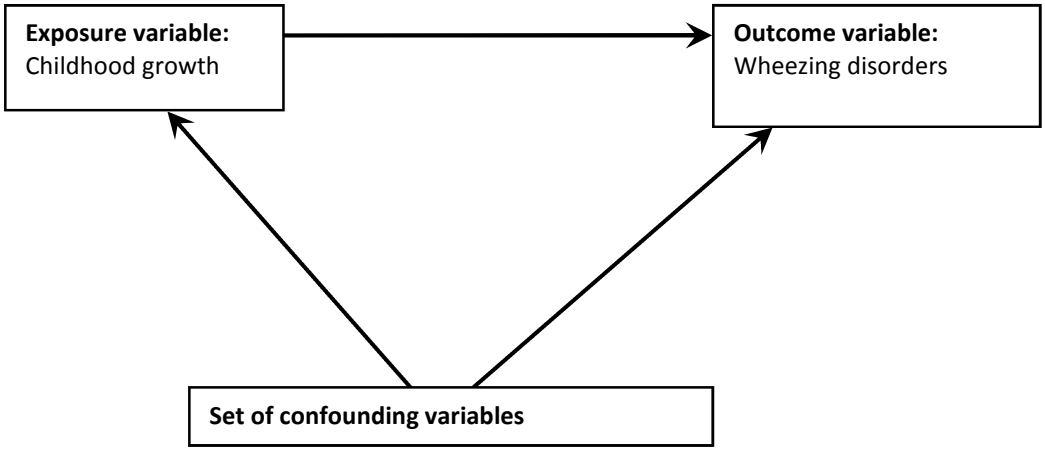
3.8.4.2 Outcome variables

Four disease definitions of wheezing disorders were used as outcome variable, namely: *asthma diagnosis*, *wheezing symptoms*, *wheezing disorders diagnosis* and *wheezing disorders treatment*, see section 3.5.2 for details on how these disease definitions were drawn up.

3.8.4.3 Confounding variables

Based on the same definition and procedure used in section 3.8.2.3; birthweight, breast feeding, ethnicity, family asthma, gender, gestational age, household environment, maternal smoking, maternal BMI, maternal age, maternal feeding habits, number of live births, outdoor playing time and parity were identified. Figure 3.7 illustrates the schematic view of adjustment for confounding when investigating the effects of childhood growth on wheezing disorders. Then, DAGitty software was used to select minimally sufficient confounding variables.

Figure 3.7 Diagrammatic view of the relationship between confounding and main variables for models that investigated the association between childhood growth and wheezing disorders



The three rectangular boxes, populated by the exposure, outcome and confounding variables, represent the nodes or vertex of the DAGs. The single headed arrows that connect the variables are arcs and indicate that the two variables are causally related. Note that the confounding variables have two single-headed arrows directed to the exposure and outcome variables.

3.9 Variables for missing data estimation

3.9.1 Growth patterns analyses of two ethnic groups (white British and Pakistani) and the BiB1000 children

Missing growth data at the 1, 3, 6, 12, 18, 24 and 36 months of age were estimated using FIML models using weight SDS at birth and the other time points (i.e. 1, 3, 6, 12, 18, 24 and 36 months).

3.9.2 Effects of birthweight, weight at the age of 3 years and growth patterns on childhood wheezing disorders

Missing data on covariates were estimated in MICE models using analysis variables (i.e. *exposure*, *outcome* and *cofounding* variables) and two additional variables (i.e. maternal hypertension and diabetes). For example, for birthweight and wheezing disorders models: birthweight, sex, ethnicity, smoking, hypertension, diabetes, SES, parity, number of births, gestational age, asthma diagnosis, asthma treatment, wheezing disorders treatment and diagnosis were included in the missing data estimation models.

3.10 Statistical methods and software

3.10.1 Incidence and burden of wheezing disorders, eczema and rhinitis

Cumulative incidence was defined as the proportion of the cohort of children with allergic conditions during 0-7 years follow up period which was calculated as the total incident cases divided by the total number of the cohort of children at risk.

Cumulative incidence rates were defined as the number of new cases of allergic conditions per the cohort population at risk in a given period calculated as the ratio of number of children diagnosed with the condition to the total person-years at risk.

Five-year period prevalence was defined as the proportion of the cohort of children with allergic conditions during 5 years follow up period which was calculated as the total incident cases divided by the total number of the cohort of children at risk.

Birth year, ethnicity and sex specific incidence rates were calculated as the ratio of number of children diagnosed with allergic conditions for the particular birth year, ethnicity and sex to the respective total person-years at risk. Ethnicity and sex specific incidence rate ratios were defined as incidence rate ratio of Pakistani to white British and boys to girls respectively.

Analyses were carried out in Stata software version 12 (StataCorp, 2011). To calculate the cumulative incidence, `proportion` Stata command was adopted. `stptime` and `stir` Stata commands were used to calculate the incidence rates and incidence rate ratios, respectively. Five per cent significance levels and 95% confidence intervals were adopted throughout.

3.10.2 Describing growth patterns of white british and Pakistani children

As described in section 3.2.3, there are three modelling options to address the problem of collinearity among repeated weight measurements, that is, GEE, MLMs and LGMs. However, LGMs were preferred over the other modelling techniques because of their flexibility to estimate missing growth data, test of heterogeneity and displaying of growth patterns graphically using Mplus software.

3.10.2.1 Latent growth model formulation

Latent Growth Curve Models (LGCMs) are the basic form of LGMs that assume the sample under study arises from a homogenous population (Duncan et al., 2006c). Growth mixture models (GMMs), which are an extension of LGCMs (Gilthorpe et al., 2014), however, assume that the population under study is made up of unknown sub-groups or latent classes (Muthen, 2001; Muthén, 2004; Duncan et al., 2006c).

LGCM can be seen as a multilevel model in which, from the LGCM in Figure 3.8; the individual repeated weight measurements at level one and the latent growth factors (i.e. *intercept* and *slope*) at level two (Duncan et al., 2006d). As such, the relationship between the growth factors, path coefficients (also known as *factor loadings*) and repeated weight measurements of linear LGCM can be expressed in multilevel notation as follows:

$$y_{ti} = \eta_{0i} + \eta_{1i}\lambda_{ti} + \varepsilon_{ti} \quad , \quad \varepsilon_{ti} \sim N(0,\sigma^2_{\varepsilon ti}) \tag{3.1}$$

$$\eta_{0i} = \alpha_0 + \zeta_{0i} \, , \quad \zeta_{0i} \sim N(0, \sigma^2_{0i}) \tag{3.2}$$

$$\eta_{1i} = \alpha_1 + \zeta_{1i} \, , \zeta_{1i} \sim N(0, \sigma^2_{1i}) \tag{3.3}$$

where y_{ti} is the weight measured (e.g. W_1 in Figure 3.8) for the i^{th} individual at time t; η_{0i} and η_{1i} are the latent growth factors, that is, intercept and slope (also known as *velocity*), respectively; the λ_ts are time scores; the ε_{ti} is a composite error term representing both random measurement error and time specific influence of the i^{th} individual. α_0 ($alpha_0$) is the model estimated overall mean of the initial weight and α_1 ($alpha_1$) is the linear average rate of weight change over time. ζ_{0i} and ζ_{1i} are error terms representing the variations among individuals in terms of initial weight measurements and the linear changes over time, respectively. Equation 3.1 is the within subject model whereas equations 3.2 and 3.3 are the between subject models. Note that ε_{ti} is assumed to be normally and independently distributed with its mean (i.e. equal to zero) and variance (i.e. $\sigma^2_{\varepsilon ti}$). The ζ_{0i} and ζ_{1i} are normally distributed with their means (i.e. equal to zero) and their respective variances (i.e. σ^2_{0i} and σ^2_{1i}). ζ_{0i} and ζ_{1i} are possibly correlated but uncorrelated with ε_{ti}; The variances of ε_{ti} are assumed to be equal and uncorrelated across time although these restrictions can be relaxed (Muthén, 2004).

With the variations of σ^2_{0i} and σ^2_{1i} for the random intercept and linear growth rate, respectively, a 2x2 variance-covariance matrix (ψ) for a linear LCGM can be expressed as follows:

$$\Psi=\begin{bmatrix} \sigma^2_{0i} & \\ \sigma^2_{01i} & \sigma^2_{1i} \end{bmatrix} \tag{3.4}$$

where σ^2_{0i} and σ^2_{1i} are the variances and σ^2_{01i} is the covariance of the random intercept and slope (linear growth rate).

Equation 3.1 can be extended further in order to accommodate nonlinear growth patterns. For example, for a quadratic polynomial LGCM,

$$y_{ti} = \eta_{0i} + \eta_{1i}\lambda_{ti} + \eta_{2i}\lambda^2_{ti} + \varepsilon_{ti} \tag{3.5}$$

$$\eta_{2i} = \alpha_2 + \zeta_{2i} \tag{3.6}$$

where η_{2i} is the quadratic growth rate (also known as *acceleration*), α_2 is the average quadratic rate change overtime, the ζ_{2i} is the variation between individuals in terms of quadratic changes overtime, and λ^2_{ti}s are the quadratic forms of the time scores. With variations between individuals in terms of intercept, linear growth and quadratic growth rates, the 3x3 variance/covariance matrix can expressed as follows:

$$\Psi=\begin{bmatrix} \sigma^2_{0i} & & \\ \sigma^2_{01i} & \sigma^2_{1i} & \\ \sigma^2_{02i} & \sigma^2_{12} & \sigma^2_{2i} \end{bmatrix} \tag{3.7}$$

where σ^2_{0i} , σ^2_{1i} and σ^2_{1i} are the variances, and σ^2_{01i}, σ^2_{02i} and σ^2_{12i} are the covariance between random intercept and random linear growth rate, random intercept and quadratic growth rate, and linear growth rate and quadratic growth rate, respectively.

Similarly, equation 3.1 can also be modified to accommodate multiphase growth patterns. For example, a two phase piecewise LGCM can be formulated as follows (Duncan et al., 2006d).

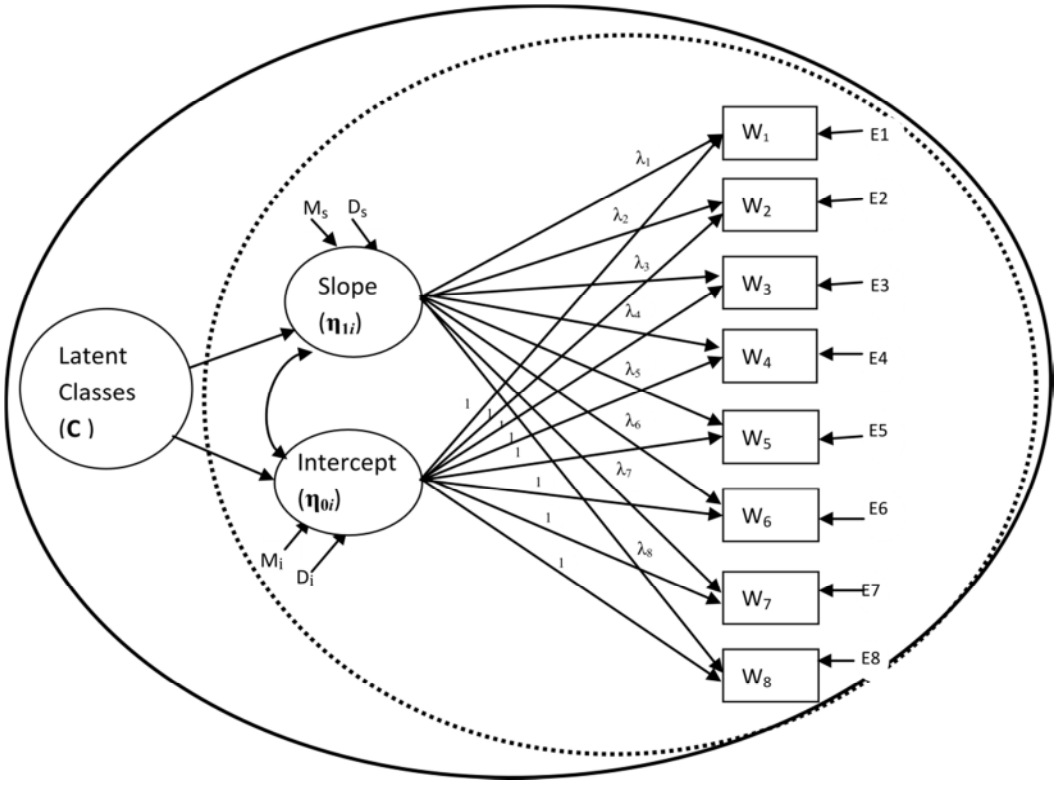
$$y_{ti} = \eta_{0i} + \eta_{1i}\lambda_{1ti} + \eta_{2i}\lambda_{2ti} + \varepsilon_{ti} \quad (3.8)$$

$$\eta_{1i} = \alpha_1 + \zeta_{1i} \quad (3.9)$$

$$\eta_{2i} = \alpha_2 + \zeta_{2i} \quad (3.10)$$

where η_{1i} and η_{2i} are the latent growth factors (linear random slopes) for phase 1 and phase 2, respectively. The λ_{1t} and λ_{2t} are phase 1 and phase 2 time scores, and the ζ_{1i} and ζ_{2i} are the variations between individuals in linear weight changes overtime for phase 1 and phase 2, respectively.

Figure 3.8 Schematic view of linear LGCM and GMM



Latent Growth curve (in dotted circle) and Growth Mixture (in solid circle) models schematic view. W_1-W_8 are repeated measurements of weight at 0, 1, 3, 6, 12, 24 and 36 months, respectively. E1-E8 are error variances of W_1-W_8 respectively. The intercept and slope are two latent (unobserved) growth factors. Arrows from the intercept factor are path coefficients (or *factor loadings*) of the intercept on each time measurement and are fixed at the value of one throughout. Arrows from the slope, represented by lambdas (λ) are path coefficients (or *factor loadings*) of the slope on the repeated measurements and are fixed by the time scores. The double headed arrow between the slope and intercept is the covariance of the latent growth factors. Mi and Ms are the means for the intercept and slope respectively; Di and Ds are the variances for the intercept and slope, respectively. The latent classes (C) are the distinct subpopulations/groups to be identified by the model.

In Figure 3.8, the individual repeated weight measurements and the latent growth factors (i.e., intercept and slope) are common to both models. Therefore, with the latent classes (C) at level three, equation 3.1 can be modified into a GMM as follows (Wang and Wang, 2012a).

$$y^{k}_{ti} = \eta^{k}_{0i} + \eta^{k}_{1i}\lambda^{k}_{ti} + \varepsilon^{k}_{ti} \qquad (3.11)$$

where k is the class number of a latent classes; y^{k}_{ti} is the weight measurement for individual i at time t for the latent class k. The η^{k}_{0i} and η^{k}_{1i} are the two class specific latent growth factors, that is, intercept and slope, respectively. Note also that in equation 3.9, a single-class GMM is equivalent to the LGCM in equation 3.1.

For a linear GMM, with the variations of σ^{2}_{0i} and σ^{2}_{1i} for the random intercept and linear growth rate, respectively, a class-based 2x2 variance-covariance matrix (ψ^{k}) can be expressed as follows:

$$\Psi^{k} = \begin{bmatrix} \sigma^{2}_{0i}{}^{k} & \\ \sigma^{2}_{01i}{}^{k} & \sigma^{2}_{1i}{}^{k} \end{bmatrix} \qquad (3.12)$$

Where $\sigma^{2}_{0i}{}^{k}$ and $\sigma^{2}_{1i}{}^{k}$ are the class-based variances and $\sigma^{2}_{01i}{}^{k}$ is the class-based covariance of the random intercept and slope (linear growth rate).

Note that in Equation 3.12, the variance and covariance of the latent growth factors (i.e. intercept and slope in linear model) are assumed to be different. However, when these parameters are assumed to be equal or are held equal; the variance-covariance structure in Equation 3.12 will reduce to variance-covariance matrix in Equation 3.4.

3.10.2.2 Latent growth model estimation

Prior to conducting model estimations, the following parameterisations were adopted. For a linear LGCM, the intercept factor loadings were fixed at 1, the linear growth rate (i.e. slope) factor loadings were fixed by the time-scores (i.e. 0, 1, 3, 6, 12, 18, 24, and 36), and the means and variance of the intercept and slope were estimated freely. For non-linear LGCMs, similar parametrisations were adopted except that the growth rate factor loadings were fixed by the non-linear form of the time-scores. For example, for a quadratic polynomial model, the growth rate factor

loadings were fixed by the squared form of the time-scores (i.e. 0^2, 1^2, 3^2, 6^2, 12^2, 18^2, 24^2, and 36^2). The same principle was followed for parameterising GMMs. Growth patterns analyses were performed in Mplus software version 7.11 (Muthén and Muthén, 2012).

Model building was carried out in several steps. First, linear LGCMs were fit to the data. Second, in order to account for any nonlinear growth patterns, three nonlinear modelling options: polynomials, piecewise and free-time score functions (Muthén and Muthén, 2012; Wang and Wang, 2012b) were explored and the best fitting function was selected. Third, using the best selected LGCM (e.g. piecewise LGCM), optimal class determination and estimating the effect of covariates on the growth classes were carried out using GMMs.

Latent growth curve models

Polynomial models

Polynomial modelling can be seen as increasing the degrees of the linear growth in order to find the best fit for the observed mean curves. Note that a latent growth factor of a linear model has a degree of 1 (*velocity*), that is, the *factor loadings* are fixed by the linear form of the time scores. For a quadratic model, the latent growth has a degree of 2 (*acceleration*), that is, the *factor loadings* are fixed by the quadratic form of the time scores.

Two polynomial functions were explored: *quadratic* and *cubic* models. First, a model with *intercept*, *linear* and *quadratic slopes* was fitted to the growth data—*quadratic* model. Second, a model that included *intercept*, *linear slope*, *quadratic slope* and *cubic slope* was fitted to the data—*cubic* model. Then, based on model fit statistics, a *cubic* model was found to be more parsimonious or fitted the data better than the quadratic model and was selected as the best among the two models. Assuming weight was measured at 0, 1, 3, 6, 12, 18, 24, and 36 months, the two polynomial models were executed in Mplus software using model commands in Table 3.1. Polynomial models are easy to implement. However, the latent growth factor parameters become difficult to interpret and model fit is poor if growth patterns follow a multimodal shape.

Table 3.1 Polynomial LGCM commands in Mplus software

Polynomial quadratic model:

```
USEVAR=zwei0 zwei1 zwei3 zwei6 zwei12 zwei18 zwei24
zwei36; !list of variables used for growth analysis

MODEL:
I S Q| zwei0@0 zwei1@.1 zwei3@.3 zwei6@.6 zwei12@1.2
zwei18@1.8 zwei24@2.4 zwei36@3.6;
!I =intercept; S=linear slope; Q=quadratic slope; and
|=regressed on
```

Polynomial cubic model:

```
MODEL:
I S Q C| zwei0@0 zwei1@.1 zwei3@.3 zwei6@.6 zwei12@1.2
zwei18@1.8 zwei24@2.4 zwei36@3.6;
!C=cubic slope
```

Free-time score (free-loading) models

In comparing *free-time score* models, two time points were fixed at the time scores and the rest were left to be estimated by the model. Fixing only two of the time points and letting the other time points to be estimated by the model would allow the growth rates (slopes) to be estimated by the model. Initially, the first two time points (i.e. birth and 1 month) were fixed at the time scores (i.e. 0 and 1) and the rest time points were left to be estimated by the model. Then, the first and last time points (i.e. 0 and 36 months) were fixed at the time scores (i.e. 0 and 36) and the rest were left to be estimated by the model. Based on model fit statistics, the model with the first and last time points fixed (i.e. at 0 and 36 months) performed better than the model with first and second time points fixed (i.e. 0 and 1 month). See Mplus free time-score model commands in Table 3.2. Free-time score models are easy to implement, and the latent growth factor parameters are easy to interpret. However, model convergence can be problematic due to the increase of free parameters to be estimated and decrease of degrees of freedom of models.

Table 3.2 Free-time score LGCM commands in Mplus software

Model with first two time-points fixed

```
MODEL:
I S | zwei0@0 zwei1@.1 zwei3* zwei6* zwei12* zwei18*
zwei24* zwei36*;
```

Model with first and last time-points fixed

```
MODEL:
I S | zwei0@0 zwei1* zwei3* zwei6* zwei12* zwei18*
zwei24* zwei36@3.6;
```

Piecewise models

For the *piecewise* models, joints or break points were created in order to get a best fitting model to the data. Fitting a *piecewise* model can be seen as a process of linearization of the growth rates in order to create a flexible predicted line that fits the data best. Thus, positioning of the joints (also known as *knots*) is a key in identifying the best fitting model to the data. In identifying the candidate positions for the *knots*, the following strategy was followed. First, a linear model was fit to the data. Second, a graph that contained the observed and estimated mean curves was outputted for visual inspection. Third, based on the shape of the observed means curve, the time-scores located at positions where the curve showed a considerable change of direction (or bending) were chosen as candidates. Hence, two sets of candidate position were selected: a) 1, 6, and 18 months; and, b) 3 and 12 months. Based on the model fit statistics, 3 and 12 months were the better set of positions for placing the knots. See Mplus piecewise model commands in Table 3.3. Piecewise models are flexible and the latent growth factor parameters become easy to interpret, however, model implementation is difficult.

Table 3.3 Piecewise LGCM commands in Mplus software

Model with knots at 3 and 12 months

```
MODEL:
I S| zwei0@0 zwei1@.1 zwei3@.3 zwei6@.3 zwei12@.3
zwei18@.3 zwei24@.3 zwei36@.3;!phase 1 growth

I S1| zwei0@0 zwei1@0 zwei3@0 zwei6@.3 zwei12@.9
zwei18@.9 zwei24@.9 zwei36@.9;!phase 2 growth

I S2| zwei0@0 zwei1@0 zwei3@0 zwei6@0 zwei12@0 zwei18@.6
zwei24@1.2 zwei36@2.4;!phase 3 growth
```

Model with knots at 1, 6 and 18 months

```
MODEL:
I S| zwei0@0 zwei1@.1 zwei3@.1 zwei6@.1 zwei12@.1
zwei18@.1 zwei24@.1 zwei36@.1;!phase 1 growth

I S1| zwei0@0 zwei1@0 zwei3@.2 zwei6@.5 zwei12@.5
zwei18@.5 zwei24@.5 zwei36@.5;!phase 2 growth

I S2| zwei0@0 zwei1@0 zwei3@0 zwei6@0 zwei12@.6
zwei18@1.2 zwei24@1.2 zwei36@1.2;!phase 3 growth

I S3| zwei0@0 zwei1@0 zwei3@0 zwei6@0 zwei12@0 zwei18@0
zwei24@.6 zwei36@1.8;!phase 4 growth
```

Selection of the best LGCM

Once the visual inspection showed that a linear model was not good fitting model, comparison between the three nonlinear LGCM models (i.e. a cubic *polynomial*, a *free-time scores* model with the first and last times scores fixed, and a *piecewise* model with knots positioned at the 3 and 12 months), was carried out. Best non-linear model was selected based on model fit statistics. See section 4.5 for details.

Fitting of LGCMs to Pakistani and white British growth data

Initially, an overall LGCM was fitted to the data in order to estimate the growth parameters and mean curves for the 1,364 white British and Pakistani children under

the assumption of homogeneity of growth patterns in the of population (Duncan et al., 2006b; Muthén and Muthén, 2012; Wang and Wang, 2012b). Then, a multi-group LGCM was fitted in order to allow ethnic-specific parameters and mean curves to be estimated under the assumption of homogeneity of growth patterns in each ethnic population.

The parameters of the growth variables (i.e., means, variances, and covariances) were freely estimated, initially. However, the residual variances of one weight measurement (weight SDS at 36 months) became negative. As a remedy to that, the variance of the variable was fixed at zero (Chen et al., 2001). There was no dramatic change of other parameter estimates due to the fixing of this parameter, that is, all models still converged well.

Fitting GMMs to Pakistani and white British growth data

In GMMs, estimations of parameters during identification of optimal number classes were performed using two approaches:

a) GMM with growth factors means freely estimated, and growth factors and error variances freely estimated but held equal across classes; and,
b) GMM with growth factors means and variances, and error variances freely estimated; no constraints for variance-covariance structure to be equal across classes.

Using approach (a), the the growth factors means, variances, and covariances were estimated in two class GMM, initially. However, the residual variances of one weight measurement (weight SDS at 36 months) became negative. Like in the LGCMs, the variance of the variable was fixed at zero (Chen et al., 2001), as a remedy. There was no dramatic change of other parameter estimates due to the fixing of this parameter, that is, all models with more than two classes still converged well. See Table 3.4 for a *piecewise* model command in Mplus software.

Table 3.4 Piecewise constrained GMM commands in Mplus software

```
Model with knots at 3 and 12 months
      CLASS= C(n); !n is the number of classes
      MODEL:
            %OVERALL%
            I S| zwei0@0 zwei1@.1 zwei3@.3 zwei6@.3 zwei12@.3
            zwei18@.3 zwei24@.3 zwei36@.3;!phase 1 growth

            I S1| zwei0@0 zwei1@0 zwei3@0 zwei6@.3 zwei12@.9
            zwei18@.9 zwei24@.9 zwei36@.9;!phase 2 growth

            I S2| zwei0@0 zwei1@0 zwei3@0 zwei6@0 zwei12@0
            zwei18@.6 zwei24@1.2 zwei36@2.4;!phase 3 growth
```

In the next step, approach (b) was used to estimate the the growth factors means, variances, and covariances by relaxing the equality of growth factors variances and error variances across classes assumption. When a two-classes GMM was fitted to the data initially, the residual variances of two weight measurement (weight SDS at 3 and 36 months) became negative, the model classification quality (entropy) was very poor <50%. The two error variances were fixed at zero and model was re-ran. However, another negative variance emerged and the model classification quality remained poor. Another attempt was made by increasing the number of random starts. However, no improvement of model convergence was observed. The problems persisted for three classes GMM. Thus, approach (b) was abandoned; it was not used as an option in the determination of optimum number of classes and further models. See Table 3.5 for a piecewise model command in Mplus software.

Table 3.5 Piecewise free GMM commands in Mplus software

```
Model with knots at 3 and 12 months
      CLASS= C(n); !n is the number of classes
      MODEL:
            %OVERALL%
            I S| zwei0@0 zwei1@.1 zwei3@.3 zwei6@.3 zwei12@.3
            zwei18@.3 zwei24@.3 zwei36@.3;!phase 1 growth

            I S1| zwei0@0 zwei1@0 zwei3@0 zwei6@.3 zwei12@.9
            zwei18@.9 zwei24@.9 zwei36@.9;!phase 2 growth

            I S2| zwei0@0 zwei1@0 zwei3@0 zwei6@0 zwei12@0
            zwei18@.6 zwei24@1.2 zwei36@2.4;!phase 3 growth

            %C#1%
            I S S1 S2 zwei0-zwei36;
            %C#2%
            I S S1 S2 zwei0-zwei36;
            .
            .
            .
            .
            %C#n%
            I S S1 S2 zwei0-zwei36;
```

Determination of optimal number of classes

First, 2-9 class models were ran to identify a model with the optimal number of latent classes using a Lo-Mendell-Rubin Likelihood Ratio Test (LMR LRT). Results were then confirmed using Bootstrapped Likelihood Ratio Test (BLRT) tests. However, owing to the high computational time needed for BLRT estimation, only 2-5 class models were re-ran and selected for comparison based on Bayesian Information Criterion (BIC) and classification quality (entropy) values from LMR LRT results.

Estimating the effects of covariates on growth classes

Estimating the effects of covariates was carried out in two steps. First, GMMs were fitted in order to allow for variability (heterogeneity) in the population; see Figure 3.8 for the schematic view of the LGCM and GMM.

Second, after determining the model with the optimal number of classes, the growth models were re-ran by including the covariates using a *three-step* approach (Vermunt, 2010; Asparouhov and Muthén, 2012) in order to estimate the multinomial logistic regression coefficients of the latent classes on the covariates (Table 3.6). See section 3.2.3 for details about the *three-step* approach.

Table 3.6 Piecewise constrained GMM commands in Mplus software

Model with knots at 3 and 12 months

```
AUXILIARY=ethnicity(R3STEP)mother_edu(R3STEP)
   smoking(R3STEP)parity(R3STEP); !list of covariates
   for logistic regression model
CLASS= C(n); !n is the number of classes
MODEL:
      %OVERALL%
      I S| zwei0@0 zwei1@.1 zwei3@.3 zwei6@.3 zwei12@.3
      zwei18@.3 zwei24@.3 zwei36@.3;!phase 1 growth

      I S1| zwei0@0 zwei1@0 zwei3@0 zwei6@.3 zwei12@.9
      zwei18@.9 zwei24@.9 zwei36@.9;!phase 2 growth

      I S2| zwei0@0 zwei1@0 zwei3@0 zwei6@0 zwei12@0
      zwei18@.6 zwei24@1.2 zwei36@2.4;!phase 3 growth
```

3.10.2.3 Missing data estimation

Two missing data estimation approaches were explored: *multiple imputation* and *maximum likelihood*. Growth data of white British and Pakistani children were analysed using LGCM based on MICE and FIML missing data estimation techniques and were compared in terms of model fit statistics. The results indicated

that models based on FIML were more parsimonious than MICE (see Appendix G). Therefore, in all subsequent growth pattern analysis models, missing growth data were estimated using a FIML method in which parameters are estimated based on all available observations in the dataset (Enders and Bandalos, 2001; Schafer and Graham, 2002). FIML is implemented as a default in Mplus software unless a user specifies a list-wise deletion or complete case analysis. Thus, all model estimation Mplus commands listed above have also missing estimation functionality.

3.10.2.4 Model fit evaluation

Latent growth modelling is a complex process. Usually, combinations of a variety of model fit statistics are used during model fit assessments (Duncan et al., 2006a). In fact, there are over 20 available fit indexes to choose from depending on the statistical software that one uses (Duncan et al., 2006a). Therefore, a combination of model fit statistics was used based on Mplus software (Muthén and Muthén, 2012).

When evaluating the goodness of LGCM, Comparative Fit Index (Bentler, 1990), Tucker-Lewis Index (Tucker and Lewis, 1973), Root Mean Square error of approximation (Browne and Cudeck, 1992), Standardised Root Mean Square Residual (Muthén and Muthén, 2012), Log-likelihood (LL), Akaike Information Criterion (Akaike, 1987) and Bayesian Information Criterion (Schwarz, 1978) were used in combination.

When selecting best fitting GMM and optimal number of classes, the Log-likelihood (LL), Akaike Information Criterion (Akaike, 1987), Bayesian Information Criterion (Schwarz, 1978), Bootstrapping Likelihood Ratio Test (McLachlan, 1987) and the classification quality or entropy (Akaike, 1998) model fit statistics were used in combination. In addition, interpretability was also considered where models were rejected if they consist of a class with less ≤1% of the total population.

Attainment of good GMM convergence or global maxima was assumed if the best log-likelihood value was replicated at least 5 times, to be conservative, although a replication of two is also acceptable (Muthén and Muthén, 2012). Where best log-likelihood did not replicate, number of random starts was increased until global maxima was attained or replication of the best log-likelihood was achieved. In significance testing, 5% significance levels and 95% confidence intervals were adopted.

3.10.3 Growth patterns of the BiB1000 cohort children

Prior to investigating the association between childhood growth patterns and wheezing disorders, growth patterns analyses of the BiB1000 children were carried out. The missing data and model estimation techniques carried out in identifying growth patterns of white British and Pakistani children in section 3.10.2 were the basis for the analysis of growth patterns of the BiB1000 children. Therefore, piecewise LGCM and GMM were preferred over other nonlinear models. Model fit indices and information criteria were used in combination, and 5% significance levels and 95% confidence intervals were adopted throughout.

3.10.4 Effects of birthweight, weight at the age of 3 years and childhood growth patterns on wheezing disorders

3.10.4.1 Missing data estimation

Prior to carrying out imputations, a brief check on the variables of analyses showed that birthweight, gestational age and outcome variables (i.e. asthma diagnosis, wheezing symptoms, wheezing disorder treatment and wheezing disorder diagnosis) were completely observed. To further explore if imputations were necessary or beneficial, dummy variable (i.e. yes or no) were created as a missing data indicator for each variable with missing observations.

The missingness indicators and outcome variables were tested for correlations and the results consistently showed that there were no significant associations between them which also indicate that complete cases analysis can produce unbiased, albeit less precise, parameter estimates (Sterne et al., 2009). However, there were consistent significant associations between the missing indicators and other confounding variables which also suggest that imputations with inclusion of these covariates may improve the precision of the parameter estimates (Collins et al., 2001; Sterne et al., 2009).

In the models that investigated the effects of birthweight and weight at the age of 3 years on wheezing disorder, missing data estimation was carried out in Stata software using MICE models. Meaning, the `ice` and `mim` Stata commands were used to multiply impute *m* sets of data and combine results of m analyses using Rubin's rules, respectively (Table 3.7).

Imputations were carried out under MAR assumption that the missingness on outcome variables does not depend on the outcome variables themselves but can be explained by (or related to) other variables included in the imputation models (Collins et al., 2001; Schafer and Graham, 2002; Sterne et al., 2009) In deciding how many datasets to be imputed, the number of imputations (***m)*** were set to be greater than the percentage or fraction of incomplete cases (Graham et al., 2007; Royston and White, 2011).

Table 3.7 Data imputation and analysis commands in Stata software

Model to identify the proportion of individuals with complete information for all variables:

```
Ice ethnicity smoking hypertension diabetes///
imdq_national imdq_bradford parity sex n_births///
birthweight gestational_age asthma_diagnosis///
asthma_trt wheez_diagnosis wheez_treatmentt, dryrun///
seed(10011) clear
```

Imputation model:

```
Ice ethnicity smoking hypertension diabetes///
imdq_national imdq_bradford parity sex n_births///
birthweight gestational_age asthma_diagnosis///
asthma_treatmentt wheez_diagnosis wheez_treatmentt,///
m(40) cycles (30)/// seed(10011) clear
```

Individual analysis of 40 datasets and combining each the results using Rubin’s rules for the birthweight and asthma diagnosis GLM model as an example:

```
xi: mim: glm asthma_diagnosis i.birthweight///
i.ethnicity i.sex i.gestational_age i.nbirths///
 i.smoking i.parity i.imdq_nat,///
fam(bin) link(log) nolog eform
```

Monte Carlo errors for parameter consistency test

```
mim,mcerror
```

Imputation models were chosen based on the type of variables to be imputed, that is, linear regression for continuous variables; logistic regression for binary variables; multinomial logistic regression for unordered and ordered categorical variables were adopted (White et al., 2011). However, note that none of the variables to be imputed was a continuous so no transformation of data was sought to address any non-normality issues.

In childhood growth and wheezing disorder models, missing data estimation was carried out in two stages. First, missing growth data on the BiB1000 children were estimated using FIML (see sections 3.10.2 and 3.10.3) in Mplus software. Then, estimated latent class membership data were saved and merged with the original data. Note that the *classify-analyse* version of the *three-step* approach was preferred over the conventional *three-step* approach used in the analysis of growth classes and their relationship with covariates for the white British and Pakistani children growth data in section 3.10.2.2. This was because, the *classify-analyse* version of the *three-step* allowed for missing data on covariates to be estimated and used for further analyses.

Second, using the merged data, missing data on covariates were estimated using MICE models in Stata software. That is, the `ice` Stata command was utilised to multiply impute *m* sets of data. Then, analyses of the *m* datasets and combining of *m* results was carried out using `mim` Stata command (Table 3.7).

3.10.4.2 Model estimation and evaluation

Generalised Linear Models (Nelder and Baker, 1972) were used to derive the relative risks (RR) where the distribution and link function were specified to be binomial and log respectively. Models were fitted in Stata software version 12. Five per cent significance levels and 95% confidence intervals were adopted throughout.

For all models, consistency of results produced from individuals imputed datasets were assessed using Monte Carlo errors from a jackknife procedure (Royston et al., 2009); where the errors of the beta coefficient (or risk estimate), test statistic (t-value) and the p-value are less than 10% of the standard error, less than 0.1, and less than 0.2 respectively (White et al., 2011), the performances of models were considered as optimal. The `mim,mcerror` post-estimation stata command was utilised to calculate Monte Carlo errors.

CHAPTER 4

RESULTS

4.1 Chapter overview

This chapter is about the results of a series of analyses using BiB cohort data. In section 4.2, results of incidence and burden of allergic diseases analyses in the BiB cohort (13,734 children) are presented and described. Section 4.3 presents results of effects of birthweight on childhood wheezing disorders analysis using 13,734 children data.

Section 4.4 describes results of effects of weight at the age of 3 years on childhood wheezing disorders analysis for 1,598 children. In section 4.5, results of growth patterns analysis of 1,364 children from two ethnic backgrounds (i.e. white British and Pakistani) are presented and described in detail.

Sections 4.6 presents a detailed description of results from effects of childhood growth patterns on childhood growth analysis using 1,598 children growth data. Finally, in section 4.7, comparison between results of complete case and imputed data analyses of section 4.3, 4.4 and 4.6 are made.

4.2 Incidence and burden of wheezing disorders, eczema and rhinitis

The cohort consisted of 13,734 children born between April 2007 and June 2011. There were 5,117 (37.3%) Pakistani and 4,501 (32.8%) white British children; and, 6,917 (50.4%) boys and 6,490 (47.3%) girls (Table 4.1). The cohort yielded a total follow up period of 74,940 person years. The median follow up period was 5.55 years, ranging from 7 days to 7.6 years.

Of the 13,734 cohort children, 140 had missing information on date of censoring. The majority of them were white British (41.4%) and boys (47.9%). The proportion of children with allergic diseases in these 140 children was lower than the overall cohort (Table 4.2)

Table 4.1 Cumulative number of incident cases and percentages for 13,734 BiB cohort children

	Number of children (%)	Allergic conditions (%)		
		Wheezing disorders	Eczema	Rhinitis
Overall	13,734 (100)	3,035 (22.1)	7,192 (52.4)	2,646 (19.3)
Ethnicity				
Pakistani	5,117 (37.2)	1,150 (22.5)	2,995 (58.5)	1,255 (24.5)
White British	4,501 (32.8)	1,074 (23.9)	2,010 (44.7)	543 (12.1)
Other	1,733 (12.6)	308 (17.8)	948 (54.7)	352 (20.4)
Missing	2,383 (17.4)	503 (21.1)	1,239 (52.0)	495 (20.8)
Sex				
Boys	6,917 (50.4)	1,775 (25.7)	3,662 (52.9)	1,445 (20.9)
Girls	6,490 (47.3)	1,190 (18.3)	3,382 (52.1)	1,150 (17.7)
Missing	327 (2.3)	70 (21.4)	148 (45.3)	51 (15.6)
Birth year				
2007	2,082 (15.2)	507 (24.4)	1,085 (52.7)	490 (23.5)
2008	3,669 (26.7)	836 (22.8)	1,935 (53.0	779 (21.2)
2009	3,817 (27.8)	872 (22.8)	2,021 (51.8)	725 (19.0)
2010	3,432 (25.0)	693 (20.2)	1,779 (50.7)	551 (16.1)
2011	734 (5.3)	127 (17.3)	372 (50.7)	101 (13.8)

4.2.1 Wheezing disorders

There were 3,035 incident cases (cumulative incidence = 22.1%, 95% CI: 21.4 to 22.8%) of wheezing disorder during the follow up period (Table 4.1). Of these, 1,422 (47%) were diagnosed during the first 12 months (Table 4.3). There was no significant difference in the cumulative incidence between white British (23.9%; 95% CI: 22.6% to 25%) and Pakistani (22.5%; 95% CI: 21.3% to 23.6%) children. However, boys (25.7%; 95% CI: 24.6% to 26.7%) were more likely to have been diagnosed with wheezing disorders than girls (18.3%; 95% CI: 17.4% to 19.3%) during the follow up period (Table 4.1).

Table 4.2 Characteristics of 140 BiB cohort children with missing information on date of censoring

	Number of children (%)	Allergic conditions (%)		
		Wheezing disorders	Eczema	Rhinitis
Overall	140 (100)	13 (9.3)	26 (18.6)	2 (1.4)
Ethnicity				
Pakistani	22 (15.7)	1 (4.5)	2 (9.1)	0 (0.0)
White British	58 (41.4)	7 (12.1)	12 (20.7)	1(1.7)
Other	20 (14.3)	2 (10.0)	3 (15.0)	0 (0.0)
Missing	40 (28.6.4)	3 (7.5)	9 (22.5)	1 (2.5)
Sex				
Boys	67 (47.9)	7 (10.4)	17 (25.4)	2 (3.0)
Girls	42 (30.0)	6 (14.3)	9 (21.4)	0 (0.0)
Missing	31 (22.1)	0 (0.0)	0 (0.0)	0 (0.0)
Birth year				
2007	26 (18.6)	3 (11.5)	6 (23.1)	1 (3.8)
2008	31 (22.1)	2(6.5)	4 (12.9)	1 (3.2)
2009	35 (25.0)	3(8.6)	4 (11.4)	0 (0.0)
2010	40 (28.6)	5 (12.5)	10 (25.0)	0 (0.0)
2011	8 (5.7)	0 (0.0)	2 (25.0)	0 (0.0)

The overall incidence rate of wheezing disorders was 40.3 (95% CI: 38.9 to 41.8) per 1000 person years (Table 4.4). The rate was significantly higher for boys than girls (incidence rate ratio=1.41; 95% CI: 1.31 to 1.51). However, there was no significant difference between Pakistani and white British children (Incidence rate ratio = 0.94; 95% CI: 0.87 to 1.02), see Table 4.4. Although the cumulative incidence showed substantial decrease between 2007 and 2011 birth years (Table 4.1), there was a considerable increase in the incidence rate during the same period (Table 4.4).

Table 4.3 **Cumulative number of incident cases of allergic conditions for 13,734 BiB cohort children based on the age of children when a diagnosis occurred**

		Wheezing disorders (%)	Eczema (%)	Rhinitis (%)
Age in Years	(0-1]	1,422 (46.8)	5,542 (77.1)	839 (31.7)
	(1-2]	676 (22.3)	1,044 (14.5)	749 (28.3)
	(2-3]	419 (13.8)	348 (4.8)	498 (18.8)
	(3-4]	319 (10.5)	146 (2)	339 (12.8)
	(4-5]	130 (4.3)	74 (1)	140 (5.3)
	(5-6]	49 (1.6)	29 (0.4)	68 (2.6)
	>6	20 (0.6)	9 (0.1)	13 (0.5)
	Total	3,035 (100)	7,192 (100)	2,646 (100)

4.2.2 Eczema

There were a total of 7,192 (cumulative incidence: 52.4%, 95% CI: 51.5 % to 53.2%) childhood eczema incident cases during the follow up period (Table 4.1). 5,542 (77.1%) of these were diagnosed during their first year (Table 4.3). There were more incident cases of Pakistani (58.5%; 95% CI: 57.2% to 59.9%) than the white British children (44.6%; 95% CI: 43.2% to 46.1%). However, there was no significant difference between boys (52.9%; 95% CI: 51.8% to 54.1%) and girls (52.1%; 95% CI: 50.9% to 53.3%), see Table 4.1.

The overall incidence rate of eczema was 95.6 (95% CI: 93.4 to 97.9) per 1000 person years. The rate was significantly higher in Pakistani than the white British (1.31 (95% CI: 1.24 to 1.39), but no significant difference between boys and girls (Table 4.4).

4.2.3 Rhinitis

There were 2,646 incident rhinitis cases (cumulative incidence: 19.3%; 95% CI: 18.6 to 19.9%) during the follow period (Table 4.1). 31.7% of the cases were diagnosed during the first 12 months (Table 4.3). There were more cases of rhinitis of Pakistani and boys than white British and girls, respectively (Table 4.1).

The overall incidence rate of rhinitis in the cohort was 35.3 per 1000 person years (95% CI: 34.0 to 36.7/1000 person years). The incidence rate was higher in Pakistani and boys as compared white British and girls, respectively (Table 4.4)

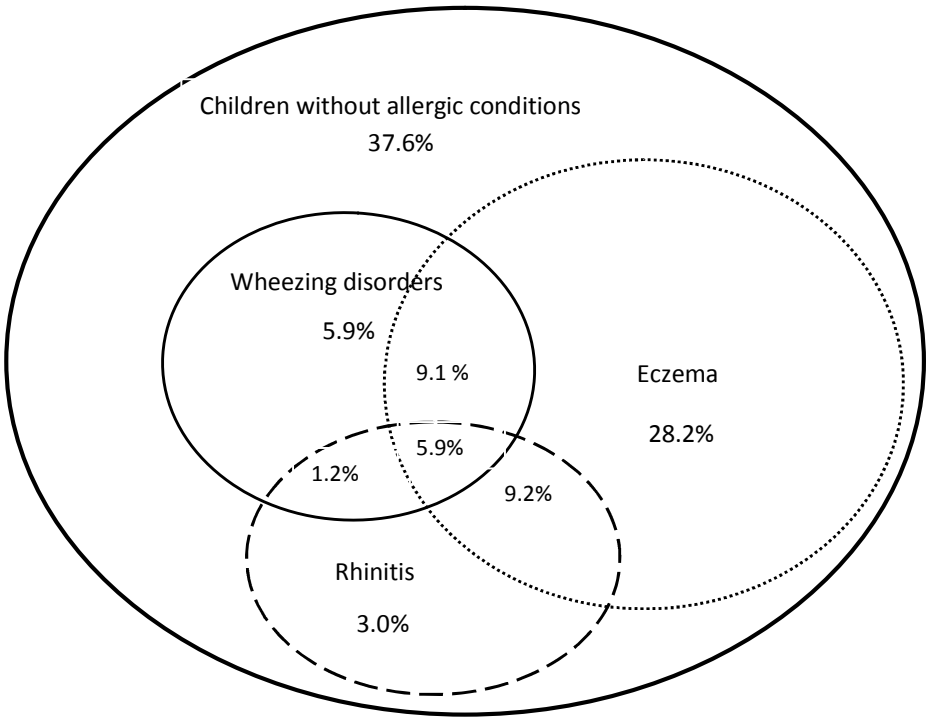
Table 4.4 Age, birth year, ethnicity and sex specific person years and incidence rates of allergic conditions for 13,734 BiB cohort children

	Incidence rate per 1000 person years (95% CI)		
	Wheezing disorders	**Eczema**	**Rhinitis**
Overall	40.3 (38.9 to 41.8)	95.6 (93.4 to 97.9)	35.3 (34.0 to 36.7)
Ethnicity			
White British	43.7 (41.2 to 46.4)	81.9 (78.4 to 85.6)	22.2 (20.4 to 24.2)
Pakistani	41.2(39.0 to 43.7)	107.4 (103.6 to 111.3)	45.0 (42.6 to 47.6)
Pakistani: white British	0.94 (0.87 to 1.03)	1.31 (1.24 to 1.39)	2.03 (1.83 to 2.25)
Sex			
Boys	46.8 (44.7 to 49)	96.5 (93.4 to 99.7)	38.2 (36.3 to 40.2)
Girls	33.3 (31.5 to 35.3)	94.8 (91.7 to 98)	32.3 (30.5 to 34.3)
Boys: girls	1.41 (1.31 to 1.51)	1.02 (0.97 to 1.07)	1.18 (1.09 to 1.28)
Birth Year			
2007	34.6 (31.7 to 37.7)	74.0 (69.7 to 78.6)	33.5 (30.7 to 36.6)
2008	36.3 (34.0 to 38.9)	84.1 (80.5 to 88.0)	33.9 (31.6 to 36.4)
2009	43.8 (41.0 to 46.8)	101.7 (97.3 to 106.2)	36.5 (34.0 to 39.3)
2010	46.4 (43.0 to 50.0)	119.2 (113.8 to 124.9)	37.1 (34.2 to 40.2)
2011	46.4 (39.0 to 55.3)	135.3 (122.2 to 149.8)	36.9 (30.4 to 44.9)

4.2.4 Incidence of multiple allergic conditions

Of the overall cohort of 13,734 children, a total of 5,085 were affected by one of the three allergic conditions; and, a total of 2,682 and 808 children had suffered from two and three allergic conditions at the same time, respectively. The cumulative incidence for only one, two and three allergic conditions was 37% (95% CI: 36.2% to 37.8%), 19.5% (95% CI: 18.9% to 20.2%) and 5.9% (95% CI: 5.5% to 6.3%) respectively (Figure 4.1 and Table 4.5).

Figure 4.1 Cumulative incidence of multiple allergic conditions for 13,734 BiB cohort children



Although Pakistani children were more likely to be diagnosed with one or two conditions than the white British, they were less likely to have three allergic conditions simultaneously. Boys and girls were equally to have been diagnosed from a single condition, but, boys were more likely to be diagnosed from two and three allergic conditions than girls (Table 4.5).

The overall incidence rate for at least one, two and three allergic conditions was 67.5 (95% CI: 65.7 to 69.4), 35.7 (95% CI: 34.4 to 37.1) and 10.8 (95% CI 10.1 to 11.5) per 1000 person years respectively. Boys and Pakistanis were more likely to be affected by multiple allergic conditions than girls and white British children, respectively (Table 4.6). A consistent increase of trend in the incidence rate of single and two allergic conditions between 2007 and 2011 birth years was observed (Table 4.6).

Table 4.5 Cumulative number of multiple incident cases and percentages for 13,734 BiB cohort children during the follow-up period

	Number of children (%)	Allergic conditions (%)		
		One condition	Two conditions	Three conditions
Overall	13,734 (100)	5,085 (37)	2,682 (19.5)	808 (5.9)
Ethnicity				
Pakistani	5,117 (37.3)	1,963 (38.4)	1,144 (22.4)	383 (4.1)
White British	4,501 (32.8)	1,647 (36.6)	711 (15.8)	186 (7.5)
Other	1,733 (12.6)	628 (36.2)	348 (20)	95 (5.5)
Missing	2,383 (17.4)	847 (35.5)	479 (20.1)	144 (6.0)
Sex				
Boys	6,917 (50.4)	2,539 (36.7)	1,456 (21.0)	477 (6.9)
Girls	6,490 (47.3)	2,444 (37.6)	1,174 (18.1)	310 (4.8)
Missing	327 (2.3)	102 (31.2)	52 (15.9)	21 (6.4)
Birth year				
2007	2,082 (15.2)	730 (35.1)	457 (22)	146 (7.0)
2008	3,669 (26.7)	1,340 (36.5)	730 (19.9)	250 (6.8)
2009	3,817 (27.8)	1,438 (37.7)	739 (19.4)	234 (6.1)
2010	3,432 (25.0)	1,294 (37.7)	626 (18.2)	159 (4.6)
2011	734 (5.3)	283 (38.6)	130 (17.7)	19 (2.6)

Table 4.6 Age, birth year, ethnic and sex specific person years and incidence rates of at least one and multiple allergic conditions

	Incidence rate per 1000 person years (95% CI)		
	One condition	**Two conditions**	**Three conditions**
Overall	67.5 (65.7 to 69.4)	35.7 (34.4 to 37.1)	10.8 (10.1 to 11.5)
Ethnicity			
White British	67.2 (64.0 to 70.5)	29.0 (26.9 to 31.2)	7.6 (6.6 to 8.76)
Pakistani	70.3 (67.3 to 73.5)	41.0 (38.7 to 43.5)	13.7 (12.4 to 15.2)
Pakistani: white British	1.05 (0.98 to 1.12)	1.41 (1.29 to 1.56)	1.81 (1.52 to 2.17)
Sex			
Boys	66.8 (64.3 to 69.5)	38.4 (36.5 to 40.5)	12.6 (11.5 to 13.8)
Girls	68.5 (65.8 to 71.3)	32.9 (31.1 to 34.9)	8.7 (7.8 to 9.5)
Boys: Girls	1.02 (0.96 to 1.08)	1.17 (1.08 to 1.26)	1.45 (1.25 to 1.67)
Birth Year			
2007	49.7 (46.2 to 53.5)	31.3 (28.5 to 34.3)	9.9 (8.4 to 11.7)
2008	58.3 (55.2 to 61.5)	31.7 (29.5 to 34.1)	10.9 (9.6 to 12.3)
2009	72.2 (68.6 to 76.1)	37.2 (34.6 to 40.0)	11.8 (10.4 to 13.4)
2010	86.6 (82.0 to 91.5)	42.0 (38.8 to 45.4)	10.7 (9.2 to 12.5)
2011	102.8 (91.4 to 115.5)	47.5 (40.0 to 56.5)	6.9 (4.4 10.9)

4.2.5 Five-year period prevalence

Of the total 13,734 children, 9,079 (66.1%) had a complete follow-up from birth until 5 years. Of these, 3,382 (37.2%) were Pakistani and 2,865 (31.6%) were white British children. 4,590 (50.6%) were boys and 4,298 (47.3%) were girls. Hence, there was no significant difference between the subset and the overall cohort.

Of those 9,079 children, there were a total of 2,135 (23.5%), 4,867 (53.6%) and 1,939 (21.4%) prevalent cases of wheezing disorders, eczema and rhinitis, respectively (Table 4.7). Eczema and rhinitis were more prevalent in Pakistani than the white British children, although there was no significant difference in wheezing disorders. All three allergic conditions were more prevalent in boys than in girls (Table 4.7).

Table 4.7 Five-year period prevalence of allergic conditions in a subset of 9,079 BiB cohort children

	Number of children (%)	Allergic condition prevalence (%)		
		Wheezing disorders	Eczema	Rhinitis
Overall	9,079 (100)	2,135 (23.5)	4,867 (53.6)	1,939 (21.4)
Ethnicity				
Pakistani	3,382 (37.3)	826 (24.4)	2,041 (60.3)	909 (26.9)
White British	2,865 (31.6)	703 (24.5)	1,292 (45.1)	394 (13.8)
Other	1,093 (12.0)	214 (19.6)	617 (56.5)	244 (22.3)
Missing	1,739 (19.1)	392 (22.7)	917 (52.7)	392 (29.1)
Sex				
Boys	4,590 (50.6)	1,234 (26.9)	2,489 (54.2)	1,067 (23.2)
Girls	4,298 (47.3)	859 (20.0)	2,288 (53.2)	833 (19.4)
Missing	191 (2.1)	42 (22.0)	90 (47.1)	39 (20.4)

4.3 Effects of birthweight on childhood wheezing disorders

Demographics

The cohort was made up of 13,734 children that yielded 74,940 person years of follow-up. In total, 37.3% and 32.8% were Pakistani and white British origin respectively; 12.6% were minority and 17.3% with missing ethnicity data. In total, 50.4% and 47.3% were boys and girls respectively, and, 2.3% of children had missing information on sex. In total, 82.6%, 9.1% and 8.3% of the cohort were 'normal', 'high' and 'low' birthweight children respectively (Table 4.8). Data on birthweight, gestational age, and wheezing disorders were complete. Approximately, 23% of the total children had missing information on at least one covariate (Table 4.8).

Out of 13,734 children, 841 (6.1%) were diagnosed as asthmatic, 1994 (14.5%) had wheezing symptoms, 2347 (17.1%) were either diagnosed for asthma or had wheezing symptoms, and 3035 (22.1%) children were treated with asthma drugs based on primary care data available up to November 2014 (Table 4.8).

Selecting minimally sufficient sets of confounding variables

DAG model output using DAGitty software resulted in two sets of *minimally sufficient* confounding variables (Figure 4.2). That is, each of the two sets contains *minimally sufficient* set of confounding variables. Since the list of variables in the DAG model was constructed retrospectively (i.e. after the BiB cohort data were collected), comparison between *minimally sufficient* sets was mainly based on the availability of data about the variables within each set. That is, a set that has a lowest number of variables without data was considered to be as the best option. Thus, the set that contains: ethnicity, family asthma, gender, gestational age, maternal smoking, number of live births, parity, and SES was selected. However, information on family asthma was missing from the BiB cohort data so models were not adjusted for this variable.

Low birthweight and wheezing disorders

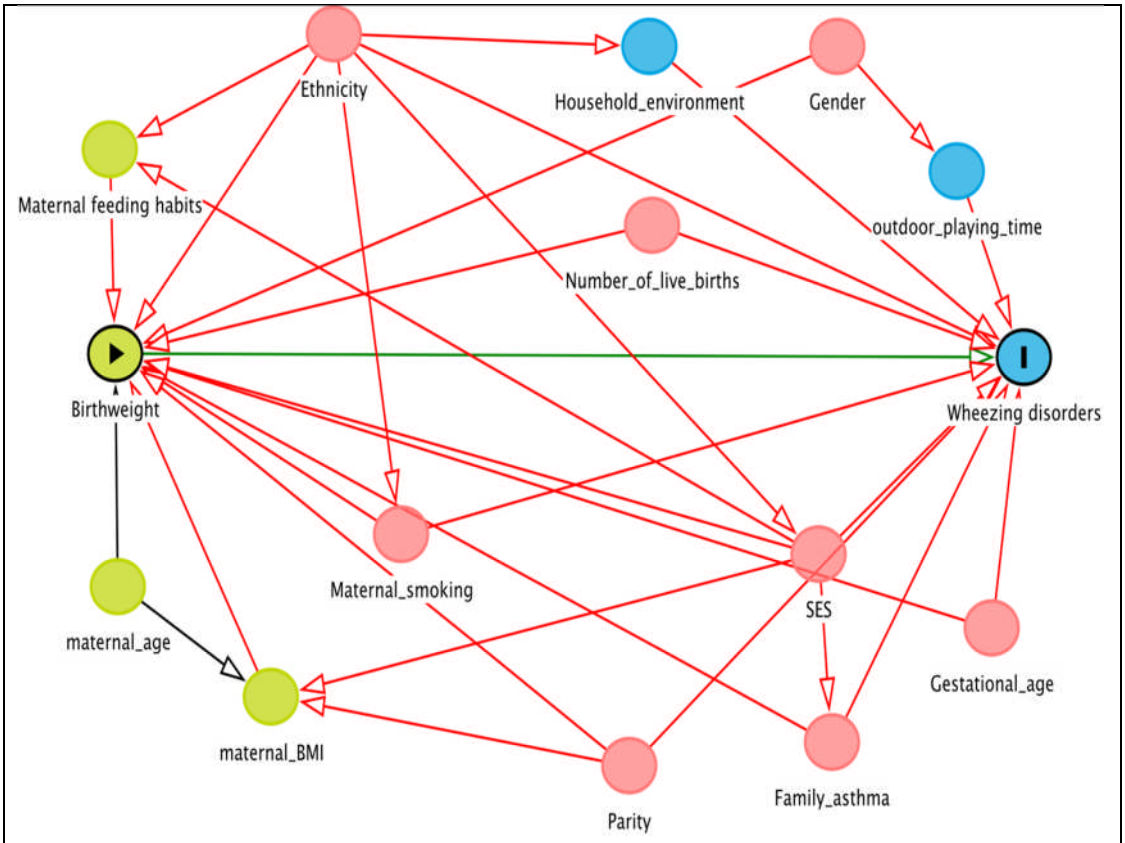
There was a significant increased risk of wheezing disorders for low birthweight children in all four disease definitions (Table 4.9). The unadjusted RRs for *asthma diagnosis*, *wheezing symptoms*, *wheezing disorder diagnosis* and *wheezing disorder treatment* 1.55 (95% CI: 1.27 to 1.89), 1.28 (95% CI: 1.13 to 1.46), 1.28 (95% CI:

1.14 to 1.45) and 1.27 (95% CI: 1.15 to 1.40). The respective adjusted RRs 1.53 (95% CI: 1.20 to 1.96), 1.29 (95% CI: 1.10 to 1.52), 1.29 (95% CI: 1.12 to 1.50) and 1.25 (95% CI: 1.10 to 1.42).

High birthweight and wheezing disorders

Based on the adjusted risk estimates, there was a consistent but non-significant reduction of risk in all four wheezing disorders disease definitions for high birthweight children (Table 4.9). The unadjusted RRs for *asthma diagnosis*, *wheezing symptoms*, *wheezing disorder diagnosis* and *wheezing disorder treatment* were 0.93 (95% CI: 0.73 to 1.19), 0.91 (95% CI: 0.78 to 1.06), 0.92 (95% CI: 0.80 to 1.05) and 1.04 (95% CI: 0.93 to 1.16). The respective adjusted RRs were 0.95 (95% CI: 0.75 to 1.22), 0.90 (95% CI: 0.77 to 1.04), 0.91(95% CI: 0.79 to 1.04) and 0.99 (95% CI: 0.89 to 1.11).

Figure 4.2 **DAG model output of confounding adjustment for models that investigated the effects of birthweight on wheezing disorders**



Key:

=Exposure variable; =Outcome variable; =Ancestor of exposure and outcome; = Ancestor of outcome; =Causal path; =Biasing path

Minimal sufficient adjustment sets for estimating the total effect of birthweight on wheezing disorders:

- Ethnicity, family asthma, gender, gestational age, maternal smoking, number of live births, parity and SES.
- Ethnicity, family asthma, gestational age, maternal smoking, number of live births, parity, SES and outdoor playing time.

Table 4.8 Characteristics of 13,734 children with complete data on wheezing disorders and covariates

	Asthma diagnosis		Wheezing symptoms		Wheezing disorder diagnosis		Wheezing disorder treatment	
	Yes/ no	Yes (%)	Yes/ no	Yes (%)	Yes/ no	Yes (%)	Yes/ no	Yes (%)
Birthweight (kg)								
Normal (2.5-4.0)	668/10,673	5.9	1,622/9,719	14.3	1,907/9,434	16.8	2,444/8,897	21.6
Low (<2.5)	104/1,035	9.1	209/930	18.3	246/893	21.6	311/828	27.3
High (>4.0)	69/1,185	5.5	163/1,091	13.0	194/1,060	15.5	280/974	22.3
Ethnicity								
White British	217/4,284	4.8	586/3,915	13.1	706/3,795	15.7	1,074/3,427	23.9
Pakistani	382/4,735	7.5	857/4,260	16.7	985/4,132	19.2	1,150/3,967	22.5
Others	86/1,647	5.0	207/1,526	11.9	243/1,490	14.0	308/1,425	17.8
Gender								
Male	502/6,415	7.3	1,220/5,697	17.6	1,416/5,501	20.5	1,775/5,142	25.7
Female	318/6,172	4.9	742/5,748	11.4	890/5,600	13.7	1,190/5,300	18.3
Gestational age								
Term	769/12,100	6.0	1,841/11,028	14.3	2,166/10,703	16.8	2,792/10,077	21.7
Pre-term	72/793	8.3	153/712	17.7	181/684	20.9	243/622	28.1
Number of births								
Singleton	803/12,281	6.1	1,923/11,161	14.7	2,262/10,822	17.3	2,911/10,173	22.2
Twins	17/297	5.4	38/276	12.1	43/271	13.7	52/262	16.6
Triplets	0/9	0	1/8	11.1	1/8	11.1	2/7	22.2

	Asthma diagnosis		Wheezing symptoms		Wheezing disorder diagnosis		Wheezing disorder treatment	
	Yes/ no	Yes (%)	Yes/ no	Yes (%)	Yes/ no	Yes (%)	Yes/ no	Yes (%)
Maternal smoking								
No	520/7,371	6.6	1,162/6,729	14.7	1,359/6,532	17.2	1,710/6,181	21.7
Yes	167/3,295	4.8	490/2,972	14.2	578/2,884	16.7	823/2,639	23.8
Parity								
Primiparous	292/4,823	5.7	686/4,429	13.4	821/4,294	16.1	1,128/3,987	22.1
Multiparous	489/7,311	6.3	1,210/6,590	15.5	1,401/6,399	18.0	1,728/6,072	22.2
IMD 2010 quintile score								
1	487/7,048	6.5	1,182/6,353	15.7	1,372/6,163	18.2	1,721/5,814	22.8
2	115/1,939	5.6	253/1,801	12.3	304/1,750	14.8	435/1,619	21.2
3	59/1,196	4.7	148/1,107	11.8	177/1,078	14.1	247/1,008	19.7
4	18/317	5.4	41/294	12.2	53/282	15.8	84/251	25.1
5	8/184	4.2	30/162	15.6	33/159	17.2	49/143	25.5

IMD=index of multiple deprivation at a national level with 1 and 5 indicating the least and most deprived scores respectively.

Table 4.9 Adjusted relative risks and 95% confidence intervals of covariates using 40 imputed datasets of 13,734 BiB cohort children

	Asthma diagnosis	Wheezing symptoms	Wheezing disorder diagnosis	Wheezing disorder treatment
Birthweight (kg)				
Normal (2.5-4.0)	1	1	1	1
High (>4.0)	0.95 (0.75 to 1.22)	0.90 (0.77 to 1.04)	0.91(0.79 to 1.04)	0.99 (0.89 to1.11)
Low (<2.5)	1.53 (1.20 to 1.96)	1.29 (1.10 to 1.52)	1.29 (1.12 to 1.50)	1.25(1.10 to 1.42)
Ethnicity				
White British	1	1	1	1
Pakistani	1.36 (1.11 to 1.66)	1.26(1.12 to 1.42)	1.21(1.08 to 1.35)	0.95 (0.87 to 1.05)
Others	0.96 (0.74 to 1.25)	0.93 (0.79 to 1.08)	0.90 (0.78 to 1.04)	0.76 (0.67 to 0.85)
Gender				
Male	1	1	1	1
Female	0.67(0.58 to 0.76)	0.64 (0.59 to 0.70)	0.66 (0.61 to 0.72)	0.71 (0.67 to 0.76)
Gestational age				
Term	1	1	1	1
Pre-term	1.11(0.83 to 1.48)	1.08 (0.90 to 1.30)	1.09 (0.92 to 1.29)	1.16 (1.01 to 1.34)
Number of births				
Singleton	1	1	1	1
Twins	0.68(0.42 to 1.10)	0.71 (0.52 to 0.97)	0.68 (0.51 to 0.90)	0.63 (0.49 to 0.81)
Triplets	-	0.57 (0.09 to 3.60)	0.48 (0.08 to 3.03)	0.75 (0.22 to 2.56)
Maternal smoking				
No	1	1	1	1
Yes	0.86(0.70 to 1.05)	1.10 (0.98 to 1.24)	1.07 (0.97 to 1.19)	1.05 (0.97 to 1.15)

	Asthma diagnosis	Wheezing symptoms	Wheezing disorder diagnosis	Wheezing disorder treatment
Parity				
Primiparous	1	1	1	1
Multiparous	1.04 (0.91 to 1.20)	1.14 (1.04 to 1.24)	1.10 (1.02 to 1.19)	1.02 (0.95 to 1.08)
IMD 2010 quintile score	0.96 (0.88 to 1.05)	0.95 (0.90 to 1.00)	0.95 (0.91 to 1.00)	0.97 (0.93 to 1.00)

4.4 Effects of weight at the age of 3 years on wheezing disorders

Demographics

A total of 1,598 BiB1000 singletons, term children were included in this analysis. They were 778 (48.7%) boys and 820 (51.3%) girls; and 762 (47.7%) Pakistani and 602 (37.7%) white British (Table 4.10). A total of 1,043 (65.3%) had complete weight data for the 36 months questionnaire. The age of the children ranged between 35.4 and 40.6 months.

The total number of children who had 'asthma' diagnosis, 'wheezing' symptoms, 'wheezing disorders' diagnosis and 'wheezing disorders' treatment were 113 (7.1%) , 252 (15.8%), 300 (18.8%) and 369 (23.1%) respectively, slightly higher than the BiB cohort (Table 4.10).

Selecting minimally sufficient sets of confounding variables

DAG model output using DAGitty software resulted in three sets of *minimally sufficient* confounding variables (Figure 4.3). Construction of the list of variables that go into the DAG model was carried out retrospectively (i.e. after the BiB cohort data were collected). That is, comparison between *minimally sufficient* sets was mainly based on the availability of data about the variables of each set. Thus, birthweight, breast feeding, ethnicity, family asthma, gender, maternal smoking, parity, and SES were selected as *minimally sufficient* set of confounding variables. However, information on breast feeding and family asthma was not available so models were not adjusted for these variables.

Weight at the age of 3 years and wheezing disorders

Underweight children had an insignificant reduced risk of wheezing disorders (Table 4.11). The unadjusted RRs and 95% confidence intervals of underweight for *asthma diagnosis*, *wheezing symptoms*, *wheezing disorders diagnosis* and *wheezing disorder treatment* were 0.58 (0.23 to 1.44), 0.97 (0.60 to 1.57), 0.89 (0.57 to 1.39) and 0.79 (0.51 to 1.23), respectively when compared with the normal weight children. The respective adjusted RRs and 95% CI underweight children were 0.55 (0.22 to 1.38), 0.95 (0.59 to 1.53), 0.87 (0.56 to 1.37) and 0.78 (0.50 to 1.21)

Overweight children showed an insignificant increase in the risk of wheezing symptoms, wheezing disorder diagnosis and wheezing disorder treatment whilst

insignificant decreased risk of asthma diagnosis when compared with the normal weight children (Table 4.11). The unadjusted RRs and 95% confidence intervals of underweight for *asthma diagnosis*, *wheezing symptoms*, *wheezing disorders diagnosis* and *wheezing disorder treatment* 0.79 (0.40 to 1.55), 1.32 (0.89 to 1.96), 1.23 (0.86 to 1.76) and 1.14 (0.82 to 1.58). The respective adjusted RRs and 95% CI were 0.85 (0.43 to 1.65), 1.31 (0.88 to 1.94), 1.23 (0.86 to 1.75) and 1.12 (0.80 to 1.57).

Obese children had an insignificant increased risk of wheezing disorders (Table 4.11). The unadjusted RRs and 95% confidence intervals of underweight for *asthma diagnosis*, *wheezing symptoms*, *wheezing disorders diagnosis* and *wheezing disorder treatment* were 1.29 (0.66 to 2.50), 0.98 (0.58 to 1.66), 1.10 (0.72 to 1.76) and 1.14 (0.78 to 1.67). And, the respective adjusted RRs and 95% CI were 1.31 (0.67 to 2.56), 1.01 (0.60 to 1.70), 1.12 (0.73 to 1.70) and 1.14 (0.80 to 1.57).

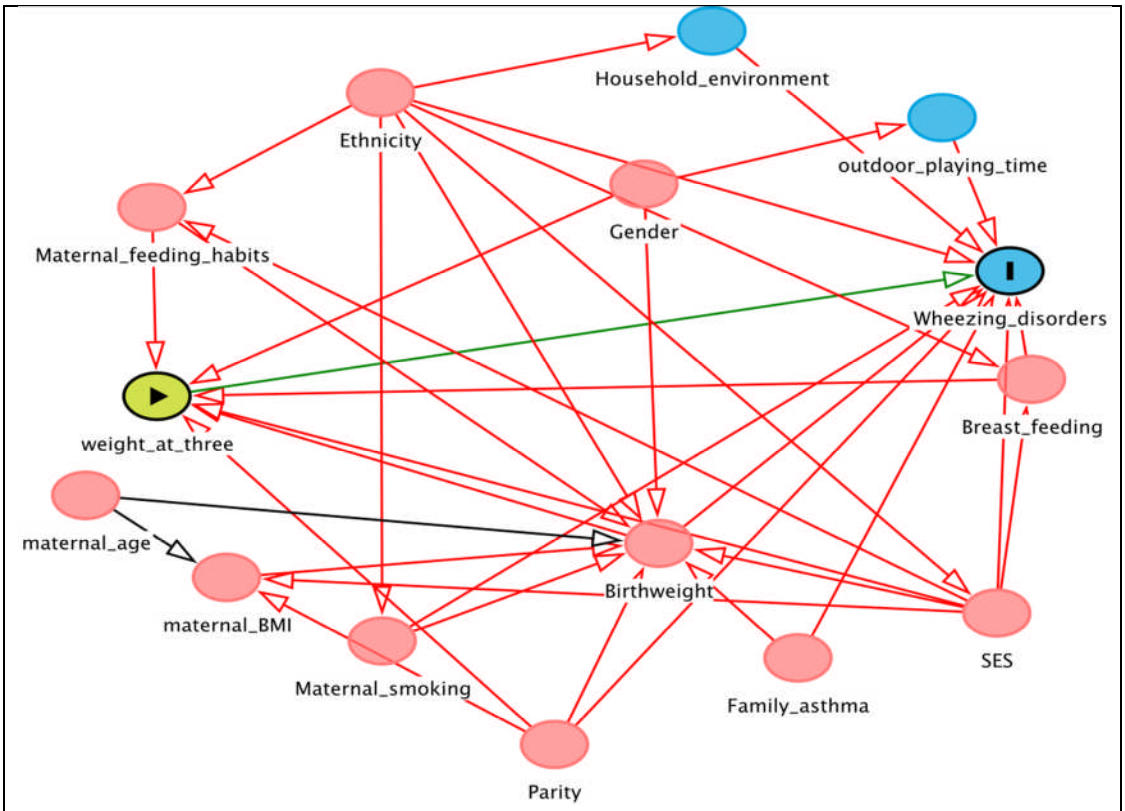
Table 4.10 Characteristics of 1,598 BiB1000 children with complete data on wheezing disorders and covariates

	Asthma diagnosis		Wheezing symptoms		Wheezing disorder diagnosis		Wheezing disorder treatment	
	Yes/ no	Yes (%)	Yes/ no	Yes (%)	Yes/ no	Yes (%)	Yes/ no	Yes (%)
Birthweight (kg)								
Normal (2.5-4.0)	101/1,314	7.1	221/1,194	15.6	264/1,151	18.7	321/ 1,094	22.7
Low (<2.5)	6/64	8.6	14/56	20.0	16/54	22.9	20/50	28.6
High (>4.0)	6/107	5.3	17/96	15.0	20/93	17.7	28/85	24.8
Weight at 3 years								
Normal (≥5th and <85th centiles)	66/667	9.0	117/616	16.0	146/587	19.9	185/548	25.2
Underweight (<5th centile)	5/92	5.2	15/82	15.5	17/80	17.5	19/78	19.6
Overweight (≥85th and <95th centiles	9/118	7.0	26/101	20.5	30/97	23.6	35/92	27.6
Obese (≥95thcentile)	10/76	11.6	13/73	15.2	17/80	22.1	25/61	29.1
Ethnicity								
White British	24/578	4.0	82/520	13.6	95/507	15.8	141/461	23.4
Pakistani	73/689	9.6	134/628	17.6	164/598	21.5	175/587	23.0
Others	16/216	6.9	36/196	15.5	41/191	17.7	53/179	22.8
Gender								
Male	70/708	9.0	159/619	20.4	185/593	23.8	212/566	27.2
Female	43/777	5.2	93/727	11.3	115/705	14.0	157/663	19.1
Maternal smoking								
No	90/1,051	7.9	177/964	15.5	213/928	18.7	256/885	22.4
Yes	23/433	5.0	74/382	16.2	86/370	18.9	112/344	24.6

	Asthma diagnosis		Wheezing symptoms		Wheezing disorder diagnosis		Wheezing disorder treatment	
	Yes/ no	Yes (%)	Yes/ no	Yes (%)	Yes/ no	Yes (%)	Yes/ no	Yes (%)
Parity								
Primiparous	41/571	6.7	87/ 525	14.2	106/ 506	17.3	144/468	23.5
Multiparous	70/ 892	7.3	163/ 799	16.9	191/ 771	19.9	218/744	22.7
IMD 2010 quintile score								
1	83/ 998	7.7	183/898	16.9	217/864	20.1	255/826	23.6
2	19/ 271	6.6	37/253	12.8	45/ 245	15.5	64/226	22.1
3	10/ 158	6.0	23/ 145	13.7	28/140	16.7	36/132	21.4
4	1/34	2.9	3/32	8.6	4/31	11.4	6/ 29	17.1
5	0/24	0	6/18	25.0	6/18	25.0	8/16	33.3

IMD=Index of multiple deprivation at national level with 1 and 5 indicating the least and most deprived scores respectively.

Figure 4.3 **DAG model output of confounding adjustment for models that investigated the weight at 3 years and risk of wheezing disorders**



Key:

=Exposure variable; =Outcome variable; =Ancestor of exposure and outcome; = Ancestor of outcome; =Causal path; =Biasing path

Minimal sufficient adjustment sets for estimating the total effect of weight at the age of 3 years on wheezing disorders:

- Birthweight, breast feeding, ethnicity, family asthma, gender, maternal smoking, parity and SES.
- Birthweight, breast feeding, ethnicity, family asthma, maternal smoking, parity, SES and outdoor playing time.
- Birthweight, breast feeding, gender, maternal feeding habits, parity and SES.

Table 4.11 Adjusted relative risks and 95% confidence intervals of covariates using 40 imputed datasets of 1,598 BiB1000 children

	Asthma diagnosis	Wheezing symptoms	Wheezing disorder diagnosis	Wheezing disorder treatment
Weight at 3 years				
Normal (≥5th and <85th centiles)	1	1	1	1
Underweight (<5th centile)	0.55 (0.22 to 1.38)	0.95 (0.59 to 1.53)	0.87 (0.56 to 1.36)	0.78 (0.50 to 1.21)
Overweight (≥85th and <95th centiles)	0.85 (0.43 to 1.65)	1.31 (0.88 to 1.94)	1.23 (0.86 to 1.75)	1.12 (0.80 to 1.57)
Obese (≥95thcentile)	1.31 (0.67 to 2.56)	1.01 (0.60 to 1.70)	1.12 (0.73 to 1.70)	1.14 (0.80 to 1.57)
Birthweight (kg)				
Normal (2.5-4.0)	1	1	1	1
High (>4.0)	0.83 (0.37 to 1.86)	0.95 (0.60 to 1.50)	0.94 (0.62 to 1.43)	1.04 (0.75 to 1.47)
Low (<2.5)	1.27 (0.57 to 2.83)	1.34 (0.83 to 2.17)	1.28 (0.82 to 2.00)	1.34 (0.91 to 1.97)
Ethnicity				
White British	1	1	1	1
Pakistani	2.23 (1.32 to 3.75)	1.37 (1.02 to 1.85)	1.41 (1.08 to 1.85)	1.04 (0.83 to 1.31)
Others	1.65 (0.87 to 3.12)	1.19 (0.81 to 1.74)	1.15 (0.81 to 1.62)	1.01 (0.75 to 1.35)
Gender				
Male	1	1	1	1
Female	0.57 (0.39 to 0.82)	0.56 (0.44 to 0.71)	0.59 (0.47 to 0.73)	0.71 (0.59 to 0.85)
Maternal smoking				
No	1	1	1	1
Yes	0.97 (0.52 to 1.81)	1.39 (1.01 to 1.91)	1.25 (0.93 to 1.69)	1.17 (0.91 to 1.50)
Parity				
Primiparous	1	1	1	1
Multiparous	0.97 (0.66 to 1.41)	1.16 (0.90 to 1.48)	1.10 (0.88 to 1.37)	0.97 (0.81 to 1.17)

	Asthma diagnosis	Wheezing symptoms	Wheezing disorder diagnosis	Wheezing disorder treatment
IMD 2010 Quintile score	0.89 (0.69 to 1.14)	0.95 (0.82 to 1.10)	0.95 (0.83 to 1.08)	0.97 (0.87 to 1.09)

4.5 Describing the growth patterns of white British and Pakistani children

Demographics

There were a total of 1,364 singletons, term children with 48.5% boys and 51.5% girls; 44% of white British and 56% of Pakistani origin, that is, 602 white British children (293 boys and 309 girls), and 762 Pakistani children (368 boys and 394 girls).

Growth characterstics

Over all, missing data was substantial for the ages of 3 and 24 months (Table 4.13). Although the missing rate for the two ethnicities was similar during the first three periods of measurements (i.e. birth, 1 and 3 months), it was slightly better during 12, 18 and 24 months but slightly worse during 6 and 36 months for the Pakistani children when compared with the white British.

The overall observed means at birth, 1 month and 3 months were below the 50^{th} centile whereas from 6 months onwards above the 50^{th} centile (Table 4.13), according to the WHO growth standards (WHO, 2006). The correlation among the repeated weight measurements was between 0.346 and 0.934 (Table 4.13). The covariance coverage (the proportion of data present in variable x given variable y) was between 0.089 (9%) and 1.00 (100%). This is a reflection of complete case covariance matrix. Given that Mplus software was used for analyses of growth patterns, this information was used for fine tunning of the "coverage" in the analysis command. The default minimum covariance coverage value in Mplus is 0.10, thus, the covariance coverage of the GMMs had to be fixed at the lowest covariance coverage value (i.e. 0.089) of the data. Otherwise, the models would not converge.

Figure 4.4 depicts individual observed growth trajectories of 1,364 children; and Figure 4.5A and Figure 4.5B illustrate individual growth trajectories of Pakistani and white British children, respectively. Approximately, 95% of the of the children had weight SDSs between +2.5 and -3.5 throughout the follow up period (Figure 4.4). The Pakistani children had a larger range of birthweight (+3.8SDS to -4.4SDS) than the white British (+3SDS to -3.7SDS) although the difference was attenuated by the age of three years (Figure 4.5A and Figure 4.5B)

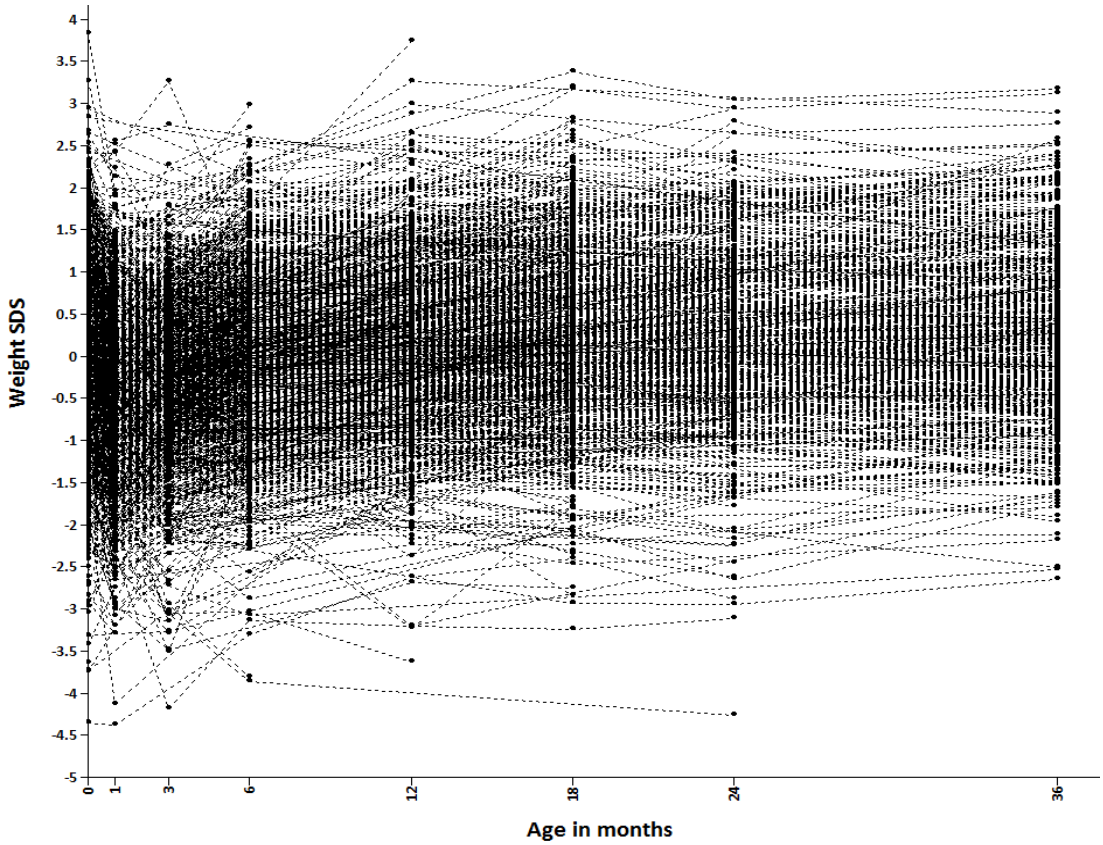
Table 4.12 Complete weight measurements for 1,364 Pakistani and white British children

	Pakistani (%)	white British (%)	All (%)
Birth	762 (100)	602 (100)	1364 (100)
1 month	519 (68.1)	426 (70.8)	945 (69.3)
3 months	190 (24.9)	155 (25.7)	345 (25.3)
6 months	425 (55.8)	361 (60.1)	786 (57.6)
12 months	363 (47.3)	262 (43.5)	624 (45.7)
18 months	387 (50.8)	262 (43.5)	649 (47.6)
24 months	257 (33.7)	184 (30.6)	441 (32.3)
36 months	264 (34.6)	270 (44.9)	534 (39.1)

Table 4.13 Observed means, correlations and covariance coverage of repeated weight SDS measurements of 1,364 Pakistani and white British children

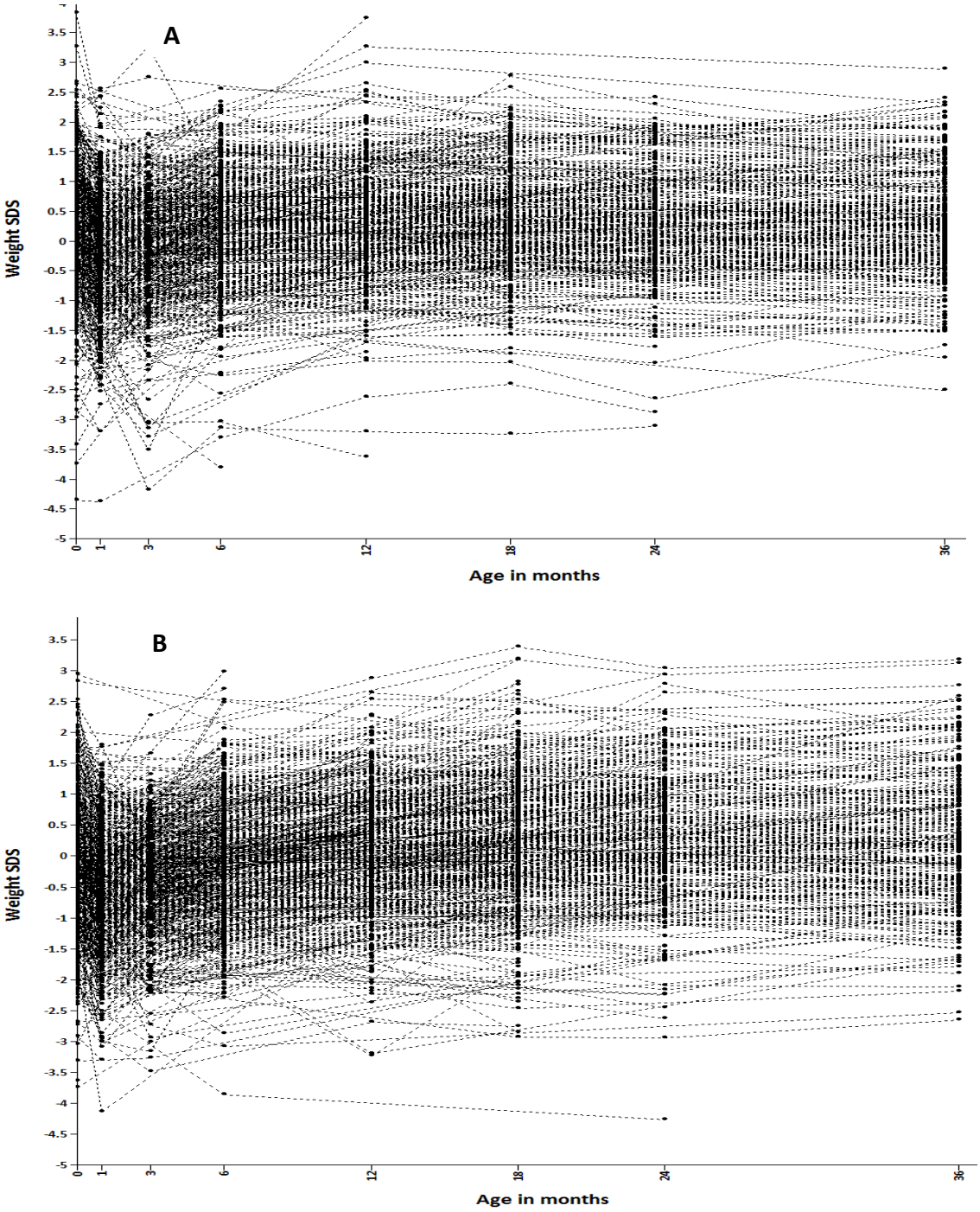
Months	Mean	Correlation Matrix							
		0	1	3	6	12	18	24	36
0	-0.087	1.00							
1	-0.426	0.825	1.00						
3	-0.412	0.588	0.786	1.00					
6	0.015	0.469	0.611	0.853	1.00				
12	0.237	0.415	0.514	0.738	0.889	1.00			
18	0.252	0.410	0.489	0.688	0.815	0.914	1.00		
24	0.272	0.378	0.456	0.656	0.755	0.850	0.934	1.00	
36	0.240	0.346	0.447	0.612	0.709	0.795	0.867	0.913	1.00
		Covariance coverage							
0		1.00							
1		0.693	0.693						
3		0.253	0.226	0.253					
6		0.576	0.420	0.156	0.576				
12		0.457	0.328	0.121	0.303	0.457			
18		0.476	0.346	0.130	0.309	0.284	0.476		
24		0.323	0.231	0.089	0.211	0.183	0.213	0.323	
36		0.391	0.284	0.106	0.271	0.223	0.243	0.179	0.391

Figure 4.4 Individual observed growth trajectories of 1,364 Pakistani and white British children



Note: X and Y axes denote age of children when measurtement was recorded and the standardised weight scores, respectively.

Figure 4.5 **Individual observed growth trajectories of Pakistani (A) and white British (B) children**



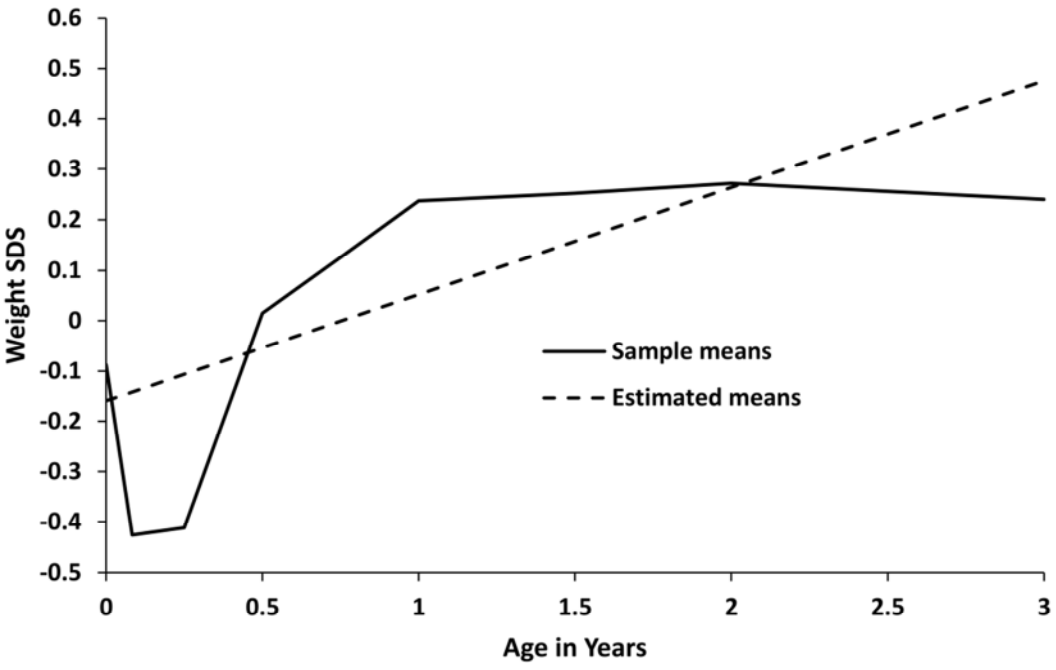
Note: Both in figre A and B, X and Y axes denote age of children when measurtement was recorded and the standardised weight scores, respectively.

4.5.1 Latent growth curve model

4.5.1.1 The linear model

According to the linear LGCM results (Table 4.14), white British and Pakistani children had an overall estimated mean of intercept (i.e. birthweight) that fell just below the 44th centile and shifted up by 0.176 SDS at every subsequent month (i.e. slope or velocity). The latent growth factors (i.e. slope and intercept) had a statistically significant inverse relationship (i.e. Covariance = -0.039; P-value = 0.01). However, as can be noted from the model fit statistics results (Table 4.14) and the graph of sample and estimated means curves (Figure 4.6), the linear LGCM fitted the data poorly. For example, the chi-squared value was very large (1765, df = 31); the Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) were lower than the recommended cut off values; and Root Mean Square Error Approximation (RMSEA) and Standardised Root Mean square Residuals (SRMR) were above the cut off values. Likewise, there was a substantial gap (residual) between the estimated means and sample means lines (Figure 4.6).

Figure 4.6 Sample and estimated mean curves of linear latent growth curve model



Note: X and Y axes denote age of children when measurtement was recorded and the standardised weight scores, respectively.

Table 4.14 Parameter estimates and model fit statistics of linear latent growth curve model

	Estimate	Standard error	P-value
Means			
Intercept	-0.159	0.028	<0.01
Slope	0.176	0.016	<0.01
Covariance			
Intercept ↔ slope	-0.039	0.014	0.01
Variances			
Intercept	0.712	0.037	<0.01
Slope	0.083	0.013	<0.01
E1 (birth)	0.561	0.044	<0.01
E2 (1 month)	0.407	0.043	<0.01
E3 (3 months)	0.353	0.045	<0.01
E4 (6 months)	0.260	0.027	<0.01
E5 (12 months)	0.211	0.026	<0.01
E6 (18 months)	0.119	0.017	<0.01
E7 (24 months)	0.073	0.018	<0.01
E8 (36 months)	0.224	0.058	<0.01
Model fit statistics			
AIC	13,214		
BIC	13,282		
CFI	0.588		
-2LL	13,188		
RMSEA	0.203		
SRMR	0.0158		
TLI	0.628		
Chi-squared statistic	1765 (df=31)		<0.01

AIC=Akaike Information Criterion; BIC= Bayesian Information Criterion; RMSEA= Root Mean Square Error Approximation; CFI=Comparative Fit Index; TLI=Tucker-Lewis Index; SRMR=Standardised Root Mean Square Residuals and -2 LL= -2 X Log-likelihood; df=degrees of freedom. Recommended cut-off values: CFI>0.95; TLI>0.95; RMSEA≤0.05; and SRMR<0.05.

4.5.1.2 Comparison between non-linear latent growth curve models

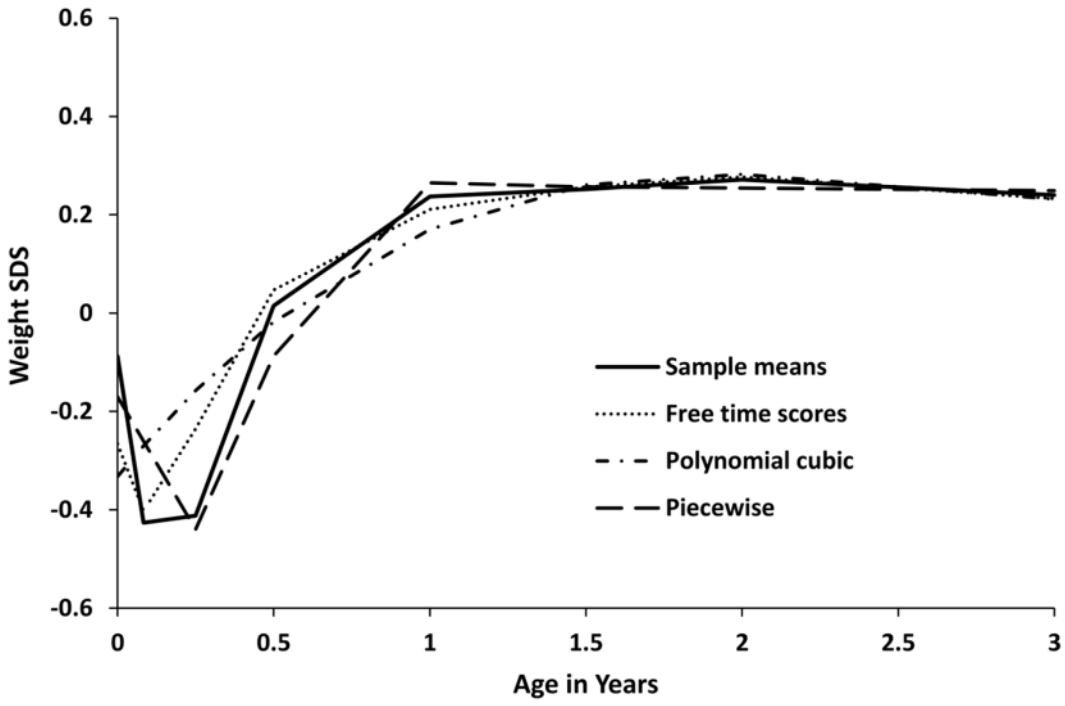
Comparison of three modelling techniques showed that a piecewise model with two joints (i.e. at 3 months and 12 months) performed better than polynomials and free time score functions (Table 4.15 and Figure 4.7). The Log-likelihood, Akaike information criterion (AIC) and Bayesian information criterion (BIC) were all optimal, and residuals were relatively smaller, albeit inconsistent, when compared to the polynomials and free time score functions (Table 4.15 and Figure 4.7). However, although some improvements were seen when compared to the linear and two nonlinear models, the RMSEA, CFI, TLI and SRMR values remained suboptimal (Table 4.15).

Table 4.15 Model fit statistics for non-linear latent growth curve models

	Free-time scores model	Piecewise model	Polynomial model
Fit index			
AIC	11,675	11,409	11,753
BIC	11,767	11,517	11,861
RMSEA	0.131	0.106	0.147
CFI	0.86	0.919	0.845
TLI	0.901	0.901	0.811
SRMR	0.084	0.040	0.074
-2LL	11,639	11,367	11,711
Residuals (observed –expected means)			
$WSDS_0$	0.145	0.027	0.190
$WSDS_1$	-0.039	-0.231	-0.201
$WSDS_3$	-0.207	-0.056	-0.282
$WSDS_6$	-0.056	0.162	0.024
$WSDS_{12}$	0.022	-0.034	0.073
$WSDS_{18}$	-0.004	-0.011	-0.005
$WSDS_{24}$	-0.002	0.016	-0.018
$WSDS_{36}$	0.005	-0.002	0.008

WSDS= standardised weight scores; subscripts are age in months.

Figure 4.7 Sample and estimated mean curves of non-linear latent growth curve models



Note: x and y axes denote age of children when measurtement was recorded and the standardised weight scores, respectively.

4.5.1.3 Piecewise latent growth curve model

According to an overall (one group) growth model, the white British and Pakistani children had an average birthweight of 3.3kgs (-0.116 SDS) with a significant downward slope or decrease in velocity (-0.707 SDS) between birth and 3 months, a significant upward slope or velocity (0.665 SDS) between 3 and 12 months, and a non significant downward slope (-0.012 SDS) between 12 and 36 months of age (Table 4.16 and Figure 4.7). The children had significant variations interms of birthweight and velocities between birth and age of one year (Table 4.17). On average, low birthweight children had higher growth velocities than the high birthweight children and vice versa. Likewise, children that had a higher growth veleocity during the first three months showed a lower growth velocity between the ages of one and three years than those who had lower velecity during the same period (Table 4.17).

Table 4.16 Estimated means of latent growth factors for the overall and multi-group piecewise latent growth curve models

Model		Estimate		P-value
		value	95% CI	
Means				
Overall (one group) model				
	Intercept	-0.116	-0.174 to -0.058	<0.01
	$Slope_{0-3}$	-0.707	-0.970 to -0.443	<0.01
	$Slope_{3-12}$	0.665	0.583 to 0.748	<0.01
	$Slope_{12-36}$	-0.012	-0.035 to 0.011	0.32
Multi-group model				
White British				
	Intercept	0.119	0.030 to 0.209	<0.01
	$Slope_{0-3}$	-0.881	-1.289 to -0.473	<0.01
	$Slope_{3-12}$	0.578	0.451 to 0.705)	<0.01
	$Slope_{12-36}$	-0.057	-0.086 to -0.028	<0.01
Pakistani				
	Intercept	-0.299	-0.377 to -0.221	<0.01
	$Slope_{0-3}$	-0.709	-1.006 to -0.412	<0.01
	$Slope_{3-12}$	0.770	0.684 to 0.857	<0.01
	$Slope_{12-36}$	0.031	-0.003 to 0.066	0.07

Note: subscripts are age in months; Intercept=birthweight; slope=velocity.

Results from the multi-group model showed that the Pakistani children were 191grams (i.e. 0.498 SDS) lighter than the white British children at birth. Furthermore, the white British children had statistically significant downward and upward trends between birth and 36 months ($slope_{0-3}$ =-0.881 SDS, $slope_{3-12}$=0.578 SDS, $slope_{12-36}$=-0.057 SDS), whereas, the Pakistani children had a statistically significant change of trend between birth and 12 months ($slope_{0-3}$=-0.709 SDS; $slope_{3-12}$=0.770 SDS) but a non-statistically significant change of trend between 12 and 36 months ($slope_{12-36}$=0.031 SDS), see Table 4.16.

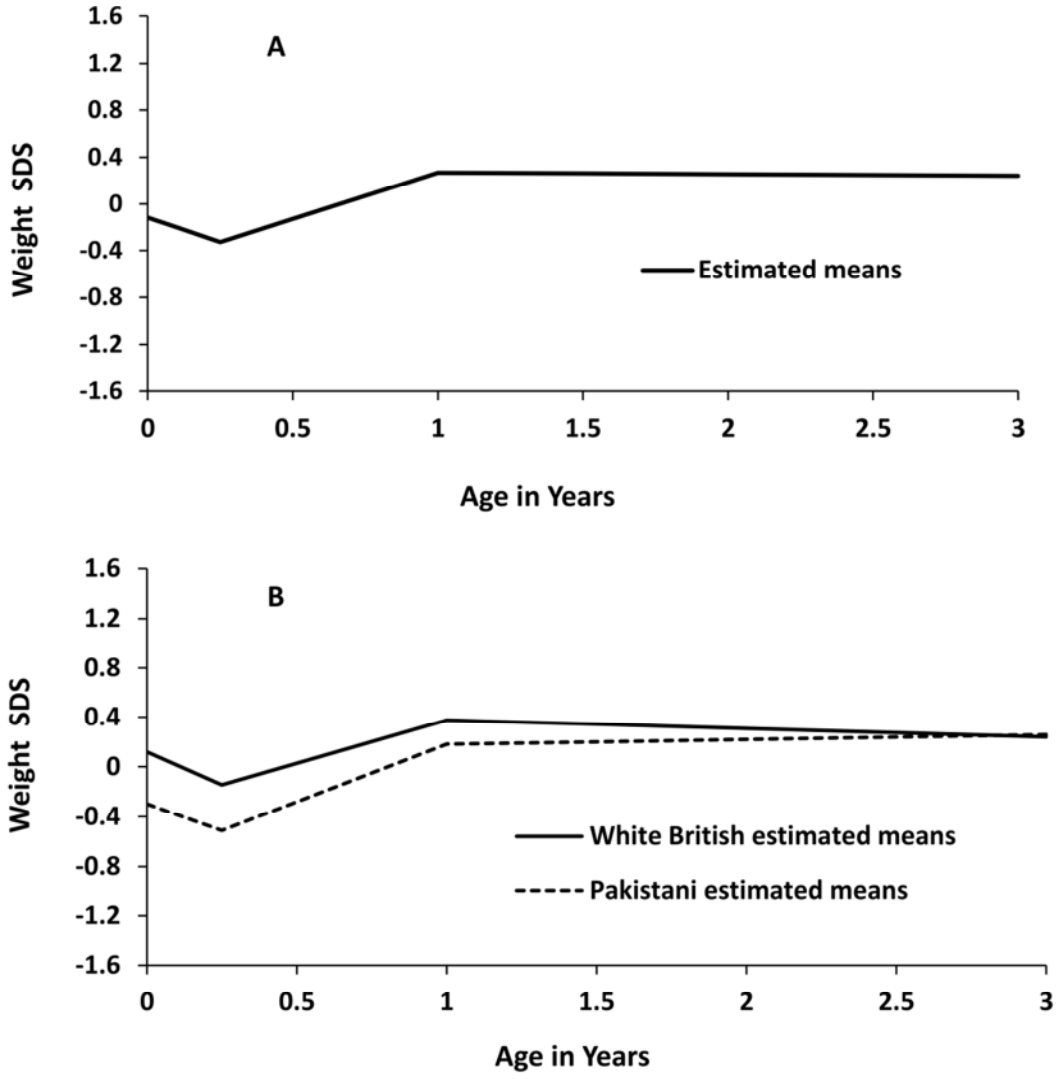
The multi-group analysis results also confirmed that the two ethnicities had distinct growth curves (Figure 4.8B). According to the estimated latent growth parameter estimates (Table 4.16), the Pakistani children had a significantly lower birthweight SDS than the white British, that is, -0.299 SDS and 0.119 SDS, respectively. There were significant variations among Pakistani and white British children in terms of birthweight and growth velocities (Table 4.17). In both ethncities, there were significant inverse relationsships between birthweight and velocities between birth and 12 months.

Table 4.17 Variance and covariance estimates for the overall and multi-group latent growth curve models

	Parameter	Estimate	Standard error	P-value
Overall model				
Variances	Intercept	1.041	0.052	<0.01
	$Slope_{0-3}$	9.232	0.805	<0.01
	$Slope_{3-12}$	0.465	0.070	<0.01
	$Slope_{12-36}$	0.065	0.006	0.30
Covariance	Intercept ↔ $slope_{0-3}$	-1.484	0.174	<0.01
	Intercept ↔ $slope_{3-12}$	-0.172	0.039	<0.01
	Intercept ↔ $slope_{12-36}$	-0.026	0.042	0.07
	$Slope_{0-3}$ ↔ $slope_{3-12}$	-0.040	0.159	0.80
	$Slope_{0-3}$ ↔ $slope_{12-36}$	-0.129	0.042	<0.01
	$Slope_{3-12}$ ↔ $slope_{12-36}$	-0.002	0.013	0.87
White British				
Variance	Intercept	0.980	0.083	<0.01
	$Slope_{0-3}$	9.553	1.172	<0.01
	$Slope_{3-12}$	0.588	0.129	<0.01
	$Slope_{12-36}$	0.050	0.006	<0.01
Covariance	Intercept ↔ $slope_{0-3}$	-1.277	0.263	<0.01
	Intercept ↔ $slope_{3-12}$	-0.168	0.054	<0.01
	Intercept ↔ $slope_{12-36}$	-0.015	0.021	0.48
	$Slope_{0-3}$ ↔$slope_{3-12}$	-0.612	0.272	0.03
	$Slope_{0-3}$ ↔$slope_{12-36}$	-0.115	0.064	0.07
	$Slope_{3-12}$ ↔ $slope_{12-36}$	-0.017	0.021	0.43
Pakistani				
Variances	Intercept	1.001	0.065	<0.01
	$Slope_{0-3}$	8.987	1.037	<0.01
	$Slope_{3-12}$	0.348	0.069	<0.01
	$Slope_{12-36}$	0.077	0.010	<0.01
Covariance	Intercept ↔ $slope_{0-3}$	-1.602	0.224	<0.01
	Intercept ↔ $slope_{3-12}$	-0.129	0.050	<0.01
	Intercept ↔$slope_{12-36}$	-0.017	0.019	0.36
	$Slope_{0-3}$ ↔ $slope_{3-12}$	0.370	0.162	0.02
	$Slope_{0-3}$ ↔ $slope_{12-36}$	-0.137	0.058	0.02
	$Slope_{3-12}$ ↔ $slope_{12-36}$	-0.001	0.016	0.93

Note: subscripts are age in months.

Figure 4.8 **Estimated mean curves of overall (A) and multi-group (B) latent growth curve models**



4.5.2 Piecewise growth mixture model

4.5.2.1 Determination of optimal class number

The goodness fit indices for the classification models did not agree (Table 4.18). However, none of the model fit indices favoured one class (i.e. equivalent to the LGCM). While the log-likelihood and AIC favoured the highest class model, the sample size adjusted BIC indicated that the three classes model was optimal. According to simulation studies by Nylund et al. (2007) and Yang (2006), BIC and sample size adjusted BIC were found to be superior to all Information Criteria indices. Of the two likelihood ratio tests (i.e. Lo-Mendell-Rubin Likelihood Ratio Test (LMR LRT) and Bootstrapped Likelihood Ratio Test (BLRT)), the BLRT was discovered to be superior (Nylund et al., 2007). In line with the recommendation of these simulation studies, both the adjusted and non adjusted LMR LRTs rejected the K and K+1 (i.e. K is class number) class models consistently. The selection of the optimal number of classes was, therefore, guided mainly by sample size adjusted BIC and BLRT values. Owing to the high computational time needed for BLRT estimation, only 2-5 class models were run and selected for comparison based on ABIC and classification quality (entropy) values of the classes.

Table 4.18 Model fit results for selection of optimal number of classes of growth mixture model

	Model fit Criterion				**Classification quality**	**Likelihood ratio test**
	-2LL	AIC	ABIC	df	Entropy	BLRT (-2LL diff; df diff; and P-values)
1 class	11,886.2	11,928.2	11,971.1	21	N/A	N/A
2 classes	11,839.8	11,891.9	11,945.0	26	0.92	46.34; 5; <0.001
3 classes	11,805.4	11,867.5	11,930.7	31	0.91	34.42; 5; 0.002
4 classes	11,786.6	11,858.6	11,932.1	36	0.89	18.80; 5; 0.070
5 classes	11,766.6	11,848.7	11,932.4	41	0.70	19.85; 5; 0.065
6 classes	11,749.6	11,841.5	11,935.4	46	0.70	-
7 classes	11,731.8	11,833.6	11,937.7	51	0.69	-
8 classes	11,716.6	11,828.6	11,942.9	56	0.64	-
9 classes	11,702.6	11,824.6	11,949.2	61	0.66	-

LL= Log-likelihood; AIC=Akaike Information Criterion; ABIC= sample size adjusted Bayesian Information Criterion; BLRT= bootstrapped likelihood ratio test; -2LL diff=2 times the Log-likelihood difference, df=degrees of freedom (number of free parameters); df diff= difference in the degree of freedom.

Based on class numeration results (Table 4.18), the BiB1000 children had three optimal classes (Figure 4.9). In Table 4.19, the average latent class assignment probabilities of individuals in each of the three classes are outputted. These figures are indicators of classification quality of the model's class assignment. A classification model with main diagonal matrix values closer to 1 is considered to be more reliable; the recommended cut-off point is 0.70 (Nagin, 2005). For example, the average probability of individuals in latent class 1 to be correctly assigned to their most likely class (i.e. class 1) was 0.975 (97.5%). In other words, on average, there was a 2.5% of class assignment error in class 1. Likewise, the average probabilities of individuals in latent class 2 and class 3 to be correctly assigned to their respective latent classes were 81.3% and 86.3%, that is, the class assignment errors in class 2 and 3 were 19% and 16%, respectively. The classification quality (entropy) for the three classes model was 0.91 (Table 4.18) which is above the recommended adequate cut-off point of 0.80 (Clark, 2010).

Table 4.19 Average latent class probabilities for most likely latent class membership (row) by latent class (column)

	Class 1	**Class 2**	**Class 3**
Class 1	**0.975**	0.015	0.010
Class 2	0.171	**0.813**	0.016
Class 3	0.137	0.000	**0.863**

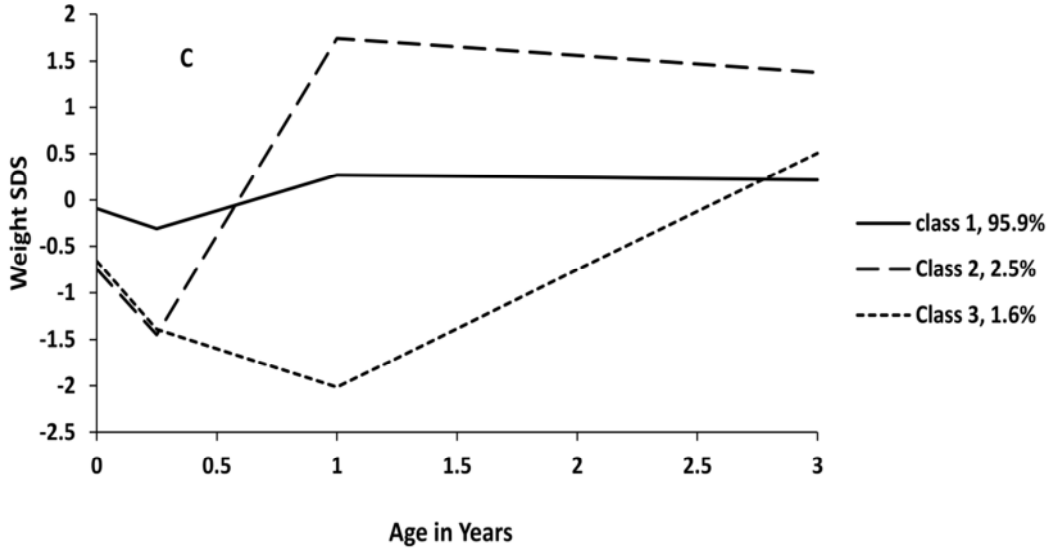
Class 1, which comprised 95.9% of the sample population, were characterised by consistent growth from birth until the age of 3 (Figure 4.9). This group of children had a birthweight SDS (intercept) of -0.095; and, statistically significant downward $slope_{0\text{-}3}$ (-0.726 SDS, 95% CI: -0.977 to -0.474), upward $slope_{3\text{-}12}$ (0.646 SDS, 95% CI: 0.571, 0.721), and an insignificant downward $slope_{12\text{-}36}$ (-0.022 SDS, 95% CI: -0.047 to 0.002), see Table 4.20. Based on the growth patterns, the group can be classified as 'normal growth' group. Generally speaking, the means weight SDS (i.e., from birth to 36 months) of this group of children were within the 38^{th} and 61^{st} centile range when compared to the WHO growth standards (WHO, 2006).

Latent class 2, which comprised 2.5% of the population, had the lowest mean birthweight SDS and showed the fastest growth from three months until 12 months when compared to the other two classes (Figure 4.9). The group had an estimated

mean birthweight SDS (intercept) of -0.746. Between birth and 3 months, the group showed a non-statistically significant drop ($slope_{0-3}$= -2.344 SDS, 95% CI: -7.975 to 3.287), then significant change to an upward trend ($Slope_{3-12}$=3.547 SDS, 95% CI: 2.438 to 4.655) between 3 and 12 months, and then a non-significant downward trend ($slope_{12-36}$ SDS=-0.151, 95% CI: -0.504 to 0.202) until the age of 3 years (Table 4.20). When compared to the WHO growth charts, the group's estimated mean standardised weight at birth was 22nd centile. Then by the age of three months, the estimated mean was at the 7th centile, and by the age of one year, it was at the 96th centile (WHO, 2006). This group can be categorised as 'fast growth' group that were observed to be overweight from 1 to 3 years of age.

The children in class 3, comprising 1.6% of the population, are those who showed a consistent downward trend from birth until 12 months, that is, $slope_{0-3}$=-2.434 SDS (95% CI: -5.496 to 0.628) and $slope_{3-12}$=-0.692 SDS (95% CI: -1.790 to 0.406). Between 12 and 36 months, they showed a significant upward trend ($slope_{12-36}$=1.050 SDS, 95% CI: 0.534 to 1.565). Subsequently, they consistently gained weight until 3 years. When compared with the WHO growth charts, their estimated mean birthweight was just above the 25th centile. By the age of 12 months, this dropped to the 2nd centile, and then at the age of 3, their mean sharply increased to the 69th centile (WHO, 2006). Generally speaking, the group can be categorised as 'slow growth' group.

Figure 4.9 Estimated mean curves of three classes GMM for 1,364 Pakistani and white British children



Collectively, children in all classes had significant variations in terms of birthweight (i.e, intercept) and growth velocities (i.e. $slope_{0-3,}$ $slope_{3-12}$ and $slope_{12-36}$). Furthermore, birthweight had a statistically significant inverse relationship with the velocity of growth between birth and three months (i.e. $slope_{0-3}$), and between three and twelve months (i.e. $slope_{3-12}$), see Table 4.20. Note that the variance and covariance were held equal across the three classes.

When the probabilities of the two ethnicities were compared respective to the three classes (reference=class 1), the Pakistani children had a higher probability of being in the 'faster growth' and 'slow growth' groups than the white British, that is, ORs of 2.90 (95% CI: 0.91 to 9.25) and 15.63 (95% CI: 1.06 to 230) for the 'fast growth' and 'slow growth' respectively (Table 4.21).

Table 4.20 Parameter estimates of latent growth factors of growth mixture model of Pakistani and white British children

		Estimate		P-value
		value	95% CI	
Means				
Class 1	Intercept	-0.095	-0.164 to -0.025	<0.01
	$Slope_{0-3}$	-0.726	-0.977 to -0.474	<0.01
	$Slope_{3-12}$	0.646	0.571 to 0.721	<0.01
	$Slope_{12-36}$	-0.022	-0.047 to 0.002	0.07
Class 2	Intercept	-0.746	-1.641 to 0.149	0.10
	$Slope_{0-3}$	-2.344	-7.975 to 3.287	0.42
	$Slope_{3-12}$	3.547	2.438 to 4.655	<0.01
	$Slope_{12-36}$	-0.151	-0.504 to 0.202	0.40
Class 3	Intercept	-0.660	-1.551 to 0.230	0.15
	$Slope_{0-3}$	-2.434	-5.496 to 0.628	0.12
	$Slope_{3-12}$	-0.692	-1.790 to 0.406	0.22
	$Slope_{12-36}$	1.050	0.534 to 1.565	<0.01
Variances*				
	Intercept	1.014	0.903 to 1.126	<0.01
	$Slope_{0-3}$	8.938	7.394 to 10.482	<0.01
	$Slope_{3-12}$	0.297	0.206 to 0.387	<0.01
	$Slope_{12-36}$	0.057	0.047 to 0.067	<0.01
Covariance*				
	Intercept ↔ $slope_{0-3}$	-1.448	-1.789 to -1.108	<0.01
	Intercept ↔ $slope_{3-12}$	-0.157	-0.258 to -0.057	<0.01
	Intercept ↔ $slope_{12-36}$	-0.015	-0.043 to 0.012	0.27
	$Slope_{0-3}$ ↔ $slope_{3-12}$	0.018	-0.319 to 0.356	0.92
	$Slope_{0-3}$ ↔ $slope_{12-36}$	-0.128	-0.210 to -0.046	<0.01
	$Slope_{3-12}$ ↔ $slope_{12-36}$	0.012	-0.009 to 0.034	0.27

*= Parameter estimates were held equal across classes; subscripts are age in months.

Table 4.21 Results of categorical latent variable multinomial logistic regressions using 3-step procedure for the three classes of Pakistan and white British children

		Odds ratio (95% CI)	P-value
Class 2 (fast growth**)***			
	Ethnicity (ref=white British)	2.90 (0.91 to 9.25)	0.072
	Smoking (ref=yes)	0.23 (0.04 to 1.29)	0.095
	Mother's education(ref=5 GSCEs)	1.87 (0.87 to 4.01)	0.111
	Parity(ref=primiparous)	0.30 (0.08 to 1.21)	0.092
Class 3 (slow growth**)***			
	Ethnicity (ref=white British)	15.63 (1.06 to 230)	0.045
	Smoking (ref=yes)	0.15 (0.02 to 1.01)	0.051
	Mother's education(ref=5 GSCEs)	1.03 (0.53 to 2.01)	0.934
	Parity(ref=primiparous)	1.42 (0.17 to 11.88)	0.747

*= reference is class 1 (the normal growth)

4.6 Effects of childhood growth patterns on wheezing disorders

The BiB1000 follow-up cohort consisted of 1,598 children that contributed a total of 8,683 person years of follow-up. The overall observed means at birth, 1 month and 3 months were below the 50th centile whereas from 6 months onwards above the 50th centile (Table 4.22), according the WHO growth standards (WHO, 2006). The correlation among the repeated weight measurements was between 0.342 and 0.936 (Table 4.22). The covariance coverage (the proportion of data present in variable x given variable y) was between 0.085 (~9%) and 1.00 (100%). There was a substantial amount of missing weight data for the ages 3 , 24 and 36 months (Table 4.23).

Table 4.22 Means, correlations and covariance coverage of repeated weight SDS measurements of 1,598 BiB1000 children

Months	Mean*	Correlation Matrix							
		0	1	3	6	12	18	24	36
0	-0.103	1.00							
1	-0.428	0.851	1.00						
3	-0.404	0.588	0.770	1.00					
6	0.006	0.472	0.613	0.861	1.00				
12	0.220	0.411	0.506	0.735	0.879	1.00			
18	0.238	0.399	0.480	0.694	0.814	0.913	1.00		
24	0.253	0.374	0.443	0.654	0.762	0.850	0.936	1.00	
36	0.209	0.342	0.428	0.613	0.710	0.788	0.863	0.911	1.00
		Covariance coverage							
0		1.00							
1		0.681	0.681						
3		0.250	0.220	0.250					
6		0.569	0.409	0.152	0.569				
12		0.452	0.317	0.118	0.298	0.452			
18		0.477	0.343	0.130	0.303	0.280	0.477		
24		0.315	0.223	0.085	0.205	0.174	0.206	0.315	
36		0.391	0.282	0.105	0.270	0.220	0.245	0.174	0.391

* = observed values

Table 4.23 Complete weight measurements for 1,598 BiB1000 children

Measurement period	Complete data (%)
Birth	1,598 (100)
1 month	1,092 (68.3)
3 months	399 (25)
6 months	910 (56.9)
12 months	722 (45.2)
18 months	763 (47.7)
24 months	504 (31.5)
36 months	625 (39.1)

After estimating missing growth data using FIML, 1.6% of the BiB1000 children had missing information on at least one covariate. Fewer than 2% and 10% of the children were diagnosed with or treated for wheezing disorders during the first three months and the first six months, respectively (Table 4.24). The total number of children who had 'asthma' diagnosis, 'wheezing' symptoms, 'wheezing disorders' diagnosis and 'wheezing disorders' treatment were 113 (7.1%) , 252 (15.8%), 300 (18.8%) and 369 (23.1%) respectively, slightly higher than the whole BiB cohort (Table 4.10).

Table 4.24 Period of diagnosis or treatment initiation of 1,598 BiB1000 children

	Period in months			
	First 3 months	First 6 months	First 9 months	First 12 months
Wheezing disorders diagnosis	1.3%	8.3%	17.0%	27.7%
Wheezing disorders treatment	2.1%	16.8%	33.1%	46.1%
Asthma diagnosis	0%	1.8%	2.7%	4.4%
Wheezing symptoms	1.59	7.9%	19.8%	31.8%

4.6.1 Piecewise LGCM for 1,598 BiB1000 children

4.6.1.1 Describing growth velocities

Based on the LGCM, the BiB1000 children had an average birthweight of 45th centile (SDS =-0.128) based on the WHO growth charts (WHO, 2006). Overall, the children showed a significant downward trend during the first 3 months ($velocity_{0-3}$=-0.761) and upward trend between 3 and 12 months ($velocity_{3-12}$=0.684).

There was a significant variation among 1,598 children on birthweight and velocities between birth and 36 months. Birthweight was inversely related with the velocity of growth, that is, children who had a higher birthweight were seen to have lower velocity of growth and vice versa (Table 4.25 and Figure 4.10). While there was a significant inverse relationship between velocities of the first 3 months and after 1 year, no significant association was observed between any of the velocities (Table 4.25).

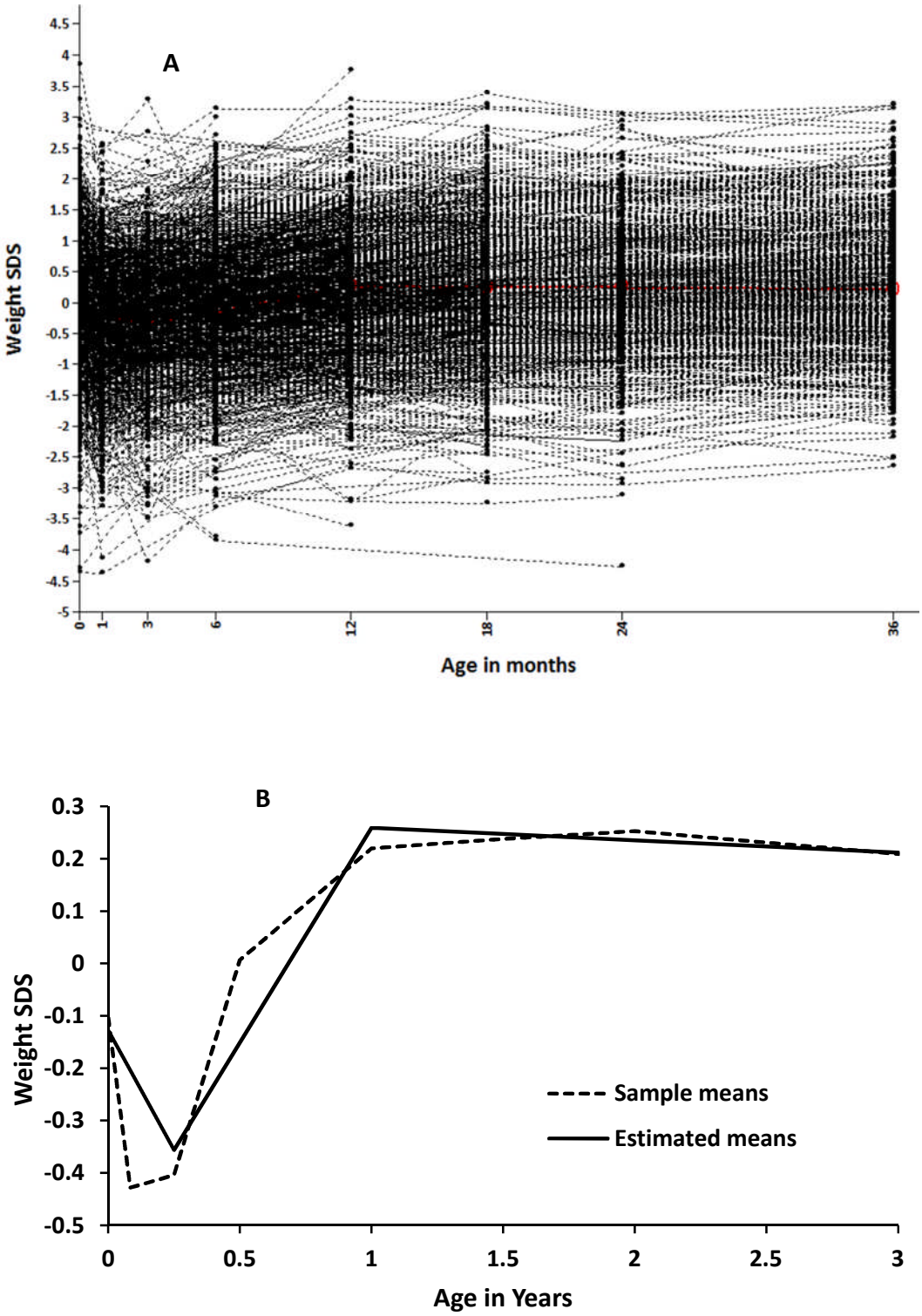
Although RMSEA and SRMR values are just at the border, CFI and TLI values are slightly lower than the recommended (Table 4.25). This may indicate that LGCM is not the optimal model for the BiB1000 children's growth data.

Table 4.25 Parameter estimates and model fit statistics of a piecewise latent growth curve model of 1,598 BiB1000 children

	Parameter	Estimate	Standard error	P-value
Means				
	Birthweight	-0.128	0.027	<0.01
	$Velocity_{0-3}$	-0.762	0.112	<0.01
	$Velocity_{3-12}$	0.684	0.035	<0.01
	$Velocity_{12-36}$	-0.020	0.011	0.08
Variances				
	Birthweight	1.009	0.047	<0.01
	$Velocity_{0-3}$	9.180	0.726	<0.01
	$Velocity_{3-12}$	0.503	0.065	<0.01
	$Velocity_{12-36}$	0.066	0.006	<0.01
Covariance				
	Birthweight ↔ $velocity_{0-3}$	-1.401	0.157	<0.01
	Birthweight ↔ $velocity_{3-12}$	-0.176	0.035	<0.01
	Birthweight ↔ $velocity_{12-36}$	-0.027	0.013	0.03
	$Velocity_{0-3}$ ↔ $velocity_{3-12}$	-0.081	0.147	0.58
	$Velocity_{0-3}$ ↔ $velocity_{12-36}$	-0.120	0.040	<0.01
	$Velocity_{3-12}$ ↔ $velocity_{12-36}$	-0.009	0.013	0.47
Residual Variances (errors)				
	E_1 (birth)	0.059	0.029	0.04
	E_2 (1 month)	0.229	0.022	<0.01
	E_3 (3 months)	0.073	0.033	0.03
	E_4 (6 months)	0.165	0.017	<0.01
	E_5 (12 months)	0.079	0.016	<0.01
	E_6 (18 months)	0.070	0.012	<0.01
	E_7 (24 months)	0.106	0.011	<0.01
	E_8 (36 months)	0.013	0.027	0.62
Fit indices				
	AIC	13,836		
	BIC	13,949		
	-2LL	13,794		
	RMSEA	0.05		
	SRMR	0.05		
	CFI	0.89		
	TLI	0.86		

Note: subscripts are age in months; birthweight=Intercept; Velocity=Slope; AIC=Akaike Information Criterion; BIC= Bayesian Information Criterion; RMSEA= Root Mean Square Error Approximation; CFI=Comparative Fit Index; TLI=Tucker-Lewis Index; SRMR=Standardised Root Mean Square Residuals and -2 LL= -2 X Log-likelihood; df=degrees of freedom. Cut-off values: CFI>0.95; TLI>0.95; RMSEA≤0.05; SRMR<0.05.

Figure 4.10 Individual observed growth curves and estimated mean (A) and sample and estimated mean curves (B) of 1,598 BiB1000 children

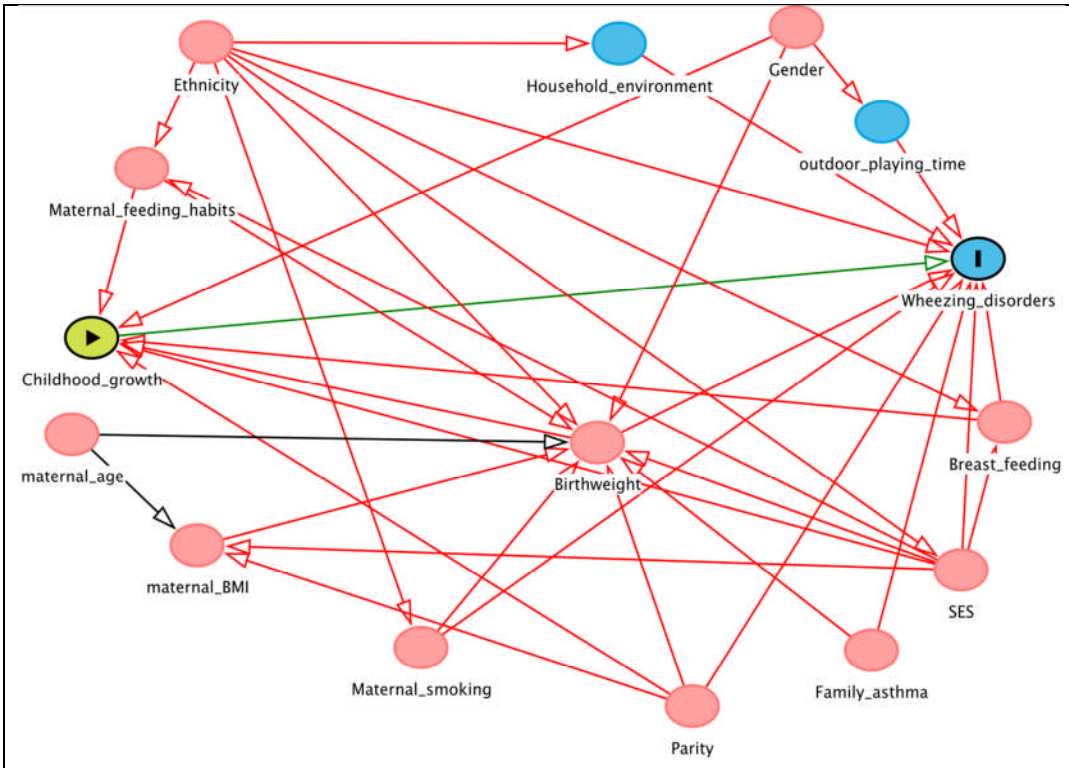


Note: Red line in A is the estimated mean curve and the black lines are individual oberved curves. The solid and dashed lines in B are estimated and sample (observed) mean curves. Both in A and B, the X and Y axes denote age of children when measurtement was recorded and the standardised weight scores, respectively.

Selection of minimally sufficient sets of confounding variables

Based on the DAG model output, three sets of minimally sufficient sets of confounding variables were eligible for selection. The set that contained birthweight, ethnicity, family asthma, breast feeding, gender, maternal smoking, parity, and SES was selected as 'minimally sufficient' set of confounding variables (Figure 4.11). However, information on family asthma and breast feeding was not available so childhood growth and wheezing disorder models were not adjusted for these variables.

Figure 4.11 DAG model output of confounding adjustment for models that investigated the effects of childhood growth on wheezing disorders



Key:

=Exposure variable; =Outcome variable; =Ancestor of exposure and outcome; = Ancestor of outcome; =Causal path; =Biasing path

Minimal sufficient adjustment sets for estimating the total effect of childhood growth on wheezing disorders:

- Birthweight, breast feeding, ethnicity, family asthma, gender, maternal smoking, parity and SES.
- Birthweight, breast feeding, ethnicity, family asthma, maternal smoking, parity, SES and outdoor playing time.
- Birthweight, breast feeding, gender, maternal feeding habits, parity and SES.

4.6.1.2 Effects of growth velocities on risk of wheezing disorders

When the velocities of growth were assessed for wheezing disorders risk, a slow growth during the first 3 months and fast growth between three and 12 months were associated with significant increased risk of all four wheezing disorders irrespective of adjusting for confounding variables (Table 4.26). For example, for every 1SDS decrease between birth and three months had an associated 5% risk of wheezing disorder diagnosis (adjusted RR= 1.05, 95% CI: 1.04 to 1.07). The respective increased risk for the upward velocity between 3 and 12 months was 26% (RR= 1.26, 95% CI: 1.18 to 1.35).

Table 4.26 Adjusted and unadjusted relative risks and 95% confidence intervals of velocities between birth and 36 months from 10 imputed datasets of 1,598 BiB1000 children

	Unadjusted RR (95% CI; p-value)	Adjusted RR (95% CI; p-value)
Velocity (-1SDS) between birth and 3 months		
Asthma diagnosis	1.09 (1.07 to 1.12; 0.04)	1.08 (1.05 to 1.11; <0.01)
Wheezing symptom	1.06 (1.06 to 1.08; <0.01)	1.07 (1.05 to 1.08; <0.01)
Wheezing disorder diagnosis	1.06 (1.05 to 1.07; 0.01)	1.05 (1.04 to 1.07; <0.01)
Wheezing disorder treatment	1.05 (1.04 to 1.07; 0.01)	1.05 (1.04 to 1.07; <0.01)
Velocity(+1SDS) between 3 and 12 months		
Asthma diagnosis	1.49 (1.34 to 1.66; <0.01)	1.39 (1.24 to 1.56; <0.01)
Wheezing symptom	1.29 (1.20 to 1.39; <0.01)	1.29 (1.19 to 1.39; <0.01)
Wheezing disorder diagnosis	1.28 (1.20 to 1.36; <0.01)	1.26 (1.18 to 1.35; <0.01)
Wheezing disorder treatment	1.19 (1.13 to 1.26; <0.01)	1.20 (1.13 to 1.27; <0.01)
Velocity(-1SDS) between 12 and 36 months		
Asthma diagnosis	1.35 (0.96 to 1.88; 0.08)	0.91 (0.64 to 1.29; 0.60)
Wheezing symptom	0.68 (0.54 to 0.84; <0.01)	0.54 (0.43 to 0.68; <0.01)
Wheezing disorder diagnosis	0.79 (0.65 to 0.97; 0.02)	0.64 (0.52 to 0.78; <0.01)
Wheezing disorder treatment	0.83 (0.70 to 0.99; 0.04)	0.78 (0.65 to 0.94; <0.01)

All models were adjusted for birthweight, ethnicity, gender, maternal smoking, parity and maternal SES;

4.6.2 Piecewise GMM for BiB1000 children

4.6.2.1 Growth patterns analysis

According to the optimal number of class determination results, a four class model was best (Table 4.27). However, a three class model was preferred on an interpretability basis.

Table 4.27 Model fit results for selection of optimal number of classes of 1,598 BiB1000 children

	Model fit Criterion				Classification quality	Likelihood ratio test
	-2LL	AIC	ABIC	df	Entropy	BLRT (-2LL diff; df diff; and P-values)
1 class	13,794	13,836	13,883	21	N/A	N/A
2 classes	13,752	13,805	13,862	26	0.94	42; 5; <0.01
3 classes	13,724	13,785	13,853	31	0.90	29; 5; <0.01
4 classes	13,698	13,770	13,849	36	0.88	24; 5; 0.02
5 classes	13,680	13,763	13,853	41	0.88	17; 5; 0.70

Based on class numeration results (Table 4.27), the BiB1000 children had three optimal classes (Figure 4.12). Based on the average latent class assignment result, the probability of individuals to be correctly assigned to class 1, class 2 and class 3, was 97%, 80% and 86%, respectively (Table 4.28). In other words, on average, there was a 3%, 20% and 24% of class assignment error in class 1, class 2 and class 3, respectively. However, the average latent class probability were above minimum the recommended cut-off point is 0.70 (Nagin, 2005). The classification quality (entropy) for the three classes model was 0.90 (Table 4.27) which is above the recommended adequate cut-off point of 0.80 (Clark, 2010).

Table 4.28 Average latent class probabilities for most likely latent class membership (row) by latent class (column)

	Class 1	Class 2	Class 3
Class 1	**0.974**	0.014	0.012
Class 2	0.188	**0.804**	0.007
Class 3	0.000	0.242	**0.758**

Class 1 (95.8%) was composed of children whose mean birthweight was at the 46th centile and were just over the 60th centile at the age of 1 year and stayed around 60th centile afterwards according to WHO growth standards (WHO, 2006). Class 2 (2.2%) was composed of children whose mean weight at birth was on the 28th centile then increased to the 96th centile at one year of age and persisted to be overweight until the age of three. Class 3 (2.0%) were a group of children whose mean birthweight was on the 29th centile, who subsequently showed very slow growth, their mean weight reaching the 3rd centile at the of 1 year, then 56th centile by the age of three years. Class 1, class 2 and class 3, could be characterised as 'normal', 'fast' and 'slow' growth respectively (Table 4.29 and Figure 4.12).

Table 4.29 Estimated mean and percentiles of the three class piecewise growth mixture model for 1,598 BiB1000 children

	Growth classes		
	Class 1	**Class 2**	**Class 3**
Birth	46th (-0.11 SDS)	28th (-0.59 SDS)	29th (-0.56 SDS)
1 month	43rd (-0.18 SDS)	19th (-0.89 SDS)	23rd (-0.75 SDS)
3 months	38th (-0.31 SDS)	7th (-1.48 SDS)	13th (-1.13 SDS)
6 months	45th (-0.12 SDS)	34th (-0.40 SDS)	8th (-1.39 SDS)
12 months	61st (0.27 SDS)	96th (1.75 SDS)	3rd (-1.91 SDS)
18 months	60th (0.25 SDS)	94th (1.57 SDS)	8th (-1.40 SDS)
24 months	59th (0.23 SDS)	92nd (1.39 SDS)	19th (-0.88 SDS)
36 months	58th (0.20 SDS)	85th (1.02 SDS)	56th (0.14 SDS)

Based on the latent growth factors parameter estimates, the BiB1000 children had a significant variation in terms of birthweight (i.e. intercept) and velocity of growth between birth and the age of three years (Table 4.30). The results also showed that birthweight and growth velocities between birth and age of 1 year were inversely related.

Figure 4.12 Estimated mean curves of three class pricewise growth mixture model

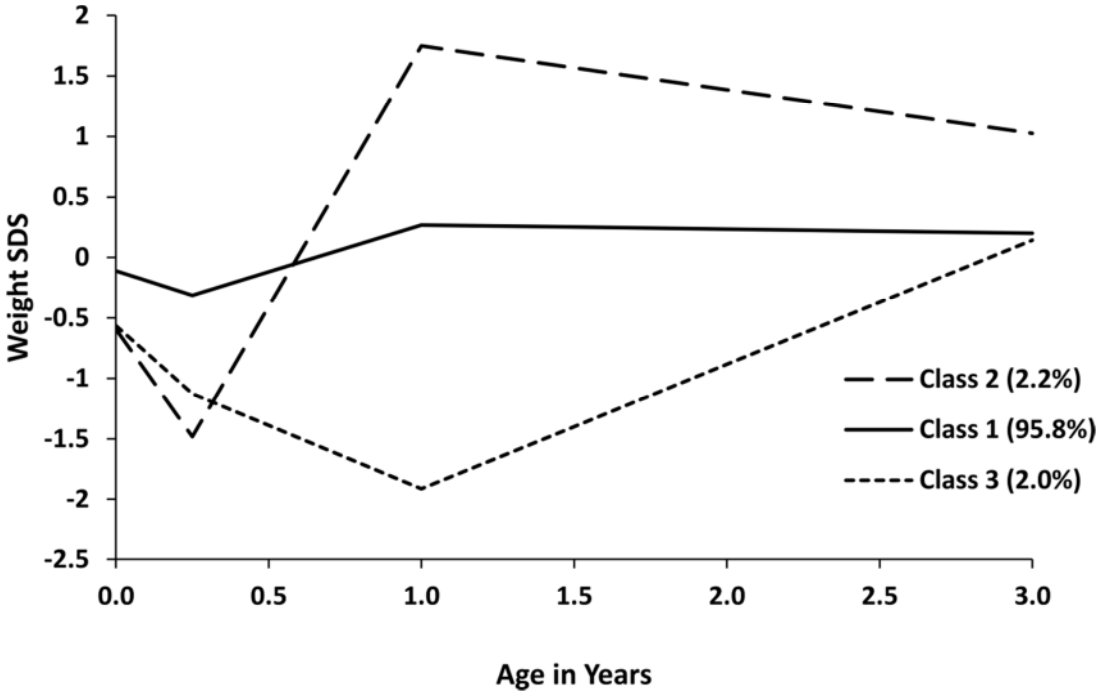


Table 4.30 Latent growth factor parameter estimates of three class piecewise growth mixture model

	Mean/variance/covariance	Estimate and 95% CI	P-value
Class 1	Birthweight	–0.111 (–0.170 to –0.053)	<0.01
	Velocity $_{0-3}$	–0.671 (–0.903 to –0.439)	<0.01
	Velocity $_{3-12}$	0.645 (0.578 to 0.712)	<0.01
	Velocity$_{12-36}$	–0.028 (–0.053, –0.003)	0.03
Class 2	Birthweight	–0.594 (–1.305 to 0.117)	0.10
	Velocity $_{0-3}$	–2.956 (–7.838 to 1.925)	0.24
	Velocity $_{3-12}$	3.588 (2.850 to 4.326)	<0.01
	velocity $_{12-36}$	–0.302 (–0.993 to 0.390)	0.39
Class 3	Birthweight	–0.564 (–1.146 to 0.018)	0.06
	Velocity $_{0-3}$	–1.878 (–3.980 to 0.225)	0.08
	Velocity $_{3-12}$	–0.871 (-1.950 to 0.208)	0.11
	Velocity $_{12-36}$	0.856 (0.266 to 1.446)	<0.01
Variances *	Birthweight	0.994 (0.897 to 1.090)	<0.01
	Velocity $_{0-3}$	8.945 (7.534 to 10.356)	<0.01
	Velocity $_{3-12}$	0.330 (0.224 to 0.437)	<0.01
	Velocity $_{12-36}$	0.057 (0.046 to 0.067)	<0.01
Covariance*	Birthweight ↔ velocity $_{0-3}$	-1.387 (-1.70 to -1.075)	<0.01
	Birthweight ↔velocity $_{3-12}$	-0.171(-0.247 to -0.94)	<0.01
	Birthweight ↔ velocity $_{3-12}$	-0.020 (-0.045 to 0.006)	0.13
	Velocity $_{0-3}$ ↔ velocity $_{3-12}$	-0.003 (-0.304 to 0.298)	0.98
	Velocity $_{0-3}$ ↔ velocity $_{12-36}$	-0.117 (-0.196 to -0.039)	<0.01
	Velocity $_{3-12}$↔ velocity $_{12-36}$	0.016 (-0.015 to 0.046)	0.31

Note: subscripts are age in months; *=parameter estimates were held equal across classes.

4.6.2.2 Effect of growth patterns on the risk of wheezing disorders

When the risk of wheezing disorders was compared among the growth classes, the slow growth group had an insignificant decreased risk when compared with the normal growth group. For example, the adjusted RRs of 'wheezing' symptoms, 'wheezing disorder' diagnosis and 'wheezing disorders' treatment were 0.72 (95% CI: 0.20 to 2.62), 0.60 (95% CI: 0.16 to 1.95) and 0.81 (95% CI: 0.29 to 2.25) respectively, see Table 4.31.

The fast growth group also showed an inconsistent risk of association for the four diseases definitions. For example, the adjusted RRs of the 'fast' compared to the 'normal' growth group for 'asthma' diagnosis, 'wheezing' symptoms, 'wheezing disorder' diagnosis and 'wheezing disorders' treatment were 0.81 (95% CI: 0.12 to 5.46), 1.59 (95% CI: 0.67 to 3.71), 1.30 (95% CI: 0.56 to 3.06) and 0.77 (95% CI: 0.20 to 2.51), respectively (Table 4.31).

Table 4.31 Adjusted and unadjusted relative risk and 95% confidence intervals using 10 imputed datasets of the BiB1000 cohort

		Unadjusted RR (95% CI; p-value)	Adjusted RR (95% CI; p-value)
Class 2	Asthma diagnosis	0.82 (0.12 to 5.56; 0.84)	0.81 (0.12 to 5.46; 0.83)
(fast growth)	Wheezing symptom	1.50 (0.62 to 3.56; 0.36)	1.59 (0.68 to 3.71; 0.29)
	Wheezing disorder diagnosis	1.25 (0.53 to 2.97; 0.61)	1.30 (0.56 to 3.06; 0.54)
	Wheezing disorder treatment	0.76 (0.27 to 2.14; 0.60)	0.77 (0.28 to 2.17; 0.63)
Class 3	Asthma diagnosis	1	1
(slow growth)	Wheezing symptom	0.80 (0.21 to 2.93;0.73)	0.72 (0.20 to 2.63; 0.29)
	Wheezing disorder diagnosis	0.67 (0.18 to 2.45; 0.54)	0.60 (0.16 to 2.18; 0.44)
	Wheezing disorder treatment	0.81 (0.29 to 2.25; 0.68)	0.81 (0.29 to 2.25; 0.69)

Note: models were adjusted for birthweight, ethnicity, gender, maternal smoking, parity and maternal SES; class 1 was a reference group.

4.7 Comparison between complete cases and imputed datasets analyses

4.7.1 Birthweight and the risk of wheezing disorders

The complete cases analysis for birthweight and wheezing disorders retained 10,623 out of the 13,734 BiB cohort children. The adjusted RRs of low and high birthweight for wheezing disorders from the complete cases analysis were similar but less efficient (wider confidence intervals), and inconsistent and less efficient respectively as compared to the imputed data results (Table 4.32). The similarity between complete cases and imputed data analyses results was expected given that all the outcome variables were completely observed and the missing indicator variables for the incomplete covariates did not have strong relationship with the outcome variables.

The unadjusted RRs of wheezing disorders (in all four definitions) for low and high birthweight using the complete cases data are also exactly the same with the imputed data results as expected provided that birthweight was completely observed (Table 4.9 and Table 4.32).

4.7.2 Weight at age of 3 years and the risk of wheezing disorders

The complete case analysis was based on 1,027 (64.2%) of the total 1,598 BiB1000 children. The adjusted RRs for the underweight, overweight, obese compared to the normal weight group using the complete cases analysis are similar to but less precise than the imputed data analyses results as expected (Table 4.11 and Table 4.33). The respective unadjusted RRs are also similar to but less efficient than the imputed data analysis results (Table 4.11 and Table 4.33).

Table 4.32 Relative risks and 95% confidence intervals of birthweight using complete data of 10,623 BiB cohort children

	Asthma diagnosis	Wheezing symptoms	wheezing disorder diagnosis	wheezing disorders treatment
Unadjusted model				
Normal (2.5-4.0kg)	1	1	1	1
High (>4.0kg)	0.93 (0.73 to 1.19)	0.91 (0.78 to 1.06)	0.92 (0.80 to 1.05)	1.04 (0.93 to 1.16)
Low (<2.5kg)	1.55 (1.27 to 1.89)	1.28 (1.13 to 1.46)	1.28 (1.14 to 1.45)	1.27 (1.15 to 1.40)
Adjusted model †				
Normal (2.5-4.0kg)	1	1	1	1
High (>4.0kg)	0.88 (0.63 to 1.22)	1.01 (0.84 to 1.21)	0.97 (0.82 to 1.14)	1.03(0.90 to1.18)
Low (<2.5kg)	1.63 (1.24 to 2.14)	1.26 (1.05 to 1.51)	1.30 (1.10 to 1.53)	1.22 (1.06 to 1.41)

† =model was adjusted for ethnicity, sex, gestational age, number of live births, maternal smoking, parity, and SES

Table 4.33 Relative risks and 95% confidence intervals of weight at the age of 3 years using complete data of 1,027 BiB1000 children

	Asthma diagnosis	Wheezing symptoms	Wheezing disorder diagnosis	Wheezing disorder treatment
Unadjusted model				
Normal (≥5th and <85th centiles)	1	1	1	1
Underweight (<5th centile)	0.57 (0.24 to 1.39)	0.97 (0.59 to 1.59)	0.88 (0.56 to 1.39)	0.78 (0.51 to 1.18)
Overweight (≥85th and <95th centiles	0.78 (0.40 to 1.54)	1.28 (0.88 to 1.88)	1.19 (0.84 to 1.67)	1.09 (0.80 to 1.49)
Obese (≥95thcentile)	1.29 (0.69 to 2.41)	0.95 (0.56 to 1.60)	1.11 (0.72 to 1.69)	1.15 (0.81 to 1.64)
Adjusted model ‡				
Normal (≥5th and <85th centiles)	1	1	1	1
Underweight (<5th centile)	0.56 (0.23 to 1.36)	0.96 (0.59 to 1.59)	0.88 (0.57 to 1.38)	0.75 (0.49 to 1.15)
Overweight (≥85th and <95th centiles	0.86 (0.44 to 1.67)	1.31 (0.90 to 1.93)	1.23 (0.87 to 1.73)	1.06 (0.77 to 1.46)
Obese (≥95thcentile)	1.20 (0.62 to 2.32)	1.01 (0.59 to 1.73)	1.11 (0.72 to 1.38)	1.06 (0.73 to 1.55)

‡ = model was adjusted for birthweight, ethnicity, sex, maternal smoking, parity, and SES

4.7.3 Childhood growth patterns and the risk of wheezing disorders

4.7.3.1 Piecewise latent growth curve model

The complete case analyses retained 1,572 of the 1,598 children. The results showed that for a decrease of 1SDS between birth and 3 months there was a moderate associated risk of wheezing disorders (i.e. both adjusted and unadjusted RRs of all four disease definitions), see Table 4.34. For an increase of 1SDS between 3 and 12 months, the unadjusted and adjusted RRs for complete case analysis did not completely agree (Table 4.34). There was a significant and insignificant risk of wheezing disorders if models were unadjusted and adjusted for covariates, respectively. However, there was an insignificant reduction of wheezing disorders risk for a decrease of 1SDS between 12 and 36 months. Overall, the relative risks of wheezing disorders for velocities between birth and 12 months of the complete cases analyses were almost the same although slightly imprecise when compared with results of imputed dataset results in Table 4.26.

Table 4.34 Adjusted and unadjusted relative risks and 95% confidence intervals of velocities between birth and 36 months using complete data of 1,572 BiB1000 children

	Unadjusted RR (95% CI; p-value)	Adjusted RR (95% CI; p-value)
Velocity (-1SDS) between birth and 3 months		
Asthma diagnosis	1.09 (1.02 to 1.17; 0.01)	1.08 (1.00 to 1.16; 0.04)
Wheezing symptom	1.07 (1.02 to 1.12; <0.01)	1.07 (1.02 to 1.12; <0.01)
Wheezing disorder diagnosis	1.06 (1.02 to 1.10; 0.01)	1.05 (1.01 to 1.10; 0.01)
Wheezing disorder treatment	1.05 (1.02 to 1.09; <0.01)	1.05 (1.01 to 1.10; 0.01)
Velocity (1SDS) between 3 and 12 months		
Asthma diagnosis	1.49 (1.03 to 2.15; 0.03)	1.36 (0.92 to 2.01; 0.13)
Wheezing symptom	1.29 (1.01 to 1.64; 0.04)	1.23 (0.95 to 1.59;0.10)
Wheezing disorder diagnosis	1.28 (1.03 to 1.59; 0.03)	1.20 (0.95 to 1.51; 0.11)
Wheezing disorder treatment	1.19 (0.98 to 1.44; 0.07)	1.15 (0.94 to 1.41; 0.16)
Velocity (-1SDS) between 12 and 36 months		
Asthma diagnosis	1.35 (0.44 to 4.08; 0.60)	0.83 (0.26 to 2.64; 0.75)
Wheezing symptom	0.68 (0.32 to 1.41; 0.29)	0.55 (0.26 to 1.18; 0.12)
Wheezing disorder diagnosis	0.79 (0.41 to 1.54; 0.49)	0.62 (0.32 to 1.23; 0.17)
Wheezing disorder treatment	0.83 (0.47 to 1.49; 0.54)	0.78 (0.43 to 1.43; 0.42)

All models were adjusted for birthweight, ethnicity, gender, maternal smoking, parity and maternal SES

4.7.3.2 Piecewise growth mixture model

Similar to the piecewise LGCM, 1,572 children contributed to the complete cases analyses. The adjusted RRs of the 'fast' growth group compared to the 'normal' growth group for 'asthma' diagnosis, 'wheezing' symptom, 'wheezing disorder' diagnosis and 'wheezing disorders' treatment were 0.84 (0.13 to 5.60), 1.26 (0.46 to 3.48), 1.03 (0.38 to 2.85) and 0.55 (0.15 to 2.02) respectively. And, the adjusted RRs of the 'slow' growth group compared to the 'normal' growth group for 'asthma' symptoms, 'wheezing disorder' diagnosis and 'wheezing disorder' treatment were 0.72 (95% CI: 0.20 to 2.60), 0.59 (95% CI: 0.16 to 2.15) and 0.81 (95% CI: 0.29 to 2.25), respectively (Table 4.35).

These results of complete cases analyses were very close to the imputed data analyses as expected given that all the outcome variables were completely observed and the missing indicator variables for the incomplete covariates did not have strong relationship with the outcome variables.

Table 4.35 Adjusted and unadjusted relative risks and 95% confidence intervals of growth patterns using complete data of 1,572 BiB1000 children

		Unadjusted RR (95% CI; p-value)	Adjusted RR (95% CI; p-value)
Age based weight SDS			
Class 2	Asthma diagnosis	0.82 (0.12 to 5.55; 0.84)	0.84 (0.13 to 5.60; 0.86)
(fast growth)	Wheezing symptom	1.50 (0.63 to 3.55; 0.36)	1.26 (0.46 to 3.48; 0.65)
	Wheezing disorder diagnosis	1.25 (0.53 to 2.97; 0.61)	1.03 (0.38 to 2.85; 0.95)
	Wheezing disorder treatment	0.76 (0.27 to 2.13; 0.60)	0.55 (0.15 to 2.02; 0.37)
Class 3	Asthma diagnosis	1	1
(slow growth)	Wheezing symptom	0.80 (0.22 to 2.92;0.73)	0.72 (0.20 to 2.60; 0.61)
	Wheezing disorder diagnosis	0.67 (0.18 to 2.44; 0.54)	0.59 (0.16 to 2.15; 0.42)
	Wheezing disorder treatment	0.81 (0.29 to 2.25; 0.68)	0.81 (0.29 to 2.25; 0.68)

All models were adjusted for birthweight, ethnicity, gender, maternal smoking, parity and maternal SES; class 1 was a reference group in both models.

CHAPTER 5

DISCUSSION

5.1 Chapter overview

This chapter presents a critical discussion of results, description of limitation and strengths of results, and conclusions drawn from the series of analyses using the BiB cohort data, namely: the incidence and burden of childhood allergic conditions and the effects of birthweight, weight at the age of 3 years and childhood growth patterns on wheezing disorders.

5.2 Incidence and burden of childhood allergic conditions in the BiB cohort

In this prospective cohort study, the results indicate that 1 in 2 children have suffered from eczema, and 1 in 5 children have suffered from wheezing disorders and rhinitis sometime between 0 and 7 years of age. Almost 2 in 5, 1 in 5 and 1 in16 children have suffered from one, two and three allergic conditions, respectively. While there was no significant difference for eczema by gender, boys were more likely to suffer from wheezing disorders and rhinitis than girls. Furthermore, while no difference was observed for wheezing disorders, Pakistani children were more likely to suffer from eczema and rhinitis than white British children. Boys and Pakistani children were more likely to suffer from multiple allergic conditions than girls and white British children, respectively.

The five-year prevalence estimates suggest that 1 in 5 children will have been diagnosed with a wheezing disorder and rhinitis, when they reach the age of 5 years; and, 1 in 2 of the cohort have had eczema during the same period. Eczema and rhinitis were more prevalent in Pakistani than white British children, whilst all three allergic conditions were more prevalent in boys than girls.

In a meta-analysis and systematic review of studies conducted in the UK by Netuveli et al. (2005), it was reported that 12-months period prevalence of asthma was lower in south Asian children (prevalence: 7.6%; 95% CI: 3.7 to 11.4%) as compared with black (prevalence: 15%; 95% CI: 3.5 to 26.5%) and white

(prevalence: 10.6; 95% 4.6 to 16.7%) children. These figures are significantly lower than any of the five-year prevalence (i.e. overall or ethnicity based figures) in this study which could be due to difference in the ethnic composition of the population, diagnosis terms used (i.e. 'asthma' versus 'wheezing disorders') and the prevalence period. It was also reported that in Pakistan, the prevalence of asthma, wheezing, eczema and rhinitis in school children of Karachi was 15.8%, 11.7%, 21.8% and 28.5% respectively (Hasnain et al., 2009). The prevalence figures for rhinitis are similar to the BiB Pakistani group results although it must be noted that the authors defined 'wheezing' and 'asthma' as separate terms and used questionnaires to confirm diagnoses of allergic conditions.

Punekar and Sheikh (2009), who used the national General Practice Research Database (GPRD) data, reported lower incidence rates and lifetime prevalence than the BiB cohort's findings. The incidence rates for eczema, asthma and rhinitis were 22.7, 13.7 and 6.1 per 1000 person years, respectively. The 18-year prevalence figures reported by the authors are also significantly lower for eczema (36.5%, 95% CI: 35.9 to 37.2%) and rhinitis (11.4%, 95% CI: 11.0 to 11.8%) while similar for wheezing disorders (22.9%, 95% 22.3 to 23.4%) when compared with the five-year prevalence of BiB cohort. These could be for two reasons. First, the authors used clinician-diagnosed allergic conditions and 'asthma' instead of 'wheezing disorders'. However, drugs can be prescribed for some period of time as a trial without any formal diagnosis (GINA, 2015). If the condition is transient, the child may not be formally diagnosed so this would underestimate the true burden of allergic diseases. Second, the GPRD data reflects the UK population and regions, but, the BiB data were composed of mainly Pakistani and white British who live in the district of Bradford. The district of Bradford has higher infant mortality (BDIMC, 2007), and air pollution has been a major concern in the community (Wright et al., 2013). Hence, the higher incidence of allergic conditions in the BiB cohort than national level could be due to either difference in ethnic composition or higher environmental pollution.

Although similar in direction, the cumulative incidence of wheezing disorder figures from the BiB cohort are moderately lower and higher than the Health Survey for

England (HSE) figures for wheezing occurrence and doctor diagnosed asthma respectively. The HSE reported 30% and 23% cumulative incidence of wheezing occurrences in boys and girls respectively; and, 17% and 12% of cumulative incidence of doctor-diagnosed asthma for boys and girls respectively, in 0-15 year old children (Boodhna and Hall, 2011). The disparities could be due to longer follow up and the use of questionnaires to confirm wheezing occurrences and doctor-diagnosed asthma in HSE's analysis report.

In a recent study that used data from the Millennium Cohort study (http://www.cls.ioe.ac.uk/page.aspx?sitesectionid=851) the lifetime prevalence at age 7 for eczema, wheeze and asthma were 42.9%, 25.8% and 15.1% respectively (Panico et al., 2014). Another recent study that used data from the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort (http://www.bristol.ac.uk/alspac/) and the Manchester Asthma and Allergy Study (MAAS) cohort (http://www.maas.org.uk/) also reported that the prevalence of eczema, wheeze and rhinitis at five years of age were 24.4%, 18.4% and 2.3%, respectively, for ALSPAC cohort; and, 32.4%, 22.5% and 28.1, respectively, for MAAS cohort (Belgrave et al., 2014). The 5-years prevalence figures from the BiB cohort were slightly higher than the prevalence reported from the Millennium Cohort; and, moderately higher than the ALSPAC and MAAS cohort results. The variations in prevalence of allergic diseases between the BIB and the other UK cohorts could be due to the difference in ethnic composition and the use of questionnaire based data.

From the cumulative incidence rates analysis of the BiB cohort based on the birth years, it can be noted that there was no significant change in the incidence rate of rhinitis during 2007-2011 (Table 4.4). However, there were substantial increases in the incidence rates of eczema during every subsequent birth year. The same pattern was also observed in the incidence rate of wheezing disorders although it plateaued between 2009 and 2011 birth years (Table 4.4). These results may indicate that either the impact of allergic conditions has increased during those birth years (Ghouri et al., 2008; Simpson et al., 2009; Turner et al., 2009) or there may have been changes in clinicians' prescribing habits.

5.3 Effects of birthweight on childhood wheezing disorders

Results from the analyses of the effects of birthweight on wheezing disorders using the BiB cohort (13,734 children) showed that, based on unadjusted and adjusted risk estimates, low birthweight was significantly associated with wheezing disorders (Table 4.9). However, the evidence about the effect of high birthweight on wheezing disorders remains inconclusive. Based on the adjusted risk estimates, there is a insignificant reduced risk of wheezing disorders (i.e. using all four disease definitions). Based on the unadjusted risk estimates, whilst there was an insignificant reduced risk of *asthma diagnosis*, *wheezing symptoms*, and *wheezing disorder diagnosis*, there was an insignificant increased risk of *wheezing disorders treatment* for being high birthweight compared to normal birthweight children (Table 4.9).

The findings for the effects of low birthweight on wheezing disorder diagnosis and treatment are in line with the findings of the meta-analysis and systematic review in chapter two that showed a 37% increase in wheezing disorders risk for low birthweight (unadjusted OR=1.37, 95% CI: 1.05 to 1.79), see Figure 2.4 . However, the finding of the effect of high birthweight on wheezing disorders is slightly different from the figure of the meta-analysis in chapter two (unadjusted OR=1.02, 95% CI: 0.99 to 1.04), see Figure 2.5. Only *wheezing disorders treatment* disease definition agreed with the meta-analysis finding (unadjusted RR=1.04; 95% CI: 0.93 to 1.16)

Although both results (i.e. the meta-analysis of previous epidemiologic studies in Chapter 2 and the BiB cohort data analysis) confirm that low birthweight is associated with wheezing disorders, the mechanism of this relationship remains unclear. However, some studies have reported that low birthweight is associated with an increased risk of childhood lower respiratory infection (Jackson et al., 2013; Lu et al., 2013). Low birthweight children can also experience more viral respiratory infection than normal birthweight children due to altered immune function (Raqib et al., 2007).

The contribution of respiratory viral infections (e.g. respiratory syncytial virus and rhinovirus) is two fold. First, in chronic infections, the virus-infected epithelium and

airway leukocytes release cytokines and mediators that increase airway inflammation (Singh et al., 2007; Sly et al., 2008). Second, recurrent infection can also damage the epithelial cells of the airways that cause constriction and stiffness of the mucosal muscles (Balfour-Lynn, 1996; Roche and Jeffery, 2002), hence, resulting in wheezing disorder symptoms.

The results indicate that low birthweight is a modifiable risk factor for childhood wheezing disorders. This implies that by implementing policy measures that enhance birthweight could result in the reduction of childhood wheezing disorders. However, some studies have reported that although South Asian children have lower birthweight than Caucasians or white European origin, they have higher adiposity at birth (Yajnik et al., 2002; Yajnik et al., 2003; West et al., 2013) and during childhood (Krishnaveni et al., 2005; Bansal et al., 2008; Whincup et al., 2010).

It was also reported that although South Asians were lighter, they had higher levels of type-2 diabetes precursors (i.e. glycated haemoglobin (HbA1c) and triglyceride concentrations) levels than white European children. This may suggest that increasing birthweight may lead to diabetes and cardiac disease in South Asian children (Whincup et al., 2010; West et al., 2013). Increasing birthweight may not have adverse effects on the white British and other ethnicities of the Bradford children; however, it may not be so for the South Asians in general and Pakistani in specific. Thus, policies that aim at enhancing birthweight in the Bradford community could be beneficial if they also incorporate measures to tackle high birthweight problems at the same time.

5.4 Weight at the age of 3 years and associated risk of wheezing disorders

Based on 1,598 BiB1000 children data, those who were underweight at the age of 3 years had a reduced risk of wheezing disorders (in all four disease definitions, and unadjusted and adjusted risk estimates) although not statistically significant (Table 4.11). The results also showed that overweight and obesity are associated with an insignificant increase in the risk of wheezing disorders (Table 4.11). The direction of these risks is in agreement with the results from a meta-analysis of past observational epidemiologic studies in chapter two; adjusted OR and 95% confidence intervals for underweight, overweight, and obese children when compared to normal weight children were 0.96 (0.75 to 1.23), 1.31 (1.20 to 1.42), 1.60 (1.42 to 1.80), respectively.

It can be noted that, unlike results from the meta-analysis in chapter two, analyses results based on 1,598 children data are all insignificant so the evidence is weak. This is likely due to a lack of statistical power of the tests to detect the risk of wheezing disorders given that the sample size was not so large for an exposure variable with four categories. In fact, post estimation power calculation (Fleiss et al., 2003) results showed that all tests that compared the unadjusted wheezing disorders risk between normal and the other weight groups (i.e. overweight, obese, and underweight) had a power less than 30%.

From the results of the meta-analyses and systematic reviews (see chapter two) and analysis of the risk of association between weight at the age of 3 years and wheezing disorders, high BMI/weight percentile is associated with wheezing disorders. However, the temporal relationship between BMI and wheezing disorders may not be apparent based on these findings. This is because, overweight or obesity, due to increased pressure on the airways, can have restrictive effect on breathing and can also cause gastro-oesophageal reflux that leads to transient asthmatic signs and symptoms (Sontag, 2000). In addition, in overweight or obese people, there are excess pro-inflammatory hormones that can trigger asthmatic symptoms (Guler et al., 2004; Castro-Rodríguez, 2007; Farah and Salome, 2012). At the same time, it is

also understood that children with respiratory symptoms can be less physically active that can lead to obesity (Lucas and Platts-Mills, 2006).

In a recent Mendelian Randomisation study, it has been reported that overweight/obesity precedes childhood wheezing disorders (Granell et al., 2014). However, the authors did not investigate the reverse direction, that is, if wheezing disorders cause overweight/obesity or not. Therefore, the possibility of reverse association can not be discredited as the current evidence stands.

5.5 Describing the growth patterns of white British and Pakistani children

Based on LGCM results, Pakistani children were lighter than the white British by 191 grams at birth. Although there was no difference in the change of weight in the first three months, Pakistani children showed faster growth than their white British counterparts between 3 and 36 months. It was reported that the prevalence of low birthweight in Pakistan was among the highest countries in the world (WHO, 2004). A study in the UK also reported that there is no significant difference of birthweight between first (i.e. mothers born abroad) and second generation (mothers born in the UK) South Asian (i.e. Bangladeshi, Indian or Pakistani) babies (Margetts et al., 2002; Harding et al., 2004). In addition, West et al. (2014) recently reported that Pakistani pregnant mothers had lower BMI and height than the white British mothers. Thus, it can be speculated that the disparity in birthweight between the two ethnic origin babies is due to diet and lifestyle of mothers.

Another recent study that used BiB1000 cohort data also reported that Pakistani children consumed more 'commercial sugar-sweetened baby meals' such as sweet drinks and chips than the white British during the age of 12-18 months (Sahota et al., 2015). The effect of sugary food and drinks is well documented suggesting that high consumption of sugar leads to obesity and high weight gain during childhood (Malik et al., 2006; Vartanian et al., 2007). Thus, it can be hypothesised that the faster growth during the age of 3-36 months seen in Pakistani children could be due to feeding habits.

The LGCM results also illustrate that Pakistani and white British children both tended to grow consistently slowly until 3 months of age when compared to the WHO growth standards (Figure 4.8A and Figure 4.8B). Although the WHO growth standards population represents all children in the world, it was made up of healthy breastfed children with no known health or environmental constraints to growth, and whose mothers were willing to follow WHO feeding recommendations, not smoking and not from low socioeconomic background (WHO, 2006). In the BiB1000 cohort population, however, no such constraint was imposed during selection of the participants. Therefore, it can be speculated that the slow growth observed in this

study's population could be due to difference in life-style and child feeding habits of mothers.

Based on GMM results, the sub population of BiB1000 children (i.e. only white British and Pakistani) had three distinct growth patterns: 'normal growth' (95.9%), 'fast growth' (2.5%) and 'slow growth' (1.6%). The Pakistani children were more likely to be in either the 'fast' or 'slow growth' group than the white British. The 'slow growth' group are similar to the growth trajectories shown by Eriksson et al. (2007) who also reported that low birthweight coupled with fast catch-up growth was associated with adulthood hypertension although the authors did not use WHO growth standards as a reference. However, a study by Rzehak et al. (2013) that used the same standardisation method as in this study had reported that children who persistently grew faster (i.e. similar to the fast growth groups in this study) as compared to those who grew normally have an increased risk of asthma by 30%.

The growth patterns results using the LGCM are similar to those reported by Fairley et al. (2013) who used multilevel spline modelling approach although the shape of the growth trajectories are not identical with this study's findings. The difference could be due to the use of different modelling parameterisations and the fact that the authors assumed MCAR during their analyses. In this thesis, however, FIML was used to estimate missing growth data.

Fairley et al. (2013) reported that the Pakistani children were lighter at birth and grew faster, during 9-24 months age, than the white British children which agrees with the findings of this study. However, the model fit statistics values in the determination of a model with optimal number of classes showed that a model with three classes was more parsimonious than a model with one class. In other words, the GMM fitted the data better than the LGCM. This indicates that LGCM and multilevel splines may not be the optimal choices for growth patterns analysis of the BiB1000 data.

The choice between multilevel spline and LGM depends mainly on the depth of information that one wants to derive from the data and the number of repeated measurement points. LGM (i.e. GMM) provides extra information (e.g., distinct

growth trajectories) about the study population and estimates the missing growth data using a FIML method. However, if the repeated measurement points are too many, parameter estimations using LGM can have more convergence problems than the multilevel spline models. In addition, the number of optimal classes and growth trajectories generated by the GMM may not agree with the initial hypothesis where a researcher may opt for the most interpretable number of classes despite the model identifying a different number of optimal classes. Therefore, there is a trade-off when choosing between the two modelling techniques.

5.6 The effects of growth patterns on childhood wheezing disorders

Based on the LGCM using growth data of 1,598 BiB1000 children, the results showed that every decrease of 1SDS during the first three months was associated with 5% increase in the risk of wheezing disorder diagnosis or treatment. 1SDS increase between 3 and 12 months was also associated with 26% increase in the risk of wheezing disorder diagnosis (Table 4.26). These results are in agreement with previous studies that reported a positive association between velocity of growth and wheezing disorders (Mamun et al., 2007; Pike et al., 2010; Zhang et al., 2010; van der Gugten et al., 2012; Anderson et al., 2013; Sonnenschein-van der Voort et al., 2014b; Magnus et al., 2015) but not so with some other studies that reported an inverse relationship (Mai et al., 2005; Sonnenschein-van der Voort et al., 2012; De Korte-De Boer et al., 2015).

Analyses of growth patterns using the 1,598 children's growth data using GMM showed inconclusive results for the group classified as 'fast' growth. While there was a non-significant increased risk of diagnosis for wheezing disorders, there was a non-significant reduced risk of receiving wheezing disorders treatment (Table 4.31). However, the results showed that the 'slow' growth group have an insignificant reduction for both wheezing disorders diagnosis and treatment when compared to the 'normal' growth group (Table 4.31). The growth patterns' associated risks are similar to Rzehak et al. (2013) who used GMM and reported hazard ratios of 1.22 (95% CI: 1.08 to 1.39) and 1.43 (95% CI: 0.90 to 2.27) for groups of children exhibiting rapid growth until 2 years and persistent rapid growth, respectively. The authors' rapid growth trajectory until 2 years group and its risk estimates are similar to the fast growth group of the BiB1000 children.

In the growth patterns and wheezing disorders analyses, on average, the children with lower birthweight SDS showed significant growth changes during the first 6 months (Figure 4.12 and Table 4.29) and were more likely to have experienced wheezing disorder conditions (Table 4.31). It can also be noted that children with the lowest birthweight SDS were more likely to be obese and to have experienced wheezing disorder conditions (Figure 4.12 and Table 4.31). Given that a very small proportion of wheezing disorders or treatment cases were identified in the first three

and six months (Table 4.24), during which changes in growth occurred, it may suggest that low birthweight coupled with rapid change in growth during the first six months is a risk factor for wheezing disorders.

5.7 Study strengths and limitations

The study has certain weaknesses and results need to be interpreted cautiously. First, in the analysis of incidence and burden of childhood allergic conditions, there was a moderate proportion of missing information on ethnicity (17.4%) which could possibly have impacted the ethnic-specific incidence rate and ratio results. The follow-up period for the cohort was also short (a maximum of 7 years) which could have impacted the incidence rate and ratio results as well as comparability with results from other cohorts that used longer follow-up periods.

Second, although the sample size for the effects of birthweight on wheezing disorders analysis was sufficiently large, study participants were those who were born at a single centre: the Bradford Royal Infirmary (BRI) maternity hospital. Births in the regional tertiary centre, home births and births in smaller hospitals outside Bradford will have been excluded. The participation in the sub-cohort (BiB1000) of growth patterns was also mainly driven by the mothers' willingness to participate and so there is likely to be further selection bias. In fact, if the socioeconomic status of the cohort population is considered, 55% of the BIB (Table 4.8) and 68% of the BiB1000 (Table 4.10) population had come from the least deprived part of the Bradford community (i.e. national IMD Quintile score=1) although it has been widely claimed that the Bradford district is among the most deprived areas of the country (BDIMC, 2007; BDIMC, 2014). Thus, BiB cohort population may not be a representative sample of the community so results must be read with a caution.

Third, the classes identified by GMM contained a small proportion of children that resulted in having less precise risk estimates. In fact, a post estimation power calculation (Fleiss et al., 2003) revealed that the tests had a very weak statistical power to detect the risk effect. For instance, from the 1,598 cohort children, 31 children were classified as 'fast' growth with 30% of whom diagnosed for wheezing disorders. In the 'normal' growth group, there were 1,531 children and 23% of them were diagnosed for wheezing disorders. The post estimation power test of a model based on these proportions was only 12.3% percent. For the model to have 80% power, the 'fast' growth group needed to be at least 414 children, that is, 'fast'

growth to ‘normal’ growth ratio of 0.27. Likewise, results for weight at the age of three years were based on a relatively small sample for a categorical exposure variable with four categories.

Fourth, missing levels of growth data at some ages and visits was substantial. Although the thesis has used FIML to address the growth data missing problem, the extent of bias was not explored using simulations. However, the extents of bias due to missing data on covariates of regression models was investigated by a series of comparisons between complete cases and multiply imputed data analyses.

Information on maternal asthma and breast feeding was also missing so the models were not adjusted for these potential confounding variables. However, the lack of adjustment may not have had a drastic effect on birthweight and BMI risk estimates as there were only slight differences between the unadjusted and adjusted summary ORs of wheezing disorders for birthweight and BMI studies (Figure 2.2, Figure 2.3, Figure 2.10, Figure 2.11, Figure 2.12 and Figure 2.13). Likewise, the direction of the risk estimates of unadjusted and adjusted models for those studies that investigated the association between childhood growth patterns and wheezing disorders was the same but slight difference in magnitude (Mamun et al., 2007; Scholtens et al., 2009; Pike et al., 2010; Zhang et al., 2010; Magnusson et al., 2012; Sonnenschein-van der Voort et al., 2012; van der Gugten et al., 2012; Rzehak et al., 2013; De Korte-De Boer et al., 2015; Magnus et al., 2015). Thus, the lack of adjusting for breast feeding and family asthma in childhood growth patterns model outcomes may not be substantial due to the same reason.

Fifth, during the investigation of the effects growth patterns on wheezing disorders using GMMs, *classify-analyse* approach was implemented. This would mean that the classification uncertainty was not incorporated into the models that investigated the effects of growth trajectory classes on wheezing disorders. The average class assignment error for 3 class growth models was 3%, 20% and 24% (class 1, class 2 and class 3, respectively) (Table 4.28). Thus, although the average latent class probability were above the minimum recommended cut-off point of 0.70 (Nagin, 2005) and the classification quality (entropy) for the three classes model was 0.90

(Table 4.27) which is above the recommended adequate cut-off point of 0.80 (Clark, 2010), the results must be interpreted with caution.

Nonetheless, this study has several strengths. First, this study was based on a contemporary prospective cohort data that anthropometric measurements were collected by trained workers (Raynor, 2008; Wright et al., 2013).

Second, clinical records were used in identifying cases of allergic conditions. Clinical records have minimal errors in excluding cases as opposed to questionnaires and clinical diagnosis data. The drug and prescription data need to be recorded for reimbursement which is an incentive for records to be accurate. In addition, unlike 12-months period or point prevalence which measure the disease burden during a limited period, lifetime prevalence figures provide a clearer picture about the absolute burden of disease and are therefore more helpful for health policy makers.

Third, during the analyses of childhood growth pattern, a more advanced analytic technique (latent growth modelling) was used to analyse life-course growth trajectories. Unlike mixed effects regression and multilevel spline models which both assume homogenous growth within a group, latent growth models allow for individuals to vary according to their distinct growth trajectories. These growth trajectories may provide greater insight in predicting the risk of childhood or early adulthood diseases in life-course studies.

Fourth, the application of FIML in missing data estimation to minimise parameter estimate biases was also an advantage as compared to list-wise and pair-wise deletion methods under missing data at random assumptions (Enders, 2001a; Enders and Bandalos, 2001). Likewise, the use of multiple imputations was also an advantage when compared to a complete cases analysis.

Fifth, the use of age-specific and sex-specific standardised weight scores have the advantage of clearly depicting the growth patterns of children in reference to the standard growth reference (WHO, 2006). The standard scores are convertible to percentiles (Pan and Cole, 2012) which can then be compared with the growth charts used by clinicians or growth monitoring workers in their daily practice.

Sixth, in the birthweight and wheezing disorders analyses, the sample size was reasonably large, in which the risk estimates were precise in the case of low birthweight.

Seventh, all modelling process throughout the thesis were informed by DAGs using DAGitty software, techniques that reduce potential bias due to confounding variables (Greenland et al., 1999; Tu et al., 2005; Textor et al., 2011).

5.8 Conclusion

The study shows that the burden of allergic conditions in the BiB cohort is higher than previously reported by earlier studies. Boys are more likely to suffer from wheezing disorders, rhinitis and multiple allergic conditions than girls. Pakistani children are more likely to suffer from eczema, rhinitis and multiple allergic conditions than white British children.

The study also confirmed that Pakistani and white British children have distinct growth patterns, that is, Pakistani children are lighter at birth and have a faster growth than White British children between birth and 3 years. More importantly, the study also showed that the children displayed heterogenous growth, that is, three distinct growth patterns although the size some of the growth classes was very small. However, the growth patterns may provide better insight in predicting the risk of childhood or early adulthood diseases in life course research.

There is a strong evidence to suggest that low birthweight is associated with increased risk of childhood wheezing disorders whilst there is a weak evidence to suggest that high birthweight children are associated with a reduced risk. Thus, increasing birthweight may reduce the impact of wheezing disorders in the Bradford community. However, by enhancing birthweight, there is also a possibility of increasing the likelihood of other health problems such as diabetes and cardiac diseases in South Asian children so health policies may also have to incorporate measures to tackle high birthweight problem at the same time.

Although results of BMI and growth based on the BiB1000 data remain inconclusive, keeping physical fitness of children may reduce the impact of childhood wheezing disorders and other diseases in the community of Bradford. In addition, maintaining optimal prenatal and postnatal growths may also reduce a risk of childhood wheezing disorders in the Bradford population.

CHAPTER 6

SUMMARY

6.1 Chapter overview

The purpose of this chapter is to summarise the findings of all previous chapters in order to make coherent recommendations mainly about the effects of birthweight, BMI and childhood growth patterns on childhood wheezing disorders.

6.2 Thesis overview

The thesis is made up of four parts. Part 1 consisted of Chapter 1, and provided a gentle introduction to childhood allergic diseases and childhood wheezing disorders in particular. Hypothetical mechanisms by which childhood wheezing disorders can be associated with childhood anthropometric measurements (i.e. birthweight, BMI and growth patterns) are also described. Then, the motivations of this thesis are discussed.

Part 2 (Chapter 2) investigated the association between childhood anthropometric measurements and childhood wheezing disorders using past epidemiologic studies. It described methods of systematic review and meta-analysis used along with meta-analyses results using *random effects* models, and included discussions and conclusions of the findings in detail.

Part 3 (Chapters 3, 4, and 5) mainly investigated the effect of birthweight, weight at the age of 3 years, and childhood growth patterns on wheezing disorders using BiB cohort data and novel statistical analyses techniques. Analysis variables were selected using *DAGs* and missing data were appropriately estimated using *FIML* and *MI*. *LGMs* were used for growth patterns analyses. Results from the series of analyses are described and discussed in detail.

The fourth and final part of this thesis consists of the present chapter, in which it summarises the results of the findings from past epidemiologic studies and the BiB cohort data. It also discusses what the implications of the findings are and the areas for further research.

6.3 Summary of past epidemiologic studies and BiB cohort results

This thesis has analysed and presented results from past epidemiologic studies and using a contemporary cohort data about the effects of birthweight, BMI, and childhood growth on wheezing disorders. In addition, the thesis also presented and discussed the results of the incidence and burden of allergic disease analyses using the BiB cohort data.

6.3.1 Past epidemiologic studies

The key findings from the systematic reviews and meta-analyses of past epidemiologic studies data can be summarised as follows:

a) Based on unadjusted and adjusted summary risk estimates, low birthweight children have significantly higher risk of childhood wheezing disorders.
b) Based on unadjusted summary risk estimates, high birthweight children have insignificant higher risk of wheezing disorders than normal birthweight children.
c) Based on unadjusted and adjusted summary risk estimates, overweight and obese children have significantly higher risk of childhood wheezing disorders than normal BMI children.
d) The effect of underweight on childhood wheezing disorders is inconclusive; that is, there is a significant and insignificant reduction of risk based on unadjusted and adjusted risk estimates, respectively.
e) The effect of childhood growth patterns on childhood wheezing disorders remains inconclusive; no summarising of risk estimates was conducted.

6.3.2 Born in Bradford cohort data

The key findings from the series of analyses using the birth cohort data can be summarised as follows:

a) Boys are more likely to suffer from wheezing disorders, rhinitis and multiple allergic conditions than girls.
b) Pakistani children are more likely to suffer from eczema, rhinitis and multiple allergic conditions than white British children.

c) Low birthweight children have significantly higher risk of childhood wheezing disorders, whilst there is a weak evidence for associated decreased risk for being high birthweight.
d) Overweight and obese children have an insignificant higher risk of wheezing disorders, whilst, underweight children have an insignificant reduced risk.
e) Pakistani children are lighter at birth but grow faster than white British children after birth until the age of three years.
f) Low birthweight is associated with faster growth until the age of 3 years.
g) Slow growth during the first 3 months is associated with higher risk of wheezing disorders
h) Fast growth between 3 and 12 months of age is associated with increased risk of wheezing disorders.
i) Slow growth between 12 and 36 months of age is associated with decreased risk of wheezing disorders.

6.4 Areas for further research

The important limitation of this thesis is that it was not possible to produce conclusive evidence about the effect of growth trajectories on childhood wheezing disorders. Producing a quantitative summary risk estimate from past epidemiologic studies was not possible; and, the GMMs using BiB cohort data produced inconsistent results. Meta-analysis is one way of enhancing the statistical power of analysis by combining data from different studies. However, there could be another way of aggregating data from different birth cohorts by using *harmonisation* methods. Analyses based on harmonised data can then have enough statistical power to detect any risk of wheezing disorders associated with childhood growth trajectories. Thus, further researches can be done using *harmonised* data of birth cohorts in order to explore the effect of growth trajectories on childhood wheezing disorders.

The thesis has not investigated the potential childhood wheezing disorder phenotypes in the BiB cohort. It could be of interest to examine this and possibly investigate its relationship with childhood anthropometric measurements in further research. Although missing extent on BMI and growth data of the BiB1000 (1,598

children) is substantial, birthweight and wheezing disorders of BiB cohort (13,734 children) are complete which would give enough statistical power for further analyses.

6.5 Discussion

Results from the past observational epidemiologic studies (i.e. through a systematic review and meta-analysis) and BiB cohort data agree that there is a strong evidence to suggest that low birthweight is associated with increased risk of childhood wheezing disorders. The two findings also agree that there is a weak evidence to suggest that there is increased (based on meta-analysis of studies) or decreased (based on BiB cohort data) risk of childhood wheezing disorders for high birthweight.

The summary estimates of wheezing disorders indicate that overweight and obesity are strongly associated with childhood wheezing disorders, although this was not replicated by the BiB1000 cohort results. This is probably due to lack of power in the BiB cohort data analyses. However, given that the meta-analyses results were based on more than 1 million children and the fact that the adjusted and unadjusted risk summary risk estimates agreed, it can be suggested that overweight and obesity are strongly associated with increased risk of childhood wheezing disorders.

The evidence for the effect of underweight on wheezing disorders has been inconclusive. Based on unadjusted summary risk estimates from 7 studies (>700 000 children), there is strong evidence for a reduced risk of the disease, whilst the adjusted summary estimates from 4 studies (<100 000 children) indicated that the evidence is weak (i.e. insignificant). However, risk estimates (adjusted and unadjusted) based on the BiB1000 cohort data of 1,598 children indicate that there is a weak evidence for reduced risk of childhood wheezing disorders.

The evidence about the effect of childhood growth remains inconclusive due to two main reasons. First, it was not possible to produce summary risk estimates of past epidemiologic studies as the weight measurements and developmental stages of children were diverse. Second, analyses for the effect of growth trajectory classes on childhood wheezing disorders using the BiB1000 cohort data lacked statistical

power to detect the risk. However, results based on the cohort data indicate that there is strong evidence to suggest that a decrease of 1SDS during the first 3 months, an increase of 1SDS between 3 and 12 months and between 12 and 36 months is associated with increased risk of wheezing disorders.

6.6 Conclusions

This thesis has confirmed that low birthweight, overweight and obesity are risk factors for childhood wheezing disorders. It has added that not only fast growth but also slow growth during early age can predispose to childhood wheezing disorders. In addition, it indicates that whilst boys are more likely to suffer from wheezing disorders, rhinitis and multiple allergic conditions than girls, Pakistani children are more likely to suffer from eczema, rhinitis and multiple allergic conditions than white British children.

Thus, if low birthweight and overweight/obesity are modifiable risk factors for childhood wheezing disorders, policies that aim at increasing birthweight, maintaining optimal childhood growth and keeping physical fitness could significantly reduce treating and caring cost of childhood wheezing disorders. However, it must be noted that enhancing birthweight in South Asian children may lead to diabetes and cardiac diseases. This indicates that, although increasing birthweight may not have adverse effects on the white British and other ethnicities of the Bradford children, it may not be so for the South Asians in general and Pakistani in specific. Therefore, policies that aim at enhancing birthweight in the Bradford community could be beneficial if they also incorporate measures to tackle high birthweight problems at the same time.

REFERENCES

Ahmad, N., Biswas, S., Bae, S., Meador, K.E.S., Huang, R. and Singh, K.P. 2009. Association between Obesity and asthma in US children and adolescents. *Journal of Asthma.* **46**(7), pp.642-646.

Akaike, H. 1987. Factor analysis and AIC. *Psychometrika.* **52**(3), pp.317-332.

Akaike, H. 1998. Prediction and entropy. In: Parzen, E., et al. eds. *Selected papers of Hirotugu Akaike.* New York: Springer, pp.387-410.

Akinbami, L., Moorman, J., Simon, A. and Schoendorf, K. 2014. Trends in racial disparities for asthma outcomes among children 0 to 17 years, 2001-2010. *J Allergy Clin Immunol.* **134**(3).

Al-Kubaisy, W., Ali, S.H. and Al-Thamiri, D. 2005. Risk factors for asthma among primary school children in Baghdad, Iraq. *Saudi Medical Journal.* **26**(3), pp.460-466.

Anand, D., Stevenson, C.J., West, C.R. and Pharoah, P.O.D. 2003. Lung function and respiratory health in adolescents of very low birth weight. *Archives of Disease in Childhood.* **88**(2), pp.135-138.

Anderson, E.L., Fraser, A., Martin, R.M., Kramer, M.S., Oken, E., Patel, R. et al. 2013. Associations of postnatal growth with asthma and atopy: The PROBIT Study. *Pediatric Allergy and Immunology.* **24**(2), pp.122-130.

Annesi-Maesano, I., Moreau, D. and Strachan, D. 2001. In utero and perinatal complications preceding asthma. *Allergy.* **56**(6), pp.491-497.

Arshad, S.H., Stevens, M. and Hide, D.W. 1993. The effect of genetic and environmental factors on the prevalence of allergic disorders at the age of two years. *Clinical & Experimental Allergy.* **23**(6), pp.504-511.

Asher, M., Keil, U., Anderson, H., Beasley, R., Crane, J., Martinez, F. et al. 1995. International Study of Asthma and Allergies in Childhood (ISAAC): rationale and methods. *European Respiratory Journal.* **8**(3), pp.483-491.

Asher, M.I., Montefort, S., Björkstén, B., Lai, C.K.W., Strachan, D.P., Weiland, S.K. et al. 2006. Worldwide time trends in the prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and eczema in childhood: ISAAC Phases One and Three repeat multicountry cross-sectional surveys. *The Lancet.* **368**(9537), pp.733-743.

Asparouhov, T. and Muthén, B. 2012. Auxiliary variables in mixture modeling: A 3-step approach using Mplus. *Mplus Web Notes.* **15**.

Asparouhov, T. and Muthén, B. 2014. Auxiliary Variables in Mixture Modeling: Three-Step Approaches Using Mplus. *Structural Equation Modeling: A Multidisciplinary Journal.* **21**(3), pp.329-341.

Asthma-UK (2014) 'Asthma Facts', [online], available: http://www.asthma.org.uk/asthma-facts-and-statistics. [accessed 02 February, 2015].

Azizi, B.H.O., Zulkifli, H.I. and Kasim, M.S. 1995. Indoor Air Pollution and Asthma in Hospitalized Children in a Tropical Environment. *Journal of Asthma.* **32**(6), pp.413-418.

Baker, J.L., Olsen, L.W. and Sørensen, T.I. 2007. Childhood body-mass index and the risk of coronary heart disease in adulthood. *New England Journal of Medicine.* **357**(23), pp.2329-2337.

Balfour-Lynn, I. and Openshaw, P. 2002. Viral infection. In: Silverman, M. ed. *Childhood Asthma and other Wheezing Disorders.* Second ed. Great Britain: Arnold, pp.205-217.

Balfour-Lynn, I.M. 1996. Why do viruses make infants wheeze? *Archives of disease in childhood.* **74**(3), pp.251-259.

Ballinger, G.A. 2004. Using Generalized Estimating Equations for Longitudinal Data Analysis. *Organizational Research Methods.* **7**(2), pp.127-150.

Bansal, N., Ayoola, O.O., Gemmell, I., Vyas, A., Koudsi, A., Oldroyd, J. et al. 2008. Effects of early growth on blood pressure of infants of British European and South Asian origin at one year of age: the Manchester children's growth and vascular health study. *Journal of hypertension.* **26**(3), pp.412-418.

Bateman, E.D., Hurd, S.S., Barnes, P.J., Bousquet, J., Drazen, J.M., FitzGerald, M. et al. 2008. Global strategy for asthma management and prevention: GINA executive summary. *Eur Respir J.* **31**(1), pp.143-178.

Bradford District Infant Mortality Commission (BDIMC). 2007. 'Bradford District Infant Mortality Commission Report 2006', [online], available: http://www.observatory.bradford.nhs.uk/SiteCollectionDocuments/Infant%20Mortality%20Report%202006.pdf [accessed May 20, 2015].

Bradford District Infant Mortality Commission (BDIMC). 2014. 'Bradford Joint Strategic Needs Assessment: Infant mortality in Bradford district', [online], available: http://www.observatory.bradford.nhs.uk/Documents/4_5_1_Infant_Mortality_2014.pdf [accessed May 22, 2015].

Belgrave, D.C.M., Granell, R., Simpson, A., Guiver, J., Bishop, C., Buchan, I. et al. 2014. Developmental Profiles of Eczema, Wheeze, and Rhinitis: Two Population-Based Birth Cohort Studies. *PLoS Med.* **11**(10), pe1001748.

Benicio, M.H.D.A., Ferreira, M.U., Cardoso, M.R.A., Konno, S.C. and Monteiro, C.A. 2004. Wheezing conditions in early childhood: Prevalence and risk factors in the city of Sao Paulo, Brazil. *Bulletin of the World Health Organization.* **82**(7), pp.516-522.

Bentler, P.M. 1990. Comparative fit indexes in structural models. *Psychological Bulletin.* **107**(2), p238.

Bernsen, R.M.D., De Jongste, J.C., Koes, B.W., Aardoom, H.A. and Van Der Wouden, J.C. 2005. Perinatal characteristics and obstetric complications as risk factors for asthma, allergy and eczema at the age of 6 years. *Clinical and Experimental Allergy.* **35**(9), pp.1135-1140.

Bibi, H., Shoseyov, D., Feigenbaum, D., Genis, M., Friger, M., Peled, R. et al. 2004. The relationship between asthma and obesity in children: Is it real or a case of over diagnosis? *Journal of Asthma.* **41**(4), pp.403-410.

Bjerg, A., Hedman, L., Perzanowski, M., Lundback, B. and Ronmark, E. 2011. A strong synergism of low birth weight and prenatal smoking on asthma in schoolchildren. *Pediatrics.* **127**(4), pp.e905-912.

Black, M.H., Smith, N., Porter, A.H., Jacobsen, S.J. and Koebnick, C. 2012. Higher prevalence of obesity among children with asthma. *Obesity.* **20**(5), pp.1041-1047.

Bolte, G., Schmidt, M., Maziak, W., Keil, U., Nasca, P., Von Mutius, E. et al. 2004. The relation of markers of fetal growth with asthma, allergies and serum immunoglobulin E levels in children at age 5-7 years. *Clinical and Experimental Allergy.* **34**(3), pp.381-388.

Boodhna, G. and Hall, J. (2011) 'Health Survey for England 2010: Respiratory Symptoms and Asthma in children', [online], available: http://www.hscic.gov.uk/catalogue/PUB03023/heal-surv-eng-2010-resp-heal-ch4-symp-chil.pdf [accessed 10 March 2015].

Borenstein, M., Hedges, L.V., Higgins, J.P.T. and Rothstein, H.R. 2009a. Fixed-Effect Versus Random-Effects Models. *Introduction to Meta-analysis.* 1st ed. UK: John Wiley & Sons Ltd, pp.59-102.

Borenstein, M., Hedges, L.V., Higgins, J.P.T. and Rothstein, H.R. 2009b. Fixed-Effect Versus Random-Effects Models. *Introduction to Meta-analysis.* First ed. UK: John Wiley & Sons Ltd, pp.59-133.

Borenstein, M., Hedges, L.V., Higgins, J.P.T. and Rothstein, H.R. 2009c. Heterogeneity. *Introduction to Meta-analysis.* 1st ed. UK: John Wiley & Sons Ltd, pp.105-133.

Borenstein, M., Hedges, L.V., Higgins, J.P.T. and Rothstein, H.R. 2009d. Meta-Regression. *Introduction to Meta-analysis.* 1st ed. UK: John Wiley & Sons Ltd, pp.187-203.

Borenstein, M., Hedges, L.V., Higgins, J.P.T. and Rothstein, H.R. 2009e. Subgroup analysis. *Introduction to Meta-analysis.* 1st ed. UK: John Wiley & Sons Ltd, pp.149-186.

Boyd, A., Golding, J., Macleod, J., Lawlor, D.A., Fraser, A., Henderson, J. et al. 2012. Cohort Profile: The 'Children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology.*

Brand, P.L., Baraldi, E., Bisgaard, H., Boner, A., Castro-Rodriguez, J., Custovic, A. et al. 2008. Definition, assessment and treatment of wheezing disorders in preschool children: an evidence-based approach. *European Respiratory Journal.* **32**(4), pp.1096-1110.

Brew, B.K., Marks, G.B. and Investigators, C. 2012. Perinatal factors and respiratory health in children. *Clinical & Experimental Allergy.* **42**(11), pp.1621-1629.

Brooks, A.M., Byrd, R.S., Weitzman, M., Auinger, P. and McBride, J.T. 2001. Impact of low birth weight on early childhood asthma in the United States. *Archives of Pediatrics and Adolescent Medicine.* **155**(3), pp.401-406.

Browne, M.W. and Cudeck, R. 1992. Alternative ways of assessing model fit. *Sociological Methods & Research.* **21**(2), pp.230-258.

Bryk, A.S. and Raudenbush, S.W. 1987. Application of hierarchical linear models to assessing change. *Psychological Bulletin.* **101**(1), p147.

Burgess, J.A., Walters, E.H., Byrnes, G.B., Giles, G.G., Jenkins, M.A., Abramson, M.J. et al. 2007. Childhood adiposity predicts adult-onset current asthma in females: a 25-yr prospective study. *European Respiratory Journal.* **29**(4), pp.668-675.

Burke, H., Leonardi-Bee, J., Hashim, A., Pine-Abata, H., Chen, Y., Cook, D.G. et al. 2012. Prenatal and passive smoke exposure and incidence of asthma and wheeze: systematic review and meta-analysis. *Pediatrics.* **129**(4), pp.735-744.

Bush, A. and Fleming, L. 2015. Diagnosis and management of asthma in children.

Cassol, V.E., Rizzato, T.M., Teche, S.P., Basso, D.F., Hirakata, V.N., Maldonado, M. et al. 2005. Prevalence and severity of asthma among adolescents and their relationship with the body mass index. *J Pediatr (Rio J).* **81**(4), pp.305-309.

Castro-Rodríguez, J.A. 2007. Relationship Between Obesity and Asthma. *Archivos de Bronconeumología ((English Edition)).* **43**(3), pp.171-175.

Castro-Rodriguez, J.A., Holberg, C.J., Morgan, W.J., Wright, A.L. and Martinez, F.D. 2001. Increased incidence of asthmalike symptoms in girls who become overweight or obese during the school years. *American Journal of Respiratory & Critical Care Medicine.* **163**(6), pp.1344-1349.

Centers for Disease Control and Prevention (CDC). 2009 .Pediatric and Pregnancy Nutrition Surveillance System: PedNSS Health Indicators', [online], available: http://www.cdc.gov/pednss/what_is/pednss_health_indicators.htm [accessed 27 May, 2014].

Centers for Disease Control and Prevention (CDC). 2014.': Healthy Weight - it's not a diet, it's a lifestyle: What is a BMI percentile?', [online], available: http://www.cdc.gov/healthyweight/assessing/bmi/childrens_bmi/about_childrens_bmi.html [accessed 22 July, 2014].

Chatkin, M.N. and Menezes, A.M.B. 2005. The association between low birthweight and asthma: a systematic literature review. *Revista Panamericana de Salud Publica.* **17**(2), pp.102-109.

Chatzi, L., Torrent, M., Romieu, I., Garcia-Esteban, R., Ferrer, C., Vioque, J. et al. 2007. Diet, wheeze, and atopy in school children in Menorca, Spain. *Pediatric Allergy and Immunology.* **18**(6), pp.480-485.

Chen, F., Bollen, K.A., Paxton, P., Curran, P.J. and Kirby, J.B. 2001. Improper Solutions in Structural Equation Models: Causes, Consequences, and Strategies. *Sociological Methods & Research.* **29**(4), pp.468-508.

Chen, Y.C., Dong, G.H., Lin, K.C. and Lee, Y.L. 2013. Gender difference of childhood overweight and obesity in predicting the risk of incident asthma: a systematic review and meta-analysis. *Obesity Reviews.* **14**(3), pp.222-231.

Chinn, S. and Rona, R.J. 2001. Can the increase in body mass index explain the rising trend in asthma in children? *Thorax.* **56**(11), pp.845-850.

Cibella, F., Cuttitta, G., La Grutta, S., Melis, M.R., Bucchieri, S. and Viegi, G. 2011. A cross-sectional study assessing the relationship between BMI, asthma, atopy, and eNO among schoolchildren. *Annals of Allergy, Asthma and Immunology.* **107**(4), pp.330-336.

Clark, S.L. 2010. *Mixture modeling with behavioral data*. PhD thesis, University of California. Available from: http://statmodel.com/papers_date.shtml

Cole, T.J., Bellizzi, M.C., Flegal, K.M. and Dietz, W.H. 2000. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ.* **320**(7244), p1240.

Cole, T.J., Flegal, K.M., Nicholls, D. and Jackson, A.A. 2007. Body mass index cut offs to define thinness in children and adolescents: international survey. *BMJ.* **335**(7612), p194.

Collins, L.M., Schafer, J.L. and Kam, C.-M. 2001. A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods.* **6**(4), p330.

Custovic, A., Simpson, B.M., Murray, C.S., Lowe, L., Woodcock, A., on behalf of the, N.A.C.M.A. et al. 2002. The National Asthma Campaign Manchester Asthma and Allergy Study. *Pediatric Allergy and Immunology.* **13**, pp.32-37.

Davidson, R., Roberts, S.E., Wotton, C.J. and Goldacre, M.J. 2010. Influence of maternal and perinatal factors on subsequent hospitalisation for asthma in children: Evidence from the Oxford record linkage study. *BMC Pulmonary Medicine.* **10**(14).

Davis, A., Lipsett, M., Milet, M., Etherton, M. and Kreutzer, R. 2007. An association between asthma and BMI in adolescents: results from the California Healthy Kids Survey. *Journal of Asthma.* **44**(10), pp.873-879.

De Korte-De Boer, D., Mommers, M., Thijs, C., Jaminon, M., Jansen, M., Mujakovic, S. et al. 2015. Early life growth and the development of preschool wheeze, independent from overweight: The LucKi Birth Cohort Study. *Journal of Pediatrics.* **166**(2), pp.343-349.e341.

Dempster, A.P., Laird, N.M. and Rubin, D.B. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society. Series B (methodological).* pp.1-38.

DerSimonian, R. and Laird, N. 1986. Meta-analysis in clinical trials. *Controlled Clinical Trials.* **7**(3), pp.177-188.

Dex, S. and Joshi, H. 2005. *Children of the 21st century: from birth to nine months.* Policy Press.

Duncan, T.E. and Duncan, S.C. 2004. An introduction to latent growth curve modeling. *Behavior Therapy.* **35**(2), pp.333-363.

Duncan, T.E., Duncan, S.C., Hops, H. and Stoolmiller, M. 1995. An analysis of the relationship between parent and adolescent marijuana use via generalized estimating equation methodology. *Multivariate Behavioral Research.* **30**(3), pp.317-339.

Duncan, T.E., Duncan, S.C. and Strycker, L.A. 2006a. CHAPTER ONE: Introduction. In: Marcoulides, G.A. ed. *An introduction to latent growth curve modeling: Concepts Issues and Applications.* 2nd ed. London: Lawrence Erlbaum Associates,Inc., pp.1-13.

Duncan, T.E., Duncan, S.C. and Strycker, L.A. 2006b. Growth Mixture Modeling. In: G.A., M. ed. *An introduction to Latent Growth Curve Modeling.* 2nd ed. New Jersy, : Lawrence Erlbaum associates, pp.125-149.

Duncan, T.E., Duncan, S.C. and Strycker, L.A. 2006c. *An introduction to latent variable growth curve modeling: Concepts, issues, and application.* Second ed. Routledge Academic.

Duncan, T.E., Duncan, S.C. and Strycker, L.A. 2006d. Piecewise and Pooled Interrupted Time Series LGMs. In: Marcoulides, G.A. ed. *An introduction to latent growth curve modeling: Concepts Issues and Applications.* 2nd ed. London: Lawrence Erlbaum Associates,Inc., pp.151-164.

Eczema-UK (2015) 'The National Eczema Society: What is Eczema?', [online], available: http://www.eczema.org/what-is-eczema [accessed 02 February, 2015].

Egan, K.B., Ettinger, A.S. and Bracken, M.B. 2013. Childhood body mass index and subsequent physician-diagnosed asthma: A systematic review and meta-analysis of prospective cohort studies. *BMC Pediatrics.* **13**(1).

Egger, M., Smith, G.D., Schneider, M. and Minder, C. 1997. Bias in meta-analysis detected by a simple, graphical test. *BMJ.* **315**(7109), pp.629-634.

Enders, C.K. 2001a. The impact of nonnormality on full information maximum-likelihood estimation for structural equation models with missing data. *Psychological Methods.* **6**(4), p352.

Enders, C.K. 2001b. A primer on maximum likelihood algorithms available for use with missing data. *Structural Equation Modeling.* **8**(1), pp.128-141.

Enders, C.K. and Bandalos, D.L. 2001. The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling.* **8**(3), pp.430-457.

Eriksson, J.G., Forsen, T.J., Kajantie, E., Osmond, C. and Barker, D.J.P. 2007. Childhood growth and hypertension in later life. *Hypertension.* **49**(6), pp.1415-1421.

Eriksson, J.G., Forsen, T.J., Osmond, C. and Barker, D.J.P. 2003. Pathways of infant and childhood growth that lead to type 2 diabetes. *Diabetes Care.* **26**(11), pp.3006-3010.

Fairley, L., Petherick, E.S., Howe, L.D., Tilling, K., Cameron, N., Lawlor, D.A. et al. 2013. Describing differences in weight and length growth trajectories between white and Pakistani infants in the UK: analysis of the Born in Bradford birth cohort study using multilevel linear spline models. *Archives of Disease in Childhood.* **98**(4), pp.274-279.

Farah, C.S. and Salome, C.M. 2012. Asthma and obesity: A known association but unknown mechanism. *Respirology.* **17**(3), pp.412-421.

Fergusson, D.M., Crane, J., Beasley, R. and Horwood, L.J. 1997. Perinatal factors and atopic disease in childhood. *Clinical & Experimental Allergy.* **27**(12), pp.1394-1401.

Flaherman, V. and Rutherford, G.W. 2006. A meta-analysis of the effect of high weight on asthma. *Archives of Disease in Childhood.* **91**(4), pp.334-339.

Fleiss, J.L., Levin, B. and Paik, M.C. 2003. *Statistical methods for rates and proportions.* John Wiley & Sons.

Flexeder, C., Thiering, E., Bruske, I., Koletzko, S., Bauer, C.P., Wichmann, H.E. et al. 2012. Growth velocity during infancy and onset of asthma in school-aged children. *Allergy.* **67**(2), pp.257-264.

Galli, S.J., Tsai, M. and Piliponsky, A.M. 2008. The development of allergic inflammation. *Nature.* **454**(7203), pp.445-454.

Garcia-Marcos, L., Valverde-Molina, J., Ortega, M.L.C., Sanchez-Solis, M., Martinez-Torres, A.E. and Castro-Rodriguez, J.A. 2008. Percent body fat, skinfold thickness or body mass index for defining obesity or overweight, as a risk factor for asthma in

schoolchildren: which one to use in epidemiological studies? *Matern Child Nutr.* **4**(4), pp.304-310.

Gennuso, J., Epstein, L.H., Paluch, R.A. and Cerny, F. 1998. The relationship between asthma and obesity in urban minority children and adolescents. *Archives of Pediatrics & Adolescent Medicine.* **152**(12), pp.1197-1200.

Ghouri, N., Hippisley-Cox, J., Newton, J. and Sheikh, A. 2008. Trends in the epidemiology and prescribing of medication for allergic rhinitis in England. *Journal of the Royal Society of Medicine.* **101**(9), pp.466-472.

Gilliland, F.D., Berhane, K., Islam, T., McConnell, R., Gauderman, W.J., Gilliland, S.S. et al. 2003. Obesity and the risk of newly diagnosed asthma in school-age children. *American Journal of Epidemiology.* **158**(5), pp.406-415.

Gilthorpe, M., Dahly, D., Tu, Y.-K., Kubzansky, L. and Goodman, E. 2014. Challenges in modelling the random structure correctly in growth mixture models and the impact this has on model mixtures. *Journal of developmental origins of health and disease.* **5**(03), pp.197-205.

GINA. 2015. *Global Strategy for Asthma Management and Prevention 2014. Chapter 6: Diagnosis and management of asthma in children 5 years and younger.*

Gold, D.R., Burge, H.A., Carey, V., Milton, D.K., Platts-Mills, T. and Weiss, S.T. 1999. Predictors of repeated wheeze in the first year of life: the relative roles of cockroach, birth weight, acute lower respiratory illness, and maternal smoking. *American Journal of Respiratory & Critical Care Medicine.* **160**(1), pp.227-236.

Gold, D.R., Damokosh, A.I., Dockery, D.W. and Berkey, C.S. 2003. Body-mass index as a predictor of incident asthma in a prospective cohort of children. *Pediatric Pulmonology.* **36**(6), pp.514-521.

Graham, J.W., Olchowski, A.E. and Gilreath, T.D. 2007. How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science.* **8**(3), pp.206-213.

Grammatikos, A.P. 2008. The genetic and environmental basis of atopic diseases. *Annals of Medicine.* **40**(7), pp.482-495.

Granell, R., Henderson, A.J., Evans, D.M., Smith, G.D., Ness, A.R., Lewis, S. et al. 2014. Effects of BMI, Fat Mass, and Lean Mass on Asthma in Childhood: A Mendelian Randomization Study. *PLoS medicine.* **11**(7).

Greenland, S., Pearl, J. and Robins, J.M. 1999. Causal diagrams for epidemiologic research. *Epidemiology.* **10**(1), pp.37-48.

Gregory, A., Doull, I., Pearce, N., Cheng, S., Leadbitter, P., Holgate, S. et al. 1999. The relationship between anthropometric measurements at birth: asthma and atopy in childhood. *Clinical & Experimental Allergy.* **29**(3), pp.330-333.

Grigg, J. 2002. Inflammation. In: Silverman, M. ed. *Childhood Asthma and other Wheezing Disorders.* Second ed. Great Britain: Arnold, pp.173-180.

Guibas, G.V., Manios, Y., Xepapadaki, P., Moschonis, G., Douladiris, N., Mavrogianni, C. et al. 2013. The obesity-asthma link in different ages and the role of Body Mass Index in its investigation: Findings from the Genesis and Healthy Growth Studies. *Allergy: European Journal of Allergy and Clinical Immunology.* **68**(10), pp.1298-1305.

Guler, N., Kirerleri, E., Ones, U., Tamay, Z., Salmayenli, N. and Darendeliler, F. 2004. Leptin: Does it have any role in childhood asthma? *Journal of Allergy and Clinical Immunology.* **114**(2), pp.254-259.

Halldorsson, T.I., Gunnarsdottir, I., Birgisdottir, B.E., Gudnason, V., Aspelund, T. and Thorsdottir, I. 2011. Childhood growth and adult hypertension in a population of high birth weight. *Hypertension.* **58**(1), pp.8-15.

Hamling, J., Lee, P., Weitkunat, R. and Ambühl, M. 2008. Facilitating meta-analyses by deriving relative effect and precision estimates for alternative comparisons from a

set of estimates presented by exposure level or disease category. *Statistics in Medicine.* **27**(7), pp.954-970.

Harding, S., Rosato, M. and Cruickshank, J. 2004. Lack of change in birthweights of infants by generational status among Indian, Pakistani, Bangladeshi, Black Caribbean, and Black African mothers in a British cohort study. *International Journal of Epidemiology.* **33**(6), pp.1279-1285.

Hardy, R., Wadsworth, M.E.J., Langenberg, C. and Kuh, D. 2004. Birthweight, childhood growth, and blood pressure at 43 years in a British birth cohort. *International Journal of Epidemiology.* **33**(1), pp.121-129.

Hasnain, S.M., Khan, M., Saleem, A. and Waqar, M.A. 2009. Prevalence of asthma and allergic rhinitis among school children of Karachi, Pakistan, 2007. *Journal of Asthma.* **46**(1), pp.86-90.

He, Q.Q., Wong, T.W., Du, L., Jiang, Z.Q., Qiu, H., Gao, Y. et al. 2009. Respiratory health in overweight and obese Chinese children. *Pediatric Pulmonology.* **44**(10), pp.997-1002.

Hedges, L.V. 2009a. Research Synthesis as a scientific process. In: Cooper, H., et al. eds. *The Handbook of Research Synthesis and Meta-analysis.* 2nd ed. 112 East 64th Street, New York: Russel Sage Foundation, pp.3-16.

Hedges, L.V. 2009b. Statistical considerations. In: Cooper, H., et al. eds. *The Handbook of Research Synthesis and Meta-analysis.* 2nd ed. 112 East 64th Street, New York: Russel Sage Foundation, pp.37-47.

Hernán, M.A., Hernández-Díaz, S., Werler, M.M. and Mitchell, A.A. 2002. Causal Knowledge as a Prerequisite for Confounding Evaluation: An Application to Birth Defects Epidemiology. *American Journal of Epidemiology.* **155**(2), pp.176-184.

Higgins, J.P.T. and Thompson, S.G. 2002. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine.* **21**(11), pp.1539-1558.

Hislop, A.A. and Pandya, H.C. 2002. Structural development. In: Silverman, M. ed. *Childhood Asthma and other Wheezing Disorders.* Second ed. Great Britain: Arnold, pp.37-51.

Ho, W.-C., Lin, Y.-S., Caffrey, J., Lin, M.-H., Hsu, H.-T., Myers, L. et al. 2011. Higher body mass index may induce asthma among adolescents with pre-asthmatic symptoms: a prospective cohort study. *BMC Public Health.* **11**(1), p542.

Holt, P.G. 2002. Postnatal maturation of immune and inflammatory functions. In: Silverman, M. ed. *Childhood Asthma and other Wheezing Disorders.* Second ed. Great Britain: Arnold, pp.69-77.

Howe, L.D., Tilling, K., Matijasevich, A., Petherick, E.S., Santos, A.C., Fairley, L. et al. 2013. Linear spline multilevel models for summarising childhood growth trajectories: A guide to their application using examples from five birth cohorts. *Statistical Methods in Medical Research.*

Hwang, H. and Takane, Y. 2005. Estimation of growth curve models with structured error covariances by generalized estimating equations. *Behaviormetrika.* **32**(2), pp.155-163.

Jaakkola, J.J.K., Ahmed, P., Ieromnimon, A., Goepfert, P., Laiou, E., Quansah, R. et al. 2006. Preterm delivery and asthma: A systematic review and meta-analysis. *Journal of Allergy and Clinical Immunology.* **118**(4), pp.823-830.

Jackson, S., Mathews, K.H., Pulanić, D., Falconer, R., Rudan, I., Campbell, H. et al. 2013. Risk factors for severe acute lower respiratory infections in children–a systematic review and meta-analysis. *Croatian medical journal.* **54**(2), pp.110-121.

Jacobson, J.S., Mellins, R.B., Garfinkel, R., Rundle, A.G., Perzanowski, M.S., Chew, G.L. et al. 2008. Asthma, body mass, gender, and Hispanic national origin among 517 preschool children in New York City. *Allergy: European Journal of Allergy and Clinical Immunology.* **63**(1), pp.87-94.

Jaddoe, V.W.V., Troe, E.-J.W.M., Hofman, A., Mackenbach, J.P., Moll, H.A., Steegers, E.A.P. et al. 2008. Active and passive maternal smoking during pregnancy and the risks of low birthweight and preterm birth: the Generation R Study. *Paediatric and Perinatal Epidemiology.* **22**(2), pp.162-171.

Jeong, Y., Jung-Choi, K., Lee, J.H., Lee, H.Y., Park, E.A., Kim, Y.J. et al. 2010. Body weight at birth and at age three and respiratory illness in preschool children. *Journal of Preventive Medicine & Public Health / Yebang Uihakhoe Chi.* **43**(5), pp.369-376.

Kabesch, M. and Von Mutius, E. 2002. Epidemiology and Public Health. In: Silverman, M. ed. *Childhood Asthma and other Wheezing Disorders.* Second ed. Great Britain: Arnold, pp.9-28.

Kallen, B., Finnstrom, O., Nygren, K.-G. and Otterblad Olausson, P. 2013. Association between preterm birth and intrauterine growth retardation and child asthma. *European Respiratory Journal.* **41**(3), pp.671-676.

Kiechl-Kohlendorfer, U., Horak, E., Mueller, W., Strobl, R., Haberland, C., Fink, F.-M. et al. 2007. Neonatal characteristics and risk of atopic asthma in schoolchildren: results from a large prospective birth-cohort study. *Acta Paediatrica.* **96**(11), pp.1606-1610.

Kozyrskyj, A.L., Kendall, G.E., Jacoby, P., Sly, P.D. and Zubrick, S.R. 2010. Association between socioeconomic status and the development of asthma: analyses of income trajectories. *American Journal of Public Health.* **100**(3), pp.540-546.

Kramer, M.S. 1987. Determinants of low birth weight: methodological assessment and meta-analysis. *Bulletin of the World Health Organization.* **65**(5), p663.

Krishnaveni, G., Hill, J., Veena, S., Leary, S., Saperia, J., Chachyamma, K. et al. 2005. Truncal adiposity is present at birth and in early childhood in South Indian children. *Indian pediatrics.* **42**(6), p527.

Kuschnir, F.C. and da Cunha, A.L.A. 2009. Association of overweight with asthma prevalence in adolescents in Rio de Janeiro, Brazil. *Journal of Asthma.* **46**(9), pp.928-932.

Kusunoki, T., Morimoto, T., Nishikomori, R., Heike, T., Ito, M., Hosoi, S. et al. 2008. Obesity and the prevalence of allergic diseases in schoolchildren. *Pediatric Allergy & Immunology*. **19**(6), pp.527-534.

Kwon, H.L., Ortiz, B., Swaner, R., Shoemaker, K., Jean-Louis, B., Northridge, M.E. et al. 2006. Childhood asthma and extreme values of body mass index: The Harlem Children's Zone Asthma Initiative. *Journal of Urban Health.* **83**(3), pp.421-433.

Lanza, S.T., Tan, X. and Bray, B.C. 2013. Latent Class Analysis With Distal Outcomes: A Flexible Model-Based Approach. *Structural Equation Modeling: A Multidisciplinary Journal.* **20**(1), pp.1-26.

Lau, J., Ioannidis, J.P.A. and Schmid, C.H. 1998. Summing up evidence: one answer is not always enough. *The Lancet.* **351**(9096), pp.123-127.

Leadbitter, P., Pearce, N., Cheng, S., Sears, M.R., Holdaway, M.D., Flannery, E.M. et al. 1999. Relationship between fetal growth and the development of asthma and atopy in childhood. *Thorax.* **54**(10), pp.905-910.

Lee, Y.L., Chen, Y.-C. and Chen, Y.-A. 2013. Obesity and the occurrence of bronchitis in adolescents. *Obesity.* **21**(1), pp.E149-E153.

Lewis, S., Butland, B., Strachan, D., Bynner, J., Richards, D., Butler, N. et al. 1996. Study of the aetiology of wheezing illness at age 16 in two national British birth cohorts. *Thorax.* **51**(7), pp.670-676.

Lewis, S., Richards, D., Bynner, J., Butler, N. and Britton, J. 1995. Prospective study of risk factors for early and persistent wheezing in childhood. *European Respiratory Journal.* **8**(3), pp.349-356.

Liang, K.-Y. and Zeger, S.L. 1986. Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika.* **73**(1), pp.13-22.

Lilljeqvist, A.C., Faleide, A.O. and Watten, R.G. 1997. Low birthweight, environmental tobacco smoke, and air pollution: Risk factors for childhood asthma? *Pediatric Asthma, Allergy and Immunology.* **11**(2), pp.95-102.

Lim, R.H., Kobzik, L. and Dahl, M. 2010. Risk for Asthma in Offspring of Asthmatic Mothers versus Fathers: A Meta-Analysis. *PLoS ONE [Electronic Resource].* **5(4): e10134**(4).

Liu, X., Olsen, J., Agerbo, E., Yuan, W., Cnattingius, S., Gissler, M. et al. 2014. Birth weight, gestational age, fetal growth and childhood asthma hospitalization. *Allergy, Asthma and Clinical Immunology.* **10**(1).

Lu, F.L., Hsieh, C.-J., Caffrey, J.L., Lin, M.-H., Lin, Y.-S., Lin, C.-C. et al. 2012. Body mass index may modify asthma prevalence among low-birth-weight children. *American Journal of Epidemiology.* **176**(1), pp.32-42.

Lu, Y.-P., Zeng, D.-Y., Chen, Y.-P., Liang, X.-J., Xu, J.-P., Huang, S.-M. et al. 2013. Low birth weight is associated with lower respiratory tract infections in children with hand, foot, and mouth disease. *Clin. Lab.* **9**(10), p1.

Lucas, S.R. and Platts-Mills, T.A.E. 2006. Paediatric asthma and obesity. *Paediatric Respiratory Reviews.* **7**(4), pp.233-238.

Luo, Z.-C., Wilkins, R., Kramer, M.S., Fetal, f.t. and System, I.H.S.G.o.t.C.P.S. 2006. Effect of neighbourhood income and maternal education on birth outcomes: a population-based study. *Canadian Medical Association Journal.* **174**(10), pp.1415-1420.

Magnus, M.C., Stigum, H., Haberg, S.E., Nafstad, P., London, S.J. and Nystad, W. 2015. Peak weight and height velocity to age 36 months and asthma development: the norwegian mother and child cohort study. *PLoS One.*

Magnusson, J.O., Kull, I., Mai, X.M., Wickman, M. and Bergstrom, A. 2012. Early childhood overweight and asthma and allergic sensitization at 8 years of age. *Pediatrics.* **129**(1), pp.70-76.

Mai, X.-M., Almqvist, C., Nilsson, L. and Wickman, M. 2007. Birth anthropometric measures, body mass index and allergic diseases in a birth cohort study (BAMSE). *Archives of Disease in Childhood.* **92**(10), pp.881-886.

Mai, X.-M., Gaddlin, P.-O., Nilsson, L. and Leijon, I. 2005. Early rapid weight gain and current overweight in relation to asthma in adolescents born with very low birth weight. *Pediatric Allergy & Immunology.* **16**(5), pp.380-385.

Malik, V.S., Schulze, M.B. and Hu, F.B. 2006. Intake of sugar-sweetened beverages and weight gain: a systematic review. *The American Journal of Clinical Nutrition.* **84**(2), pp.274-288.

Mamun, A.A., Lawlor, D.A., Alati, R., O'Callaghan, M.J., Williams, G.M. and Najman, J.M. 2007. Increasing body mass index from age 5 to 14 years predicts asthma among adolescents: evidence from a birth cohort study. *Int J Obes.* **31**(4), pp.578-583.

Mannino, D.M., Mott, J., Ferdinands, J.M., Camargo, C.A., Friedman, M., Greves, H.M. et al. 2006. Boys with high body masses have an increased risk of developing asthma: findings from the National Longitudinal Survey of Youth (NLSY). *International Journal of Obesity.* **30**(1), pp.6-13.

Margetts, B.M., Mohd Yusof, S., Al Dallal, Z. and Jackson, A.A. 2002. Persistence of lower birth weight in second generation South Asian babies born in the United Kingdom. *Journal of Epidemiology and Community Health.* **56**(9), pp.684-687.

Masoli, M., Fabian, D., Holt, S., Beasley, R. and Global Initiative for Asthma, P. 2004. The global burden of asthma: executive summary of the GINA Dissemination Committee Report. *Allergy.* **59**(5), pp.469-478.

Mathew, A.C., Prince, T.G., Remees, R., Saravanapandian, N., Ramalingam, S., Srikanth, K. et al. 2012. Prevalence and risk factors of asthma in school going children in South India. *Nepal Journal of Epidemiology.* **2**(1), pp.171-178.

Matos, S.M., Jesus, S.R., Saldiva, S.R., Prado, M.S., D'Innocenzo, S., Assis, A.M. et al. 2011. Overweight, asthma symptoms, atopy and pulmonary function in children of 4-12 years of age: findings from the SCAALA cohort in Salvador, Bahia, Brazil. *Public health nutrition.* **14**(7), pp.1270-1278.

McKeever, T., Lewis, S., Smith, C., Collins, J., Heatlie, H., Frischer, M. et al. 2001. Siblings, multiple births, and the incidence of allergic disease: a birth cohort study using the West Midlands general practice research database. *Thorax.* **56**(10), pp.758-762.

Mckenzie, S. 2002. The clinical features and their assessment. In: Silverman, M. ed. *Childhood Asthma and other Wheezing Disorders.* Second ed. Great Britain: Arnold, pp.106-119.

McLachlan, G.J. 1987. On Bootstrapping the Likelihood Ratio Test Stastistic for the Number of Components in a Normal Mixture. *Journal of the Royal Statistical Society. Series C (Applied Statistics).* **36**(3), pp.318-324.

McNamee, R. 2003. Confounding and confounders. *Occupational and Environmental Medicine.* **60**(3), pp.227-234.

McNutt, L.-A., Wu, C., Xue, X. and Hafner, J.P. 2003. Estimating the Relative Risk in Cohort Studies and Clinical Trials of Common Outcomes. *American Journal of Epidemiology.* **157**(10), pp.940-943.

Mebrahtu, T., Feltbower, R., Greenwood, D. and Parslow, R. (2014) 'Birth weight, childhood asthma and wheezing disorders: a systematic review and meta-analysis', [online], available: http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014010424 [accessed May 02, 2014].

Midodzi, W.K., Rowe, B.H., Majaesic, C.M., Saunders, L.D. and Senthilselvan, A. 2010. Early life factors associated with incidence of physician-diagnosed asthma in preschool children: results from the Canadian Early Childhood Development cohort study. *Journal of Asthma.* **47**(1), pp.7-13.

Mitchell, E.A., Beasley, R., Bjorksten, B., Crane, J., Garcia-Marcos, L. and Keil, U. 2013. The association between BMI, vigorous physical activity and television viewing and the risk of symptoms of asthma, rhinoconjunctivitis and eczema in children and adolescents: ISAAC Phase Three. *Clinical and Experimental Allergy.* **43**(1), pp.73-84.

Mitchell, E.A., Clayton, T., García-Marcos, L., Pearce, N., Foliaki, S., Wong, G. et al. 2014. Birthweight and the risk of atopic diseases: the ISAAC Phase III study. *Pediatric Allergy and Immunology.* **25**(3), pp.264-270.

Miyake, Y. and Tanaka, K. 2013. Lack of relationship between birth conditions and allergic disorders in Japanese children aged 3 years. *Journal of Asthma.* **50**(6), pp.555-559.

Mogensen, N., Larsson, H., Lundholm, C. and Almqvist, C. 2011. Association between childhood asthma and ADHD symptoms in adolescence – a prospective population-based twin study. *Allergy.* **66**(9), pp.1224-1230.

Moher, D., Liberati, A., Tetzlaff, J. and Altman, D.G. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ.* **339**.

Mu, M., Ye, S., Bai, M.-J., Liu, G.-L., Tong, Y., Wang, S.-F. et al. 2014. Birth Weight and Subsequent Risk of Asthma: A Systematic Review and Meta-Analysis. *Heart, Lung and Circulation.* **23**(6), pp.511-519.

Muthen, B. 2001. Latent variable mixture modeling. In: Marcoulides, G.A. and Schumacker, R.E. eds. *New developments and techniques in structural equation modeling.* Mahwah, New Jersey: Lawrence Erlbaum Associates, pp.1-33.

Muthén, B. 2004. Latent variable analysis:Growth Mixture Modeling and Related Techniques for Longitudinal Data. In: Kaplan, B.A. ed. *The handbook of quantitative methodology for the social sciences* Thousand Oaks, CA: Sage Publications, pp.345-368.

Muthén, L. and Muthén, B. 2012. *Mplus User's Guide. Seventh Edition: 1998-2012.* Los Angeles, CA: Muthén & Muthén.

Nagin, D. 2005. *Group-based modeling of development.* Harvard University Press.

Nelder, J.A. and Baker, R. 1972. Generalized linear models. *Journal of the Royal Statistical Society.* **135**(3), pp.370-384.

Nepomnyaschy, L. and Reichman, N.E. 2006. Low birthweight and asthma among young urban children.[Erratum appears in Am J Public Health. 2006 Oct;96(10):1723]. *American Journal of Public Health.* **96**(9), pp.1604-1610.

Netuveli, G., Hurwitz, B., Levy, M., Fletcher, M., Barnes, G., Durham, S.R. et al. 2005. Ethnic variations in UK asthma frequency, morbidity, and health-service use: a systematic review and meta-analysis. *Lancet.* **365**(9456), pp.312-317.

Noal, R.B., Menezes, A.M.B., Macedo, S.E.C., Dumith, S.C., Perez-Padilla, R., Araujo, C.L.P. et al. 2012. Is obesity a risk factor for wheezing among adolescents? A prospective study in southern Brazil. *Journal of Adolescent Health.* **51**(6 Suppl), pp.S38-45.

Noational Obesity Observatory (NOO). 2011. 'A simple guide to classifying body mass index in children', [online], available: http://www.noo.org.uk/uploads/doc/vid_11601_A_simple_guide_to_classifying_BMI_in_children.pdf [accessed 18 September, 2015].

Nuolivirta, K., Koponen, P., Helminen, M. and Korppi, M. 2012. Weight gain in infancy and post-bronchiolitis wheezing. *Acta Paediatrica, International Journal of Paediatrics.* **101**(1), pp.38-42.

Nylund, K.L., Asparouhov, T. and Muthén, B.O. 2007. Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling.* **14**(4), pp.535-569.

Okabe, Y., Adachi, Y., Itazawa, T., Yoshida, K., Ohya, Y., Odajima, H. et al. 2012. Association between obesity and asthma in japanese preschool children. *Pediatric Allergy and Immunology.* **23**(6), pp.550-555.

Okabe, Y., Itazawa, T., Adachi, Y., Yoshida, K., Ohya, Y., Odajima, H. et al. 2011. Association of overweight with asthma symptoms in Japanese school children. *Pediatrics International.* **53**(2), pp.192-198.

Ong, K.K.L., Preece, M.A., Emmett, P.M., Ahmed, M.L. and Dunger, D.B. 2002. Size at Birth and Early Childhood Growth in Relation to Maternal Smoking, Parity and Infant Breast-Feeding: Longitudinal Birth Cohort Study and Analysis. *Pediatr Res.* **52**(6), pp.863-867.

Orsini, N., Li, R., Wolk, A., Khudyakov, P. and Spiegelman, D. 2012. Meta-Analysis for Linear and Nonlinear Dose-Response Relations: Examples, an Evaluation of Approximations, and Software. *American Journal of Epidemiology.* **175**(1), pp.66-73.

Ortqvist, A.K., Lundholm, C., Carlstrom, E., Lichtenstein, P., Cnattingius, S. and Almqvist, C. 2009. Familial factors do not confound the association between birth weight and childhood asthma. *Pediatrics.* **124**(4), pp.e737-e743.

Osman, M., Tagiyeva, N., Wassall, H.J., Ninan, T.K., Devenny, A.M., McNeill, G. et al. 2007. Changing trends in sex specific prevalence rates for childhood asthma, eczema, and hay fever. *Pediatric Pulmonology.* **42**(1), pp.60-65.

Owen, C.G., Whincup, P.H., Orfei, L., Chou, Q.-A., Rudnicka, A.R., Wathern, A.K. et al. 2009. Is body mass index before middle age related to coronary heart disease risk in

later life? Evidence from observational studies. *International Journal of Obesity*. **33**(8), pp.866-877.

Pan, H. and Cole, T. (2012) 'LMSgrowth, a Microsoft Excel add-in to access growth references based on the LMS method. Version 2.77', [online], available: http://www.healthforallchildren.com/?product=lmsgrowth [accessed October 30, 2013].

Panico, L., Stuart, B., Bartley, M. and Kelly, Y. 2014. Asthma trajectories in early childhood: identifying modifiable factors. *PLoS One.*

Pearce, N., Aït-Khaled, N., Beasley, R., Mallol, J., Keil, U., Mitchell, E. et al. 2007. Worldwide trends in the prevalence of asthma symptoms: phase III of the International Study of Asthma and Allergies in Childhood (ISAAC). *Thorax.* **62**(9), pp.758-766.

Pike, K.C., Crozier, S.R., Lucas, J.S.A., Inskip, H.M., Robinson, S., Southampton Women's Survey Study, G. et al. 2010. Patterns of fetal and infant growth are related to atopy and wheezing disorders at age 3 years. *Thorax.* **65**(12), pp.1099-1106.

Punekar, Y.S. and Sheikh, A. 2009. Establishing the incidence and prevalence of clinician-diagnosed allergic conditions in children and adolescents using routinely collected data from general practices. *Clinical & Experimental Allergy.* **39**(8), pp.1209-1216.

Raqib, R., Alam, D.S., Sarker, P., Ahmad, S.M., Ara, G., Yunus, M. et al. 2007. Low birth weight is associated with altered immune function in rural Bangladeshi children: a birth cohort study. *The American Journal of Clinical Nutrition.* **85**(3), pp.845-852.

Rasanen, M., Kaprio, J., Laitinen, T., Winter, T., Koskenvuo, M. and Laitinen, L.A. 2000. Perinatal risk factors for asthma in Finnish adolescent twins. *Thorax.* **55**(1), pp.25-31.

Raynor, P. 2008. Born in Bradford, a cohort study of babies born in Bradford, and their parents: Protocol for the recruitment phase. *BMC Public Health.* **8**(1), p327.

Reis, G.G., Miranda, V.M., Cardoso, M.R.A., Sole, D., Barral, A. and Nascimento-Carvalho, C.M. 2015. Prevalence and risk factors for wheezing in Salvador, Brazil: A population-based study. *Qjm.* **108**(3), pp.213-218.

Remes, S.T., Patel, S.P., Hartikainen, A.L., Jarvelin, M.R. and Pekkanen, J. 2008. High birth weight, asthma and atopy at the age of 16 yr. *Pediatric Allergy & Immunology.* **19**(6), pp.541-543.

Renauld, J.-C. 2001. New insights into the role of cytokines in asthma. *Journal of Clinical Pathology.* **54**(8), pp.577-589.

Roche, R. and Jeffery, R. 2002. Remodelling and inflammation. In: Silverman, M. ed. *Childhood Asthma and other Wheezing Disorders.* Second ed. Great Britain: Arnold, pp.93-101.

Rodríguez, M.A., Winkleby, M.A., Ahn, D., Sundquist, J. and Kraemer, H.C. 2002. Identification of population subgroups of children and adolescents with high asthma prevalence: Findings from the third national health and nutrition examination survey. *Archives of Pediatrics & Adolescent Medicine.* **156**(3), pp.269-275.

Rona, R., Gulliford, M. and Chinn, S. 1993. Effects of prematurity and intrauterine growth on respiratory health and lung function in childhood. *BMJ: British Medical Journal.* **306**(6881), p817.

Ronmark, E., Perzanowski, M., Platts-Mills, T. and Lundback, B. 2002. Incidence rates and risk factors for asthma among school children: a 2-year follow-up report from the obstructive lung disease in Northern Sweden (OLIN) studies. *Respiratory Medicine.* **96**(12), pp.1006-1013.

Royston, P., Carlin, J.B. and White, I.R. 2009. Multiple imputation of missing values: new features for mim. *Stata Journal.* **9**(2), p252.

Royston, P. and White, I.R. 2011. Multiple imputation by chained equations (MICE): implementation in Stata. *Journal of Statistical Software.* **45**(4), pp.1-20.

Rubin, D.B. 1987. *Multiple imputations for non-response in surveys.* New York: John Wiley & Sons.

Rücker, G., Schwarzer, G., Carpenter, J.R. and Schumacher, M. 2008. Undue reliance on I2 in assessing heterogeneity may mislead. *BMC Medical Research Methodology.* **8**(1), p79.

Rzehak, P., Wijga, A.H., Keil, T., Eller, E., Bindslev-Jensen, C., Smit, H.A. et al. 2013. Body mass index trajectory classes and incident asthma in childhood: Results from 8 European Birth Cohorts—a Global Allergy and Asthma European Network initiative. *Journal of Allergy and Clinical Immunology.* **131**(6), pp.1528-1536.

Saha, C., Riner, M.E. and Liu, G. 2005. Individual and neighborhood-level factors in predicting asthma. *Archives of Pediatrics and Adolescent Medicine.* **159**(8), pp.759-763.

Sahota, P., Gatenby, L.A., Greenwood, D.C., Bryant, M., Robinson, S. and Wright, J. 2015. Ethnic differences in dietary intake at age 12 and 18 months: the Born in Bradford 1000 Study. *Public health nutrition.* pp.1-9.

Saxena, S., Ambler, G., Cole, T.J. and Majeed, A. 2004. Ethnic group differences in overweight and obese children and young people in England: cross sectional survey. *Archives of Disease in Childhood.* **89**(1), pp.30-36.

Schafer, J.L. and Graham, J.W. 2002. Missing Data: Our View of the State of the Art. *Psychological Methods.* **7**(2), pp.147-177.

Schaubel, D., Johansen, H., Dutta, M., Desmeules, M., Becker, A. and Mao, Y. 1996. Neonatal characteristics as risk factors for preschool asthma. *Journal of Asthma.* **33**(4), pp.255-264.

Schisterman, E.F., Cole, S.R. and Platt, R.W. 2009. Overadjustment Bias and Unnecessary Adjustment in Epidemiologic Studies. *Journal of Epidemilogy.* **20**(4), pp.488-495.

Scholtens, S., Wijga, A.H., Seidell, J.C., Brunekreef, B., de Jongste, J.C., Gehring, U. et al. 2009. Overweight and changes in weight status during childhood in relation to asthma symptoms at 8 years of age. *Journal of Allergy and Clinical Immunology.* **123**(6), pp.1312-1318.e1312.

Schwartz, J., Gold, D., Dockery, D.W., Weiss, S.T. and Speizer, F.E. 1990. Predictors of asthma and persistent wheeze in a national sample of children in the United States. Association with social class, perinatal events, and race. *American Review of Respiratory Disease.* **142**(3), pp.555-562.

Schwarz, G. 1978. Estimating the dimension of a model. *The annals of statistics.* **6**(2), pp.461-464.

Sears, M.R., Holdaway, M.D., Flannery, E.M., Herbison, G.P. and Silva, P.A. 1996. Parental and neonatal risk factors for atopy, airway hyper-responsiveness, and asthma. *Archives of Disease in Childhood.* **75**(5), pp.392-398.

Seidman, D.S., Laor, A., Gale, R., Stevenson, D.K. and Danon, Y.L. 1991. Is low birth weight a risk factor for asthma during adolescence? *Archives of Disease in Childhood.* **66**(5), pp.584-587.

Shadish, W.R. and Haddock, C.K. 2009. Combining Estimates of Effect Size. In: Cooper, H., et al. eds. *The Handbook of Research Synthesis and Meta-analysis.* 2nd ed. 112 East 64th Street, New York: Russel Sage Foundation, pp.257-277.

Shamssain, M.H. 2006. The association between overweight and respiratory symptoms in schoolchildren. *Pediatric Asthma, Allergy and Immunology.* **19**(1), pp.19-25.

Shrier, I. and Platt, R. 2008. Reducing bias through directed acyclic graphs. *BMC Medical Research Methodology.* **8**(1), p70.

Silva, R.D.C.R., Assis, A.M.O., Goncalves, M.S., Fiaccone, R.L., Matos, S.M.A., Barreto, M.L. et al. 2013. Prevalence of Wheezing and its Association with Body Mass Index and Abdominal Obesity in Children. *Journal of Asthma.* **50**(3), pp.267-273.

Silverman, M. 2002. Wheezing disorders in infants and young children. In: Silverman, M. ed. *Childhood Asthma and other Wheezing Disorders.* Second ed. Great Britain: Arnold, pp.307-326.

Simpson, C.R., Newton, J., Hippisley-Cox, J. and Sheikh, A. 2009. Trends in the epidemiology and prescribing of medication for eczema in England. *Journal of the Royal Society of Medicine.* **102**(3), pp.108-117.

Sin, D.D., Spier, S., Svenson, L.W., Schopflocher, D.P., Senthilselvan, A., Cowie, R.L. et al. 2004. The relationship between birth weight and childhood asthma: a population-based cohort study. *Archives of Pediatrics & Adolescent Medicine.* **158**(1), pp.60-64.

Singh, A.M., Moore, P.E., Gern, J.E., Lemanske, R.F. and Hartert, T.V. 2007. Bronchiolitis to Asthma. *Am J Respir Crit Care Med.* **175**(2), pp.108-119.

Slezak, J.A., Persky, V.W., Kviz, F.J., Ramakrishnan, V. and Byers, C. 1998. Asthma prevalence and risk factors in selected Head Start sites in Chicago. *Journal of Asthma.* **35**(2), pp.203-212.

Sly, P.D., Boner, A.L., Björksten, B., Bush, A., Custovic, A., Eigenmann, P.A. et al. 2008. Early identification of atopy in the prediction of persistent asthma in children. *The Lancet.* **372**(9643), pp.1100-1106.

Sonnenschein-van der Voort, A.M., Arends, L.R., de Jongste, J.C., Annesi-Maesano, I., Arshad, S.H., Barros, H. et al. 2014a. Preterm birth, infant weight gain, and childhood asthma risk: a meta-analysis of 147,000 European children. *Journal of Allergy & Clinical Immunology.* **133**(5), pp.1317-1329.

Sonnenschein-van der Voort, A.M., Howe, L.D., Granell, R., Duijts, L., Sterne, J.A., Tilling, K. et al. 2014b. Influence of childhood growth on asthma and lung function in adolescence. *J Allergy Clin Immunol.*

Sonnenschein-van der Voort, A.M., Jaddoe, V.W.V., Raat, H., Moll, H.A., Hofman, A., de Jongste, J.C. et al. 2012. Fetal and Infant Growth and Asthma Symptoms in Preschool Children. *Am J Respir Crit Care Med.* **185**(7), pp.731-737.

Sontag, S.J. 2000. Gastroesophageal reflux disease and asthma. *Journal of clinical gastroenterology.* **30**(3 Suppl), pp.S9-30.

Stanley, D. 1952. *Davidson's Principles & Practices of Medicine* 21 ed. London: CHURCHILL LINVINGSTONE, Elsevier Limited.

StataCorp. 2011. Stata statistical software: Release 12.1. College Station, TX: Stata Corporation.

Steffensen, F.H., Sorensen, H.T., Gillman, M.W., Rothman, K.J., Sabroe, S., Fischer, P. et al. 2000. Low birth weight and preterm delivery as risk factors for asthma and atopic dermatitis in young adult males. *Epidemiology.* **11**(2), pp.185-188.

Sterne, J.A., Egger, M. and Smith, G.D. 2001. Investigating and dealing with publication and other biases in meta-analysis. *BMJ.* **323**(7304), pp.101-105.

Sterne, J.A. and Habord, R.M. 2004. Funnel Plots in meta-analysis. *The Stata Journal.* **4**(2), pp.127-141.

Sterne, J.A., White, I.R., Carlin, J.B., Spratt, M., Royston, P., Kenward, M.G. et al. 2009. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ.* **338**.

Suglia, S.F., Chambers, E.C., Rosario, A. and Duarte, C.S. 2011. Asthma and obesity in three-year-old urban children: Role of sex and home environment. *Journal of Pediatrics.* **159**(1), pp.14-20.e11.

Tai, A., Volkmer, R. and Burton, A. 2009. Association between asthma symptoms and obesity in preschool (4-5 year old) children. *Journal of Asthma.* **46**(4), pp.362-365.

Taveras, E.M., Camargo Jr, C.A., Rifas-Shiman, S.L., Oken, E., Gold, D.R., Weiss, S.T. et al. 2006. Association of birth weight with asthma-related outcomes at age 2 years. *Pediatric Pulmonology.* **41**(7), pp.643-648.

Textor, J., Hardt, J. and Knuppel, S. 2011. DAGitty: a graphical tool for analyzing causal diagrams. *Epidemiology.* **22**(5), p745.

To, T., Guan, J., Wang, C., Radhakrishnan, D., McLimont, S., Latycheva, O. et al. 2012. Is large birth weight associated with asthma risk in early childhood? *Archives of Disease in Childhood.* **97**(2), pp.169-171.

Tollefsen, E., Langhammer, A., Romundstad, P., Bjermer, L., Johnsen, R. and Holmen, T.L. 2007. Female gender is associated with higher incidence and more stable respiratory symptoms during adolescence. *Respiratory Medicine.* **101**(5), pp.896-902.

Tsai, H.J. and Tsai, A.C. 2009. The association of BMI and sedentary time with respiratory symptoms and asthma in 5th grade schoolchildren in Kaohsiung, Taiwan. *J Asthma.* **46**(1), pp.9-15.

Tsai, H.J., Tsai, A.C., Nriagu, J., Ghosh, D., Gong, M. and Sandretto, A. 2007. Associations of BMI, TV-watching time, and physical activity on respiratory symptoms and asthma in 5th grade schoolchildren in Taipei, Taiwan. *Journal of Asthma.* **44**(5), pp.397-401.

Tu, Y.-K., Tilling, K., Sterne, J.A. and Gilthorpe, M.S. 2013. A critical evaluation of statistical approaches to examining the role of growth trajectories in the developmental origins of health and disease. *International Journal of Epidemiology.* **42**(5), pp.1327-1339.

Tu, Y.-K., West, R., Ellison, G.T.H. and Gilthorpe, M.S. 2005. Why Evidence for the Fetal Origins of Adult Disease Might Be a Statistical Artifact: The "Reversal Paradox" for the Relation between Birth Weight and Blood Pressure in Later Life. *American Journal of Epidemiology.* **161**(1), pp.27-32.

Tucker, L.R. and Lewis, C. 1973. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika.* **38**(1), pp.1-10.

Turner, S., Thomas, M., von Ziegenweidt, J. and Price, D. 2009. Prescribing trends in asthma: a longitudinal observational study. *Archives of Disease in Childhood.* **94**(1), pp.16-22.

van Buuren, S. 2007. Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research.* **16**(3), pp.219-242.

Van De Ven, M.O., Van Den Eijnden, R.J. and Engels, R.C. 2006. Atopic diseases and related risk factors among Dutch adolescents. *The European Journal of Public Health.* **16**(5), pp.549-558.

van der Gugten, A.C., Koopman, M., Evelein, A.M.V., Verheij, T.J.M., Uiterwaal, C.S.P.M. and van der Ent, C.K. 2012. Rapid early weight gain is associated with wheeze and reduced lung function in childhood. *European Respiratory Journal.* **39**(2), pp.403-410.

Vargas, P.A., Perry, T.T., Robles, E., Jo, C.H., Simpson, P.M., Magee, J.M. et al. 2007. Relationship of body mass index with asthma indicators in Head Start children. *Annals of Allergy, Asthma and Immunology.* **99**(1), pp.22-28.

Vartanian, L.R., Schwartz, M.B. and Brownell, K.D. 2007. Effects of Soft Drink Consumption on Nutrition and Health: A Systematic Review and Meta-Analysis. *American Journal of Public Health.* **97**(4), pp.667-675.

Vazquez-Nava, F., Morales Romero, J., Cordova Fernandez, A., Saldivar-Gonzalez, A.H., Vazquez-Rodriguez, C.F., Barrientos Gomez, M.D.C. et al. 2010. Association between obesity and asthma in preschool Mexican children. *ScientificWorldJournal.* **10**, pp.1339-1346.

Vermunt, J.K. 2010. Latent class modeling with covariates: Two improved three-step approaches. *Political Analysis.* **18**(4), pp.450-469.

Viera, A.J. 2008. Odds Ratios and Risk Ratios: What's the Difference and Why Does It Matter? *Southern Medical Journal.* **101**(7), pp.730-734.

Visness, C.M., London, S.J., Daniels, J.L., Kaufman, J.S., Yeatts, K.B., Siega-Riz, A.-M. et al. 2010. Association of Childhood Obesity With Atopic and Nonatopic Asthma: Results From the National Health and Nutrition Examination Survey 1999–2006. *Journal of Asthma.* **47**(7), pp.822-829.

von Kries, R., Hermann, M., Grunert, V.P. and von Mutius, E. 2001. Is obesity a risk factor for childhood asthma? *Allergy.* **56**(4), pp.318-322.

Wang, D., Qian, Z., Wang, J., Yang, M., Lee, Y.L., Liu, F. et al. 2014. Gender-specific differences in associations of overweight and obesity with asthma and asthma-related symptoms in 30,056 children: result from 25 districts of Northeastern China. *Journal of Asthma.*

Wang, J. and Wang, X. 2012a. Mixture modeling: 6.3 Growth mixture modeling. *Structural equation modeling : applications Using Mplus.* John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, United Kingdom, pp.340-365.

Wang, J. and Wang, X. 2012b. *Structural equation modeling : applications Using Mplus.* John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, United Kingdom.

Wang, W.-H., Chen, P.-C., Hsieh, W.-S. and Lee, Y.L. 2012. Joint effects of birth outcomes and childhood body mass index on respiratory symptoms. *European Respiratory Journal.* **39**(5), pp.1213-1219.

Warner, J. and Warner, J. 2002. Allergy. In: Silverman, M. ed. *Childhood Asthma and other Wheezing Disorders.* Second ed. Great Britain: Arnold.

Weitzman, M., Gortmaker, S. and Sobol, A. 1990. Racial, social, and environmental risks for childhood asthma. *American Journal of Diseases of Children.* **144**(11), pp.1189-1194.

Wells, G., Shea, B., O'connell, D., Peterson, J., Welch, V., Losos, M. et al. (2000) 'The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses', [online], available: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp [accessed 20 May, 2014].

West, J., Lawlor, D.A., Fairley, L., Bhopal, R., Cameron, N., McKinney, P.A. et al. 2013. UK-born Pakistani-origin infants are relatively more adipose than white British infants: findings from 8704 mother-offspring pairs in the Born-in-Bradford prospective birth cohort. *Journal of Epidemiology and Community Health.* **67**(7), pp.544-551.

West, J., Lawlor, D.A., Fairley, L. and Wright, J. 2014. Differences in socioeconomic position, lifestyle and health-related pregnancy characteristics between Pakistani and White British women in the Born in Bradford prospective cohort study: the influence of the woman's, her partner's and their parents' place of birth. *BMJ Open.* **4**(6).

Whincup, P.H., Nightingale, C.M., Owen, C.G., Rudnicka, A.R., Gibb, I., McKay, C.M. et al. 2010. Early Emergence of Ethnic Differences in Type 2 Diabetes Precursors in the UK: The Child Heart and Health Study in England (CHASE Study). *PLoS Med.* **7**(4), pe1000263.

White, I.R., Royston, P. and Wood, A.M. 2011. Multiple imputation using chained equations: issues and guidance for practice. *Statistics in Medicine.* **30**(4), pp.377-399.

Whitehead, A. and Whitehead, J. 1991. A general parametric approach to the meta-analysis of randomized clinical trials. *Statistics in Medicine.* **10**(11), pp.1665-1677.

World Health Organisation (WHO). 2004. *Low Birthweight: Country, Regional and Global Estimates.* [Online]. New York: UNICEF. [Accessed 20 May 2014]. Available from: http://www.unicef.org/publications/index_24840.html

World Health Organisation (WHO).2006. 'WHO Child Growth Standards: Methods and development: Length/height-for-age, weight-for-age, weight-for-length, weight-for-

height and body mass index-for-age.', *Child growth standards* [online], available: http://www.who.int/childgrowth/publications/technical_report_pub/en/index.html [accessed February 04, 2015].

World Health Organisation (WHO). 2013. 'Asthma Fact Sheet Number 307', [online], available: http://www.who.int/features/factfiles/asthma/en/ [accessed 24 September, 2015].

World Health Organisation (WHO). 2014. 'International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10)-2014-WHO Version for ;2014:Disorders related to length of gestation and fetal growth (P05-P08) ', [online], available: http://apps.who.int/classifications/icd10/browse/2014/en#/P05-P08 [accessed 03 December, 2014].

Wickens, K., Barry, D., Friezema, A., Rhodius, R., Bone, N., Purdie, G. et al. 2005. Obesity and asthma in 11-12 year old New Zealand children in 1989 and 2000. *Thorax.* **60**(1), pp.7-12.

Willeboordse, M., van den Bersselaar, D.L.C.M., van de Kant, K.D.G., Muris, J.W.M., van Schayck, O.C.P. and Dompeling, E. 2013. Sex Differences in the Relationship between Asthma and Overweight in Dutch Children: A Survey Study. *Plos One.* **8**(10).

Willet, K.E. and Sly, P.D. 2002. Developmental physiology. In: Silverman, M. ed. *Childhood Asthma and other Wheezing Disorders.* Second ed. Great Britain: Arnold, pp.57-64.

Wjst, M., Popescu, M., Trepka, M.J., Heinrich, J. and Wichmann, H.E. 1998. Pulmonary function in children with initial low birth weight. *Pediatric Allergy and Immunology.* **9**(2), pp.80-90.

Wood, A.M., White, I.R. and Thompson, S.G. 2004. Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. *Clinical Trials.* **1**(4), pp.368-376.

Wright, J., Small, N., Raynor, P., Tuffnell, D., Bhopal, R., Cameron, N. et al. 2013. Cohort Profile: The Born in Bradford multi-ethnic family cohort study. *International Journal of Epidemiology.* **42**(4), pp.978-991.

Xu, B., Pekkanen, J., Laitinen, J. and Järvelin, M.R. 2002. Body build from birth to adulthood and risk of asthma. *The European Journal of Public Health.* **12**(3), pp.166-170.

Xu, X.-F., Li, Y.-J., Sheng, Y.-J., Liu, J.-L., Tang, L.-F. and Chen, Z.-M. 2014. Effect of low birth weight on childhood asthma: a meta-analysis. *BMC Pediatrics.* **14**, p275.

Xu, X., Dailey, A.B., Freeman, N.C., Curbow, B.A. and Talbott, E.O. 2009. The effects of birthweight and breastfeeding on asthma among children aged 1-5 years: ORIGINAL ARTICLE. *J Paediatr Child Health.* **45**(11), pp.646-651.

Yajnik, C., Fall, C., Coyaji, K., Hirve, S., Rao, S., Barker, D. et al. 2003. Neonatal anthropometry: the thin–fat Indian baby. The Pune maternal nutrition study. *International journal of obesity.* **27**(2), pp.173-180.

Yajnik, C., Lubree, H., Rege, S., Naik, S., Deshpande, J., Deshpande, S. et al. 2002. Adiposity and hyperinsulinemia in Indians are present at birth. *The Journal of Clinical Endocrinology & Metabolism.* **87**(12), pp.5575-5580.

Yang, C.-C. 2006. Evaluating latent class analysis models in qualitative phenotype identification. *Computational Statistics & Data Analysis.* **50**(4), pp.1090-1104.

Yang, H.J., Qin, R., Katusic, S. and Juhn, Y.J. 2013. Population-based study on association between birth weight and risk of asthma: a propensity score approach. *Annals of Allergy, Asthma, & Immunology.* **110**(1), pp.18-23.

Yao, T.C., Ou, L.S., Yeh, K.W., Lee, W.I., Chen, L.C. and Huang, J.L. 2011. Associations of age, gender, and BMI with prevalence of allergic diseases in children: PATCH study. *Journal of Asthma.* **48**(5), pp.503-510.

Yiallouros, P.K., Lamnisos, D., Kolokotroni, O., Moustaki, M. and Middleton, N. 2013. Associations of body fat percent and body mass index with childhood asthma by age and gender. *Obesity*. **21**(9), pp.E474-E482.

Yuan, W., Basso, O., Sorensen, H.T. and Olsen, J. 2002. Fetal growth and hospitalization with asthma during early childhood: a follow-up study in Denmark. *International Journal of Epidemiology*. **31**(6), pp.1240-1245.

Yuan, W., Fonager, K., Olsen, J. and Sørensen, H. 2003. Prenatal factors and use of anti-asthma medications in early childhood: A population-based Danish birth cohort study. *European Journal of Epidemiology*. **18**(8), pp.763-768.

Zhang, Z., Lai, H.J., Roberg, K.A., Gangnon, R.E., Evans, M.D., Anderson, E.L. et al. 2010. Early childhood weight status in relation to asthma development in high-risk children. *Journal of Allergy and Clinical Immunology*. **126**(6), pp.1157-1162.

Appendix A Quality assessment scale for cohort studies

NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE
COHORT STUDIES

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability

Selection

1) Representativeness of the exposed cohort
   a) truly representative of the average _____ (describe) in the community ✸
   b) somewhat representative of the average _____ in the community ✸
   c) selected group of users eg nurses, volunteers
   d) no description of the derivation of the cohort

2) Selection of the non exposed cohort
   a) drawn from the same community as the exposed cohort ✸
   b) drawn from a different source
   c) no description of the derivation of the non exposed cohort

3) Ascertainment of exposure
   a) secure record (eg surgical records) ✸
   b) structured interview ✸
   c) written self report
   d) no description

4) Demonstration that outcome of interest was not present at start of study
   a) yes ✸
   b) no

Comparability

1) Comparability of cohorts on the basis of the design or analysis
   a) study controls for _____ (select the most important factor) ✸
   b) study controls for any additional factor ✸ (This criteria could be modified to indicate specific control for a second important factor.)

Outcome

1) Assessment of outcome
   a) independent blind assessment ✸
   b) record linkage ✸
   c) self report
   d) no description

2) Was follow-up long enough for outcomes to occur
   a) yes (select an adequate follow up period for outcome of interest) ✸
   b) no

3) Adequacy of follow up of cohorts
   a) complete follow up - all subjects accounted for ✸
   b) subjects lost to follow up unlikely to introduce bias - small number lost - > ____ % (select an adequate %) follow up, or description provided of those lost) ✸
   c) follow up rate < ____% (select an adequate %) and no description of those lost
   d) no statement

Appendix B Quality assessment scale for case-control studies

NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE
CASE CONTROL STUDIES

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Exposure categories. A maximum of two stars can be given for Comparability.

Selection

1) Is the case definition adequate?
   a) yes, with independent validation ✸
   b) yes, eg record linkage or based on self reports
   c) no description

2) Representativeness of the cases
   a) consecutive or obviously representative series of cases ✸
   b) potential for selection biases or not stated

3) Selection of Controls
   a) community controls ✸
   b) hospital controls
   c) no description

4) Definition of Controls
   a) no history of disease (endpoint) ✸
   b) no description of source

Comparability

1) Comparability of cases and controls on the basis of the design or analysis
   a) study controls for _____ (Select the most important factor.) ✸
   b) study controls for any additional factor ✸ (This criteria could be modified to indicate specific control for a second important factor.)

Exposure

1) Ascertainment of exposure
   a) secure record (eg surgical records) ✸
   b) structured interview where blind to case/control status ✸
   c) interview not blinded to case/control status
   d) written self report or medical record only
   e) no description

2) Same method of ascertainment for cases and controls
   a) yes ✸
   b) no

3) Non-Response rate
   a) same rate for both groups ✸
   b) non respondents described
   c) rate different and no designation

Appendix C Quality assessment scale for cross-sectional studies

NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE
(adapted for cross-sectional studies)

Selection: (Maximum 3 stars)

1) Representativeness of the sample:
- a) Truly representative of the average in the target population. ✵ (all subjects or random sampling)
- b) Somewhat representative of the average in the target population. ✵ (non-random sampling)
- c) Selected group of users.
- d) No description of the sampling strategy.

2) Non-respondents:
- a) Comparability between respondents and non-respondents characteristics is established, and the response rate is satisfactory. ✵
- b) The response rate is unsatisfactory, or the comparability between respondents and non-respondents is unsatisfactory.
- c) No description of the response rate or the characteristics of the responders and the non-responders.

3) Ascertainment of the exposure (risk factor):
- a) Validated measurement tool. ✵
- b) Non-validated measurement tool, but the tool is available or described.
- c) No description of the measurement tool.

Comparability: (Maximum 2 stars)

1) The subjects in different outcome groups are comparable, based on the study design or analysis. Confounding factors are controlled.
- a) The study controls for the most important factor (select one). ✵
- b) The study control for any additional factor. ✵

Outcome: (Maximum 2 stars)

1) Assessment of the outcome:
- a) Independent blind assessment. ✵
- b) Record linkage. ✵
- c) Self report.
- d) No description.

2) Statistical test:
- a) The statistical test used to analyze the data is clearly described and appropriate, and the measurement of the association is presented, including confidence intervals and the probability level (p value). ✵
- b) The statistical test is not appropriate, not described or incomplete.

Appendix D Summary of studies that investigated the effect of birthweight on wheezing disorders using non-standard birthweight categories

Author and year	Comparison	Unadjusted odds ratio (95% CI)	Adjusted odds ratio (95% CI)
Gold et al, 1999	<3.2kg Vs 3.2-3.8kg	1.43 (0.86, 2.39)	
	>3.8kg Vs 3.2-3.8kg	0.61 (0.33, 1.13)	
Yuan et al, 2003	<3.2kg Vs 3.2-3.8kg	1.13 (0.86, 1.49)	
	>3.8kg Vs 3.2-3.8kg	1.00 (0.79, 1.27)	
Sin et al, 2004	<2.5kg Vs 2.5-4.5kg		1.00 (0.90, 1.11)*
	>4.5kg Vs 2.5-4.5kg		1.16 (1.04, 1.29)*
Mai et al, 2007	<2.9kg Vs 2.9-4.2kg	1.70 (1.19, 2.43)	1.47 (0.87, 2.49)
	>4.2kg Vs 2.9-4.2kg	1.27 (0.86, 1.87)	1.18 (0.74, 1.87)
Garcia-Marcos et al,	<2.0kg Vs 2.0-3.5kg	0.52 (0.12, 2.22)	
2008	>3.5kg Vs 2.0-3.5kg	1.04 (0.65, 1.69)	
Davidson et al, 2010	<3.0kg Vs 3.0-4.0kg	1.17 (1.08, 1.26)	1.21 (1.13, 1.30)
	>4.0kg Vs 3.0-4.0kg	1.10 (0.97, 1.24)	1.05 (0.93, 1.18)
Jeong et al, 2010	<2.8kg Vs 2.8-3.3kg	0.29 (0.09, 0.92)	0.56 (0.16, 1.96)
	>3.3kg Vs 2.8-3.3kg	0.45 (0.17, 1.22)	0.29 (0.05, 1.59)
Brew and Marks, 2012	<3.27 kg Vs 3.28-3.7kg		1.95 (1.07, 3.54)
	>3.71kg Vs 3.28-3.7kg		0.91 (0.47, 1.75)
Lu et al, 2012	<3.0 kg Vs 3.0-4.0kg	1.94 (1.78, 2.11)	1.24 (1.16, 1.33)
	>4.0kg Vs 3.0-4.0kg	1.54 (1.33, 1.77)	0.93 (0.82, 1.06)
Mathew et al, 2012	<2.7kg Vs >=2.7kg	1.88 (1.08, 3.29)	1.79 (1.08, 2.98)
Mathew et al, 2012	<2.7kg Vs >=2.7kg	1.51 (0.94,2.42)	1.09 (0.54, 2.20)
Nuolivirta et al,2012	<3.0kg Vs 3.0–4.0kg	0.65 (0.13, 3.16)	
	>4.0kg Vs 3.0–4.0kg	2.95 (1.04, 8.37)	

*=Relative risk

Appendix E List of drug names and British National Formulary chapters used to confirm diagnoses of allergic conditions

Drug class names (BNF chapters)	Drug family names
Asthma	
Antimuscarinic bronchodilators (**3.1.2**)	IPRATROPIUM BROMIDE
Selective beta-2 agonists (**3.1.1**)	FORMOTEROL FUMARATE
	SALBUTAMOL
	SALMETEROL
	TERBUTALINE SULPHATE
Leukotriene receptor antagonist (**3.3**)	MONTELUKAST
	ZAFIRLUKAST
Inhaled Corticosteroids (**3.2**)	BECLOMETASONE DIPROPIONATE
	BUDESONIDE
	CICLESONIDE
	FLUTICASONE PROPIONATE
	MOMETASONE FURATE
	SODIUM CROMOGLICATE
Eczema	
Emollients (**13.2.1**)	Proprietary emollient preparations
	Non- Proprietary emollient preparations
	Preparations containing urea
	Emollient bath and shower preparations
Topical corticosteroids (**13.4**)	ALCLOMETASONE DIPROPIONATE
	BECLOMETASONE DIPROPIONATE
	BETAMETHASONE ESTERS
	CLOBETASOL PROPIONATE
	CLOBETASONE BUTYRATE
	DIFLUCORTOLONE VALERATE
	FLUDROXYCORTIDE
	FLUOCINOLONE ACETONIDE
	FLUOCINONIDE
	FLUOCORTOLONE
	FLUTICASONE PROPIONATE
	HYDROCORTISONE
	HYDROCORTISONE BUTYRATE

Drug class names (BNF chapters)	Drug family names
	MOMETASONE FUROATE
	TRIAMCINOLONE ACETONIDE
Ichthammol (**13.51**)	ICHTHAMMOL
Rhinitis	
Nasal corticosteroids (**12.2.1**)	AZELASTINE HYDROCHLORIDE
	BECLOMETHASONE DIPROPIONATE
	BETAMETHASONE SODIUM PHOSPHATE
	BUDESONIDE
	FLUNISOLIDE
	FLUTICASONE PROPIONATE
	MOMETASONE FUROATE
	SODIUM CROMOGLICATE
	TRIAMCINOLONE ACETONIDE
Antihistamines (**3.4.1**)	ACRIVASTINE
	ALIMEMAZINE TARTRATE
	BILASTINE
	CETIRIZINE HYDROCHLORIDE
	CHLORPHENAMINE MALEATE
	DESLORATADINE
	FEXOFENADINE HYDROCHLORIDE
	HYDROXYZINE HYDROCHLORIDE
	KETOTIFEN
	LEVOCETIRIZINE HYDROCHLORIDE
	LORATADINE
	MIZOLASTINE
	PROMETHAZINE HYDROCHLORIDE
	RUPATADINE
Nasal decongestants (**12.2.2**)	EPHEDRINE HYDROCHLORIDE
	XYLOMETAZOLINE

BNF= British National formulary

Appendix F List of disease terms and Read Codes used to confirm diagnosis of wheezing disorders

Name	List of terms	Read Code	Term ID
Wheezing			
	Expiratory polyphonic wheeze	Xa83N	YaVc1
	Expiratory wheeze	Xa7uu	YaVQZ
	Expiratory wheezing	Xa7vA	YaVQt
	Inspiratory wheeze	Xa7ut	YaVQY
	Inspiratory wheezing	Xa7v9	YaVQs
	Mild wheeze	XaX5K	Yaty9
	Moderate wheeze	XaX5L	YatyA
	Nocturnal wheeze/cough	173B.	YM1gs
	Severe wheeze	XaX5M	YatyC
	Very severe wheeze	XaX5N	YatyE
	Viral wheeze	XaMe7	YapfP
	Wheeze - rhonchi	X76If	Y7DxZ
	Wheezing	XE0qs	Y7DuF
	Wheezing symptom	XM0Ci	YM1is
	Wheezy	XE0qs	Y7DuF
Asthma			
	Acute asthma	Xa9zf	YaYk2
	Allergic asthma	XE0YT	Y108G
	Asthma	H33..	Y107p
	Asthma NOS	XE0YX	Y1080
	Asthma unspecified	H33z.	Y107y
	Asthmatic bronchitis	Xa0lZ	Y108e
	Brittle asthma	Ua1AX	YMFVN
	Childhood asthma	X101t	Y107w
	Chronic asthmatic bronchitis	H3120	Y108g
	Mild asthma	663V1	YaY1o
	Moderate asthma	663V2	YaY1p
	Nocturnal asthma	XaLPE	Y1084
	Non-allergic asthma	XE0YT	Y108G
	Occasional asthma	663V0	YaY1n

Appendix G Model fit statistics results of FIML and MICE missing data methods

Fit criterion	FIML		MICE	
	Overall model	**Multi-group model**	**Overall model**	**Multi-group model**
AIC	11409	11319	16635	16555
BIC	11517	11535	16748	16781
RMSEA	0.106	0.105	0.180	0.187
CFI	0.919	0.924	0.931	0.927
TLI	0.901	0.907	0.912	0.907
SRMR	0.040	0.043	0.045	0.048
-2LL	11,367	11,234	16,590	16,466