Defining optimal infant growth for lifetime health:
a systematic review of lay and scientific literature

009/0026

**Defining optimal infant growth for lifetime health:
a systematic review of lay and scientific literature**

Contents

Review of lay perspectives and focus groups

Appendices

Appendix 1   Composition of advisory group
Appendix 2   Definition of terms

Review of life course outcomes
Appendix 3   Framework of health outcomes
Appendix 4   Search strategies
Appendix 5   Summary of grouping of outcomes for searches and numbers of abstracts identified
Appendix 6   List of experts and first authors contacted
Appendix 7   Summary of quality scoring assessment system
Appendix 8   Log of protocol changes
Appendix 9   Excluded studies (with reasons)

Appendix 10   Synthesis and quality tables

Review of lay perspectives
Appendix 11   Included study descriptions and quality assessment
Appendix 12   Search strategies
Appendix 13   Example data extraction
Appendix 14   Poster displayed in community groups visited
Appendix 15   Information for participants and consent form

Executive summary

Size and growth matter to parents in infancy and beyond. More recently, early size and growth have become a concern to policymakers, who fear that increasing levels of childhood obesity will lead to an explosion of chronic disease in adulthood.

Research evidence suggests that infant growth is linked to health in later life. However it is not clear what particular pattern of growth will be associated with good health at different ages. In order to inform policy the evidence relating to infant growth to health across the life course needs to be collated.

The aim of this project was to carry out a systematic collation of evidence to address the question: how should an infant grow for life long health?

There were two components to the project:
- a systematic review to assess the relationship of infant growth status to health and well-being across the life course
- a systematic review of lay perspectives on infant growth status combined with focus groups exploring the views of parents

The objective of the review of life course outcomes was to address the question:
- What is optimal infant growth for early and later health?

The objectives of the review of lay perspectives and focus groups were to understand:
- lay, particularly parental perspectives on infant growth
- issues that are seen as important in relation to infant growth and whether infant growth is a salient issue for these groups
- where infant growth lies among priorities for those responsible for creating and maintaining child health

Methods

Review of life course outcomes
The association of infant growth status with health and wellbeing across the life course was studied through a series of systematic reviews. Outcomes for the review were important causes of burden of disease, significant health-related behaviours and determinants of non-health related quality of life selected at all stages of the life course from infancy to adulthood.

This review was carried out according to the methods of the Centre for Reviews and Dissemination , University of York. A search of 8 electronic databases was carried out leading to identification of over 160,000 abstracts and 279 papers were assessed for inclusion by two reviewers

A mainly narrative approach to synthesis was taken but, where possible, meta-analysis was carried out according to standard procedures.

Review of lay perspectives
A search of 12 electronic databases was undertaken leading to identification of 2,694 abstracts.   Of these, 76 papers were assessed for inclusion by two reviewers.

Study findings were synthesised using both narrative and thematic approaches. The quality of studies was assessed using criteria based on existing guidelines for review of qualitative research but adapted in an iterative process for the study types included in this review.

The second part of the work was addressed through identifying individuals in childcare facilities in East London who met a range of inclusion criteria. On the basis of these, five focus groups were carried out with parents, and other family members and friends who supported them.

Results

Review of life course outcomes
119 papers based on108 studies met inclusion criteria for the review. Nearly all studies were observational and the majority were cohort studies. The quality of studies was generally not high with most being assessed as having medium risk of bias. Areas where study quality was most often poor were consideration of the effects of confounding factors and rates of follow-up.

Findings in infancy and childhood
Over half of the studies identified related to outcomes in infancy and childhood but the majority of these related to childhood development and obesity with very few studies relating to diseases of childhood. Large gaps in the literature were identified.

The studies of childhood obesity showed that the risk of obesity was increased in children who had been in the highest end of the size distribution for weight or body mass index or who had gained weight rapidly during infancy.

The studies of development revealed that:
- Larger size in infancy (head circumference, weight and length) was associated with better cognitive development
- Growth faltering was associated with poorer cognitive development
- Larger head size in infancy was associated with better motor development

Findings in adolescence
Only 8 studies relating to adolescent health and wellbeing were identified. One of these related to mental illness and showed that there was no association between growth faltering in infancy and anxiety and self-esteem in adolescence. Seven studies related to insulin dependent diabetes and showed that larger size in infancy was associated with greater risk of diabetes.

Findings in adulthood
We identified 26 studies relating to adult outcomes, of which 5 related to ischaemic heart disease and 12 to adult obesity. Other important adult outcomes such as cancer, mental illness, stroke, dementia and non-insulin dependent diabetes had little or no evidence associating them with infant growth.

The main findings were that:
- smaller size in infancy was associated with increased risk of ischaemic heart disease
- there was very little evidence in relation to stroke, non-insulin dependent diabetes and suicide but the few studies that were identified suggested that smaller size in infancy was associated with increased risk of each outcome
- larger size (weight or body mass index) during infancy, was associated with adult obesity

There was virtually no evidence relating to measures of health-related behaviours such as smoking, or non-health related quality of life such as employment or education.

Lay perspectives
Nineteen studies met the inclusion criteria for the review of lay perspectives and these related to the views of a total of 3,590 individuals across studies. The majority of studies (16) related to the views of mothers (1948). The remainder considered the views of health professionals, children and other adults. Infants in these studies ranged from preterm neonates to 2 years of age. All studies were based in developed countries

Included studies were generally of poor quality for the purposes of this review particularly with respect to study design and sampling. The views of family members other than mothers were under-represented in the studies reviewed.

The findings of the syntheses suggest that:
- monitoring the size and growth of infants was common practice and something that mothers considered important, particularly for their first child
- in contrast, reported concerns about overweight and underweight, in themselves, were small
- there was anxiety about the causes and consequences of size, for example mothers of infants with faltering growth wanted to explore the medical causes of this
- understanding of healthy infant size or growth was dominated by the concept of normality
- norms were described by referring to growth charts; clothes were used to judge appropriateness of size and monitor growth, as was comparing between children and comparing between family members
- where growth differed from the norm and a plausible explanation could not be found, growth became an important concern to parents
- a wider range of infant sizes was acceptable than for other stages in childhood and evidence from the USA suggested that parents did not consider overweight to be a problem
- the views and actions of parents were influenced by their families, other parents, health professionals and the media

The findings of the focus groups accord with the review of studies in that participants referred to the importance of normality in infant size and most saw growth as an indicator of underlying health rather than a problem in itself.

In addition:
- reactions to the possibility of intervention to alter growth were mixed. While all participants thought that intervening might be appropriate in principle, all struggled with defining an occasion when it would be acceptable in practice, particularly in infants

Conclusions

Though some beneficial health outcomes (in childhood and adulthood) and developmental milestones in childhood are associated with larger size in infancy, both insulin dependent diabetes and obesity in childhood and adulthood are also associated with larger infant size. There are many gaps in the literature particularly in relation to important disease groups such as cancer, asthma and mental illness.

The literature on lay perspectives suggested that understanding of what constituted healthy size or growth was dominated by concepts of normality. Parents were concerned that the size or growth of their baby was 'normal', but viewed these as indicators of infant health rather than of importance in itself. Findings from focus groups suggested that parents would not be in favour of intervening in an infant's growth without strong evidence of benefit. Any intervention proposed would also need to draw on parents' understandings of the causes and consequences of unhealthy growth.

The evidence is not sufficient to propose the development of an intervention to alter infant growth with a view to improving adult health.

Defining optimal infant growth for lifetime health:
a systematic review of lay and scientific literature

1. Introduction

1.1 Background

Size and growth matter to parents in infancy and beyond. More recently, early size and growth have become of concern to policymakers, who fear that the trends to increasing fatness in childhood [1;2] will lead to an explosion of chronic obesity-related disease in adulthood, such as cardiovascular disease and diabetes mellitus[3].

The research evidence linking early growth life with adult disease poses a dilemma. On the one hand, observational evidence of people born 60 to100 years ago links slower rates of infant growth to increased rates of death from coronary heart disease (CHD) in later life[4;5]. This has led some to suggest interventions to promote infant growth for the primary prevention of CHD[6].   On the other hand, research in younger age groups has suggested that rapid growth in infancy may also be associated with increased risk of later obesity, hypertension and insulin resistance, arguing against interventions to promote infant growth for later health [7]. Taken together, these strands of evidence raise an important question – how should an infant grow to achieve life-long good health?

Though considerable evidence exists to address this question, it has been difficult to use in policymaking. Firstly, there has been no systematic collation of the evidence to assess whether optimal patterns of infant growth can be defined. Such an assessment is an essential pre-requisite to designing interventions or policies to alter infant growth[8]. Secondly, public opinion on this topic is unknown but crucial, since people's values will affect both the acceptability and implementation of any policy or intervention. The benefits of policies and interventions aimed at optimising infant growth to prevent adult disease will not be realised for many years. However parents and others concerned with the welfare of infants will want to know the likely effects of interventions at earlier stages of their lives. They will also be concerned about the possibility of adverse effects of interventions particularly if they precede the benefits. Thus development of policy and interventions to optimize infant growth needs to be based on a collation of research evidence which relates to all stages of the life course and all dimensions of health.

Interpretation of research relating infant growth to later health is challenging. Growth is one of a number of inter-related social and biological factors that influence health. The growth that has occurred in utero will be a strong influence on the growth trajectory in early infancy. Babies who are large at birth are likely to have a slower rate of growth in early infancy and, conversely, babies who are small at birth are likely to grow more quickly. Infant feeding also influences growth with breast and bottle fed babies having different growth trajectories during infancy. Similarly socio-economic status and factors linked to it, such as maternal education, will be potential confounding factors. Any collation of evidence must take account of the role of factors that might confound or modify the relationship between infant growth and later health.

A further challenge relates to the measurement of infant growth. Since growth is change in size over time, longitudinal studies with serial measurements of size are the most accurate way of assessing growth. However such studies are scarce and much published research focuses on infant size rather than growth.

In this project we have addressed the question – "how should an infant grow to achieve life-long good health" - in a way which we have aimed to make relevant to policymaking, and to take into account the issues just described. We have assessed scientific evidence and lay perspectives on infant growth to assess whether patterns of optimal infant growth can be defined, and if so whether the development of an intervention to optimise infant growth is justified scientifically and is feasible, desirable and ethical. This was achieved through a systematic review of the associations of infant growth with health across the life course, integrated with a systematic review and qualitative evaluation of lay perspectives on infant growth.

1.2 Aims of the project

While we set out to consider the evidence relating to infant growth, we recognised that many studies would consider infant size alone. We took account of this in our aims:

The aims of the project were:

- To systematically assess and collate the scientific evidence on infant size or growth with a view to determining whether optimal patterns of infant growth can be defined (NB it was not assumed that a single pattern of infant growth will be found to be 'optimal')
- To systematically assess and collate evidence on lay views of infant size or growth with a view to understanding attitudes towards desirable infant size and growth

- To identify gaps in the evidence, to assess the implications of those gaps and make recommendations for addressing the gaps

If optimal patterns of infant growth could be defined as a result of this work then the next phase of our research would be directed towards the development of a testable intervention to prevent adult disease by optimising infant size or growth **and/or** to recommend a policy change to achieve this.

1.3 Themes for the review

A number of themes were considered throughout the review:
- Consideration of possible beneficial and adverse outcomes in relation to early growth
- Consideration of inequalities in people's lives, and how these influence the relationships of infant growth to health
- The balance between early and late effects of infant growth on health
- Consideration with lay people of interventions to optimise infant growth that would be acceptable to them
- The capacity of the evidence to inform public policy at local (e.g. local authority,), national (e.g. Department of Health, Food Standards Agency) or international (e.g. international aid agencies) level.

In order to inform policy it was important that we considered the balance between any beneficial or harmful effects of altering infant growth. Likewise, although the starting point for the review was the relation of early growth to adult disease, we considered it important to assess the balance between early and late effects since any early effects would be influential in determining whether interventions were ethical and whether they were acceptable to parents and young people. There are inequalities both in infant growth and in many of the outcomes that we were considering. We therefore intended to consider the influence of inequalities on the relationships between infant size or growth and each outcome and to assess potential implications of our findings for addressing inequalities.

The methods and results are divided into two sections:
- The review of life course outcomes - methods are described in chapter 2 and the results in chapters 3-5
- The review of lay perspectives - methods are described in chapter 6 and 10 and the results in chapters 6,8 and 11.

Review of life course outcomes

2. Methods

2.1 Objectives
The overall question addressed by the review was:
<u>What is optimal infant size or growth for early and later health?</u>
In addressing this objective a number of questions were set out with the intention that each should be addressed through a series of systematic reviews.

- What is the relation of infant size/growth to adult disease (including outcomes of mortality, longevity and health-related quality of life)?
- What is the relation of infant size/growth to wellbeing (including outcomes of non-health related quality of life) in adulthood?
- What is the relation of infant size/growth to health in early life (including outcomes of mortality and health-related quality of life)?
- What is the relation of infant size/growth to wellbeing (including outcomes of non-health related quality of life) and developmental status in early life?
- What is the size of effect of these relations between infant size/growth and earlier and later health?
- Can the information obtained to answer these questions be used to define "optimal" infant size/growth?

We did not assume causality in the relationships we examined. As outlined in section 1.1, we recognised that growth is one of a number of inter-related factors that are associated with health. Infant growth may be an indicator of other factors rather than causing associations in its own right. We took account of this through consideration of the effects of potential confounding factors and effect modifiers in the relationship of infant growth with each of the review outcomes as outlined in section 2.2.5.

2.2 Methods

The review was carried out according to the structure and methods developed by the Centre for Reviews and Dissemination, University of York[9].

2.2.1   Study inclusion criteria

Studies were selected based on criteria relating to study sample, infant growth status, outcomes and study design. These are detailed below.

Sample

Though infancy is usually defined as the first year of life, we included studies of children from birth to 2 years of age, as most of the variability in early growth is contained in the first 2 years, and selection of a specific time point within this period (e.g. 12 or 18 months) would have been arbitrary. We did not select studies on the basis of characteristics at birth (the beginning of infancy). Thus studies of low birthweight and premature infants were included in the review.

Infant growth status (independent variable)

We recognised that data reported in many studies was likely to be of insufficient quality or completeness to adequately define growth in terms of change in size over time. We therefore aimed to examine the literature relating to infant size as well as infant growth.

Studies were selected if they reported the relationship between infant size or infant growth and one of the review outcomes. Studies examining infant size were eligible for inclusion if they reported at least a single measurement of infant size between 3 months and 2 years.

Studies that examined infant growth were included if they reported at least 2 measurements of size, of which one was between 3 months and 2 years. Studies that examined any measure of infant size including weight, height and components of height such as leg length and sitting height, head circumference, measures of weight for height (ponderal index, body mass index) and other measures of fatness including skinfolds were eligible for inclusion. Studies examining abnormal patterns of infant growth such as stunting or growth faltering (failure to thrive) were also considered for inclusion in the review.

Outcomes

Since the review was intended to inform public health action we selected a range of outcomes that either pose a significant burden of disease or have a significant impact on health or quality of life that we believed would be important to parents and others concerned with the welfare of infants.

Outcomes were selected across all stages of the life course: infancy, childhood, adolescence and adulthood in the following categories:

- **Determinants of mortality/longevity:** diseases and conditions that account for a significant proportion of the overall mortality within a particular age group.

- **Health related quality of life:** diseases and conditions that have a significant impact on health-related quality of life. This category also included composite measures of quality of life such as the SF 36 and the 'health utilities index'.
- **Non-health related quality of life:** social and economic determinants of health in the categories 'emotional well-being', 'economic status', 'material environment' and 'education'.
- **Significant health-related behaviours:** this category was included in order to assess whether interventions that change patterns of infant or childhood growth have positive or negative effects on health-related behaviours such as smoking or alcohol intake.
- **Development:** this category, which applied to early well-being only, included measures of social, emotional, cognitive, sensory and motor development.

We recognised that some of the relationships we were going to explore based on this framework did not appear biologically plausible, for example the relationship between infant growth and employment. However biologically implausible relationships have been demonstrated in the past (for instance between lipid lowering drugs and violent death) and the public may be concerned about such associations.

Selection of outcomes

This section outlines the methods used to select outcomes for the review. The selected outcomes, listed in appendix 3, are grouped according to age group (infancy 0-2 years, childhood 1-9 years, adolescence 10-17 years and adulthood 18 years and over).

Mortality/longevity

Adults

The Global Burden of Disease Study (GBDS), carried out by the World Health Organisation (WHO) and the World Bank in 1997, provided an appropriate source for selection of adult mortality outcomes as well as morbidity outcomes as outline in the section on health-related quality of life. The GBDS reports ranked leading causes of death and disability in 1990 across different regions of the world and also reported projections for the year 2020[10]. The primary indicator used to summarise the burden of death and disability in GBDS is the disability-adjusted life year (DALY). This is the sum of life years lost due to premature mortality and years lived with disability, adjusted for severity.

We used the 2020 projections since this takes into account the fact that the target population of any intervention designed as a result of this review are currently children. Our outcomes were based on major causes of DALYs in developed countries, reflecting the fact that the review intended to focus on interventions that could be used in developed country settings. Some of the major causes of DALYs in developing countries, such as infectious disease and war injuries, would be of less relevance to the developed world. However, the findings of the review would still have some relevance to developing countries since developed and developing countries share a number of major causes of DALYs (including unipolar major depression, road traffic accidents,

ischaemic heart disease, chronic obstructive pulmonary disease and cerebrovascular disease).

Outcomes were selected if they represented one of the 10 leading causes of DALYs for developed countries. The 12 outcomes listed in the adult mortality section of appendix 3, represent the top 10 causes of DALYS combined for men and women.

<u>Children and adolescents</u>
The GBDS is inappropriate for selection of childhood and adolescent outcomes. The methods used to select diseases for inclusion in GBDS and the weighting given to years lived in disability, result in the major causes of DALYS in developed countries being largely chronic conditions that cause a greater burden of disease in adults. Having used GBDS in adults we looked for equivalent quantitative data for earlier stages of the life course.

The outcomes listed in appendix 3 represent the 10 leading causes of mortality in children and adolescents in the United States taken from statistics reported by the Centre for Disease Control and Prevention (CDC). These were selected in preference to mortality statistics reported by the Office of National Statistics (ONS for England and Wales) and WHO because CDC report leading causes of death using a well-validated technique based on rankable cause lists. In contrast, ONS and WHO report leading International Classification of Disease (ICD) chapters rather than actual causes of death. However, comparing ONS and WHO mortality statistics with CDC causes, there appeared to be reasonable agreement as to the leading causes of death in different developed countries. Outcomes in children aged 1-9 years were based on leading causes of death reported in the 1-4 and 5-9 year age groups by CDC. The corresponding percentage of deaths attributable to each outcome in each age group is shown in brackets in table 1. The age groups differed with respect to one leading cause of death and so 11 outcomes are listed to take account of this. This also applied to outcomes in the 10-17 year age group which were based on causes of death in two CDC age groups (10-14 and 15-19 years) and also differed with respect to one cause of death.

Although they are important causes of mortality, we did not consider congenital anomalies and cystic fibrosis in this review. This was because they would not be amendable to intervention during infancy given that they are present from birth and could not be predicted by patterns of infant growth or size. These outcomes are shown in brackets in appendix 3.

There were also a number of outcomes where changes in rate of growth can be caused by presence of disease before it is detected, for example respiratory infection. We included studies which assessed the relationship of infant size/growth with these outcomes, and assessed the potential for reverse causation during data extraction.

Health-related quality of life

The GBDS was also used to select adult outcomes in this category since DALYs take account of years lived with disability adjusted for severity. They therefore reflect health-related quality of life as well as important causes of mortality. Again our outcomes were the 10 leading causes of DALYS for men and women in developed countries.

We failed to identify an appropriate source of important morbidity outcomes in childhood. We therefore selected outcomes that, due to their chronic nature and prevalence in infancy, childhood and adolescence, could be considered important causes of morbidity[11]. Outcomes included asthma, insulin dependent diabetes, mental health problems, such as depression, anxiety and eating disorders, learning disability and sensory impairment.

We also aimed to consider studies that examined the relationship between infant size or growth and composite measures of health-related quality of life for inclusion in the review.

Non-health related quality of life

Selection of non-health related quality of life outcomes was based on evidence relating to the wider determinants of health[12]. Outcomes were social and economic determinants of health grouped in the following categories:

- Emotional well-being (well-being in its own right and factors influencing well-being including security at work or home, social support and control at work or home)
- Economic status (income and proxies for income (e.g. car ownership, housing tenure), employment status)
- Material environment (housing and homelessness, quality of material environment (e.g. Townsend score))
- Education (time spent in full-time education, qualifications)

Whilst there is overlap in determinants of non-health related quality of life in adults and children, given that a child's quality of life will depend on the environment provided by their carers, additional outcomes of relevance in childhood were also selected. These included 'looked after children' (or children not in parental care), 'children on the at risk register' in the UK (and equivalent in other countries), and 'exclusion from school'. Outcomes are summarised in full in appendix 3.

Significant health-related behaviours

We selected major lifestyle risk factors that are known to be important factors in determining burden of disease, as demonstrated in a recent study from GBDS which explored the contribution of lifestyle factors to the burden of disease world-wide[13]. The outcomes selected were smoking, alcohol, unhealthy eating, lack of physical activity use of illicit drugs and unsafe sex. Obesity is a proxy for both unhealthy eating and lack of physical activity and so we planned to look for studies relating to obesity as well as to unhealthy eating and lack of physical activity in their own right.

Although known to be most relevant to adults and adolescents, unhealthy eating, lack of physical activity and, to an extent, smoking and substance misuse were also relevant at earlier stages of the life course.

Childhood development
Outcomes in this category included aspects of social, emotional, cognitive, sensory and motor development.

Study design
It was our intention to include studies (in any language) which reported data on individuals. Ecological studies, and non-human studies would therefore be excluded. Study designs that were eligible for inclusion were:

- Cohort (prospective or retrospective)
- Case-control
- Cross-sectional
- Experimental studies (clinical or other controlled trials, before and after studies)

We also intended to consider systematic reviews of the relationship between infant size/growth and later outcomes for inclusion in the review. However, given that systematic reviews are secondary research the approach to reviewing them differed from that for primary research studies. If quality was not sufficient to warrant their inclusion in the review, then our intention was to use them as a source of primary studies.

Although the focus of the review was on the potential for development of interventions in developed countries, there is overlap with developing countries relating to chronic disease outcomes as outlined previously. Therefore studies carried out in developing countries were considered for inclusion in the review if they related infant growth to any of the outcomes listed in appendix 3.

Study settings and timing
We aimed to include all studies which contributed relevant information, regardless of the setting. However, we aimed to note the setting as part of data extraction and for use in narrative synthesis. We did not plan to exclude studies describing the relation of early growth to later disease/health on the basis of publication date. However, given the changing trends in childhood growth, we considered setting a cut off date (perhaps 1960) for studies describing early growth and early disease/health if the literature scoping suggested that the pre-1960s literature was large and perhaps not feasible to include this within the timescales of the review. However scoping suggested that the number of studies this applied to would be small, and so no cut-off date was set.

2.2.2 Literature search strategy

Preliminary searches
Given the large number of outcomes included in the review, some preliminary searches were carried out in order to assess the potential size of the literature.

A search strategy was developed for each outcome, and searches were carried out using Medline from 1966 to 2003. The databases of abstracts were then screened by one of two reviewers. Estimates for the number of studies relating to each outcome likely to be included in the review were calculated based on this work.

Electronic database searches
The full search involved searching a range of electronic databases. Medline was searched for all outcomes but additional databases differed according to the outcomes being considered and were selected because of their relevance to that particular outcome. For example, searches for studies relating to childhood development included ERIC, CINAHL, Sociological abstracts and PsycInfo.

Each database was searched from its starting date until January to June 2003 (due to the number of outcomes, searches were carried out over a 6-month period). Full details of the databases searched and search strategies are given in appendix 4.

Given the large number of outcomes being considered, we grouped outcomes when appropriate. So, for example, the searches for lung cancer, breast cancer and childhood cancer were all carried out as one large cancer search. Although some outcomes were specific to particular stages of the life course, we avoided using terms to specify stages of life course instead applying this criterion when screening abstracts. So, for example, our searches for mental illness in adolescence and adulthood were carried out as one search.

The bibliographies of all included studies were hand-searched. All searches were updated between April and May 2004.

We also intended to search foreign language databases such as LILACS. However, due to the large scale of the searching task and the time constraints of the project it was only possible to carry this out for the outcomes relating to childhood development as reported on page 36, section 3.3.

Contact with first authors and experts
All first authors of studies that met the review inclusion criteria were contacted by letter and by email in order to identify any further completed analyses (published or unpublished) that may be of relevance to the review. In order to achieve this each author was sent the complete list of included studies together with details of the outcomes being considered in the review and review inclusion criteria. The same approach was used to contact a number of experts. A full list of experts and authors contacted is given in appendix 5.

Just over a third (34%) of first authors replied to our request and 72% of these made suggestions regarding studies or authors that might be relevant. As a result of this process 3 unpublished papers were added to the review. However most of the other studies suggested did not meet inclusion critieria usually because they related to outcomes outside our framework or because they related to growth in childhood rather than infancy. Nearly all experts who

were contacted replied, although only one suggested any studies that had not already been identified.

2.2.3 Assessment of papers for inclusion

Appendix 6 summarises the numbers of abstracts identified by each search. Two reviewers independently assessed each title and abstract for potential relevance to the review. Papers thought to be relevant were obtained and assessed in detail against review inclusion criteria. Disagreements over inclusion were resolved through consensus and, where necessary, through discussion with a third member of the review project group.

During the review work we kept a record of interpretation of inclusion criteria and these are reported in appendix 7 together with a log of protocol changes. A list of studies that were considered but subsequently excluded together with reasons for exclusion is given in appendix 8.

2.2.4 Data extraction

Data extraction was carried out independently by 2 reviewers (JB and DF). Data were extracted into a Microsoft Access database. Two versions of an electronic data extraction form were developed; one for case-control studies and the other for cohort and other study designs.

2.2.5 Assessment of study quality

Study quality was assessed using a checklist based on CRD guidelines and the validity checklist used in 'A Systematic Review of Water Fluoridation (2000)'[14], but developed in an iterative process of extensive piloting and consultation to meet the needs of our own review. Two separate checklists were devised, one for studies using a case-control design and one for any other study designs (mostly cohort). A number of aspects of quality (11 for cohort and14 for case-control designs) were assessed according to whether they posed a low, medium or high risk of bias for a study's results and conclusions. Study quality was assessed independently by two reviewers (JB and DF), with disagreements resolved through consensus. The checklists used are summarised in appendix 9.

An important aspect of quality assessment was to determine whether studies had adequately controlled for variables that could confound or modify the relationship between infant size or growth and later outcomes. These factors varied according to the outcome being considered and the stage of the life course. During infancy, infant feeding, socio-economic status, maternal education and ethnicity were considered. The same factors were considered in childhood. During adolescence and adulthood we also considered current size, puberty and the effects of lifestyle factors including smoking, alcohol intake and physical activity. Additional factors were considered according to the outcome being reviewed and these are listed in 'quality assessment' for each section of the results.

Overall judgment on a study was also given in terms of low, medium or high risk of bias, based on a combination of performance in the checklist and by consensus by two independent reviewers. This approach enabled the

reviewers to highlight any particular problems relating to study validity that were not adequately addressed by the checklist.

2.2.6 Data synthesis
Our aim was to investigate whether effects were consistent across studies and to explore reasons for apparent differences. Given that we suspected that most of the included studies would be observational, it was necessary to use a mainly narrative approach to synthesis. Where possible, we carried out meta-analysis according to standard procedures[9].

Results
Chapters 3,4 and 5 describe the findings of the review of life course outcomes. Chapter 3 relates to the findings in infancy and childhood, chapter 4 to adolescence and chapter 5 to adulthood.

No relevant studies were identified for many of the outcomes considered. Table 1 describes the number of studies identified in relation to each outcome where we did identify studies and summarises the year of birth of subjects, and the geographical location of studies. It also gives information on the numbers of studies that considered infant size or growth.

Table 1: Summary of studies identified in relation to each outcome

STAGE OF LIFE COURSE	OUTCOME	NUMBER OF STUDIES	YEAR OF BIRTH OF SUBJECTS (RANGE)	GEOGRAPHICAL LOCATION (DEVELOPED/ DEVELOPING)	SIZE/ GROWTH*
Infancy and childhood	SIDS	8	1956-96	8/0	2/6
	Cancer	1	1984-96	0/1	0/1
	Respiratory and diarrhoeal infection	3	1982-91	0/3	1/2
	Autism	1		1/0	1/1
	Obesity	9	1950-90	8/1	5/4
	Cognitive development	44	1959-90	33/11	29/17
	Motor development	10	1977-87	7/3	8/2
Adolescence	Insulin-dependent diabetes	7	1960-90	6/1	6/1
	Mental illness	1	1987-88	1/0	0/1
Adulthood	Iscaemic heart disease	5	1911-44	5/0	5/0
	Cerebrovascular disease	1	1911-30	1/0	1/0
	Non-insulin dependent diabetes	3	1920-72	3/0	3/0
	Cancer	3	1911-46	3/0	3/1
	Osteoarthritis	1	1946	1/0	1/0
	Mental illness	2	1911-30	2/0	2/1
	Obesity	12	1927-82	11/1	9/2
	Income and educational status	1	1933-44	1/0	1/1

*Some studies considered both size and growth and so totals in this column may differ from 'number of studies'.

3. Results relating to outcomes in infancy and childhood

3.1 Mortality and health-related quality of life
Of the 16 possible outcomes or groups of outcomes relating to mortality and morbidity in infancy and childhood (appendix 3), searches identified eligible studies in relation to only 4. These were sudden infant death syndrome (SIDS), cancer, respiratory and diarrhoeal infection and autism. No studies were identified for conditions responsible for a large burden of disease in childhood such as asthma or accidents.

3.1A Sudden infant death syndrome
Tables in appendix 10 relating to SIDS are table 3.1A (i) and (ii).

3.1A.1 Selection of studies
Electronic searches identified 1714 references. Screening of titles and abstracts led to identification of 14 studies considered to potentially meet review inclusion criteria. Seven of these were excluded once papers were obtained and assessed in more detail. Seven studies met review inclusion criteria. Screening of reference lists of these papers led to identification of a further 4 studies that were assessed in detail. One of these met inclusion criteria for the review. The 8 included studies are summarised in Table 2 below.

Table 2: Summary of included studies in review of infant size or growth and SIDS

Ref. no	Title	Author (year)	Design
15	Growth velocity before sudden infant death	Williams (1990)	Case-control
16	Postnatal growth preceding sudden infant death syndrome	Brooks (1994)	Case-control
17	Growth and the sudden infant death syndrome	Williams (1996)	Case-control
18	Postnatal growth and the sudden infant death syndrome	Peterson (1974)	Case-control
19	Sudden infant death in Copenhagen 1956-1971. IV. Infant development	Jorgensen (1982)	Case-control
20	Weight gain and sudden infant death syndrome: changes in weight z scores may identify infants at increased risk	Blair (2000)	Case-control
21	Sudden infant death syndrome. A prospective study	Naeye (1976)	Case-control
22	Epidemiology of sudden unexpected death in infants ('cot death') in Northern Ireland	Froggatt (1971)	Case-control

3.1A.2 Description of studies
All 8 studies were case-control studies, 3 set in the UK, 2 in New Zealand, 2 in the USA and 1 in Denmark.

3.1A.3 Assessment of infant size/growth
Six of the studies examined the relationship between infant growth and SIDS. The 2 studies by Williams, and the studies by Brooks and Petersen compared growth velocity in cases with that in controls at different time intervals leading

up to death[15-18]. The later Williams study also compared infant size (weight) in cases and controls[17]. In the study by Blair et al size measurements were converted to z scores and odds ratios per unit change in z score were calculated at different time intervals[20]. The remaining studies by Naeye et al and Froggatt et al considered infant size only, the former assessed weight at 4 months and the latter size at death[21;22]. The study by Petersen et al looked at growth velocity in weight, length and head circumference[18]. This was the only study to explore measures of growth other than weight. However these measurements of growth were based on post mortem data which were used to estimate growth up to death.

3.1A.4 Quality assessment

The design of 5 of the studies was considered appropriate since it was based on all SIDS cases arising in a given population over a period of time compared with controls from the same population[15-17;19;20]. However the design and setting of 3 of the studies was considered a potential source of bias. The US study by Naeye was hospital based with cases drawn from 14 medical institutions taking part in a perinatal study over 7 years with controls drawn from the same perinatal study[21]. The study by Naeye is also biased since it was based on only 22% of potential cases due to lack of infant size data in the remainder who had all died before 4 months of age. The 1974 study by Petersen et al, whilst population-based in relation to ascertainment of cases, selected controls from 2 different sources one of which was a clinic serving disadvantaged infants and the other a private clinic serving infants from more privileged backgrounds[18]. The 1971 study by Froggatt et al compared a population-based sample of SIDS cases with standard populations series when exploring the influence of infant growth, instead of the control group for the study[22]. For these reasons, control selection was considered a high source of bias in the studies by Froggatt and Peterson. Nearly all of the other studies used an appropriate method of control selection either though random selection of controls or through a systematic approach which enabled matching of potential confounding factors.

Infant size or growth was measured by health visitors, often in subjects' homes, in the majority of studies. Data on size or growth were ascertained from child health records in most cases. None of the studies gave information as to whether observers abstracting the exposure data were blind to outcome status. Only one of the studies used any rounding of infant size measurements: Williams et al used a smoothing spline technique to generate a weight for each 2-week period of the study[15]. Two of the studies used size at death as their measure of infant size and compared these to living weights in controls[18;22]. These were considered to have a high risk of bias since measurements at death are known to differ from live measurements due to loss of water and other factors.

SIDS outcome was ascertained using validated techniques, involving independent assessment of autopsy and medical records by a multi-disciplinary committee in 5 of the studies[15-17;19;20]. In the other 3 studies, however, there was no independent validation of the diagnosis of SIDS other

than that assigned by a single investigator and so we considered these studies to have potential for bias in this respect[18;21;22].

The confounding factors we considered important in the relationship between infant growth and SIDS were infant feeding, gender, socio-economic status and parental smoking.   In assessing quality, we considered matching of cases and controls in relation to confounding factors together with adjustment in statistical analyses to be appropriate means of controlling for confounding factors. Two of the UK studies adequately considered the majority of the confounding factors in their analyses. However the studies by Naeye, Petersen and Froggatt, all of which were published in the 1970s, were poor in their consideration of confounding and were rated as having a high risk of bias in this respect. The 3 remaining studies were assessed as having medium risk of bias. The first study by Williams et al matched cases and controls on sex and infant feeding and used a matched pairs analysis, but did not consider socio-economic factors or parental smoking[15]. The study by Jorgensen matched analyses in relation to sex and age but did not state whether a matched analysis was used. This study also considered infant feeding but not socio-economic status or parental smoking[19]. The second study by Williams et al stratified analyses by sex and also carried out randomisation tests for maternal smoking and infant feeding but did not adjust for these in analyses[17].

3.1A.5 Summary of findings
There was a general lack of consistency in the findings of the 8 studies reviewed. Four of the studies showed that at least one element of growth or size in SIDS victims was reduced in comparison with controls[15;17;18;20]. However only the study Blair et al, which was the largest of the studies, including 325 cases and 1292 controls was consistent across all the intervals of growth considered[20]. The other 3 studies showed that one or more aspect of growth was reduced but demonstrated no association for others.

One approach to synthesis of findings was to group studies according to their risk of bias. Looking at the 4 studies assessed as having low risk of bias, 2 of the studies did suggest a negative association between infant growth and SIDS. As stated above, Blair et al demonstrated poorer growth in SIDS cases at three different intervals up to the last live measurement before death[20]. Overall the odds ratio of SIDS between birth and last live observation within 2 weeks of death was 1.89 (95% confidence intervals (CI) 1.15, 3.08 $p<0.01$) after adjusting for all potential confounders. The 1996 study by Williams et al showed significant differences in infant size between cases and controls with smaller size in cases at 3 months, 6 months and at penultimate observation before death in both sexes except for females at 6 months. The studies by Brooks et al and Jorgensen et al failed to show an association between infant growth and SIDS[16;19]. However, both of these studies were based on relatively smaller numbers of SIDS cases: 78 in the study by Brooks and 131 in the Jorgensen study compared with 325 and 309 in the Blair and Williams 1996 studies respectively. It is therefore possible then that the fact that they were underpowered might partly explain why these 2 studies failed to demonstrate a statistically significant association.

The other 4 studies, with medium risk of bias, were all inconclusive in their findings, although the studies by Williams and Petersen did show a significant inverse association for one particular interval of growth[15;18]. Williams' 1990 study showed significantly lower growth rates between the last observation and death in cases compared with controls. Petersen showed significantly lower growth rates in the SIDS cases than the advantaged controls but these findings were not replicated in the comparison with the disadvantaged controls.

These results, grouped by risk of bias, are shown in Table 3.

Table 3: Summary of SIDS findings, grouped by risk of bias

Direction of relationship with infant size/growth	No of studies (ref. nos. in parentheses)	
	Low risk of bias	Medium risk of bias
Inverse	2 [17;20]	0
None	2 [16;19]	0
Positive	0	0
Inconclusive	0	4 [15;18;21;22]
Total	4	4

Although overall the findings of these studies were inconclusive, 2 studies both with low risk of bias suggested that SIDS was associated with poorer growth during infancy.

3.1B   Cancer

3.1B.1   Selection of studies
Searches of Medline and Embase identified 12,678 references. Screening of abstracts led to identification of only one study relating to childhood cancer that met the inclusion criteria for the review (Table 4).

3.1B.2   Description of studies
The study assessed body size at time of diagnosis in children with cancer in one district in Japan who were diagnosed between the 10-year period 1985-94[23]. The study was based on all cases up to age of 19 years. However the findings relating to cases diagnosed before the age of 1 year were reported separately, so making the study eligible for inclusion.

Table 4: Summary of study included in review of infant growth status and childhood cancer

Ref. no	Title	Author (year)	Design
23	Rapid somatic growth after birth in children with neuroblastoma: A survey of 1718 patients with childhood cancer	Suminoe (1999)	Retrospective cohort

Tables in appendix 10 relating to childhood cancer are tables 3.1B (i) and (ii).

3.1B.3 Assessment of infant size/growth

Infant height and weight at diagnosis were assessed together with change in size from birth to diagnosis. Growth in the children with cancer was compared with that in the general population through calculation of weight SD scores using a survey of 700,000 Japanese children as the reference population.

3.1B.4 Quality assessment

An external comparison group was used in this study with subjects being compared with a general population sample of children. Levels of attrition were low with 90% follow-up of cases up to 19 years of age. Infant size measurements were ascertained from data recorded on the Japanese cancer registry but no information on the methods for measuring the children was reported and so the reliability of methods is unknown. Cases of childhood cancer were ascertained from the cancer register but no information is given on the methods used in diagnosis of the completeness of the register. The confounding factors we considered important in the relationship between infant growth status and childhood cancer were socio-economic status and parental smoking. There was no consideration of confounding in the study and it was considered to have high risk of bias in this respect. Overall we assessed the study as having medium risk of bias.

3.1B.5 Summary of findings

The findings of the study of childhood cancer were inconclusive[23]. The change in weight SD score from birth to diagnosis was not significant in the total group of 292 childhood cancer cases. However, in the subgroup of 175 neuroblastoma cases growth was increased in the first year of life compared with the reference population (p for difference 0.03). However these findings were largely due to the 103 cases detected due to introduction of mass screening (p for difference in this group <0.01).

In summary, there is limited evidence that children who subsequently develop neuroblastoma grow faster during the first year of life. However this is based on the findings of a single study which is not of high quality.

3.1C Respiratory and diarrhoeal infection

3.1C.1 Selection of studies

Searches of Medline and Embase identified 14,495 references relating to respiratory outcomes in the review. These included influenza and pneumonia, chronic lower respiratory disease, asthma and chronic obstructive pulmonary disease. Screening of abstracts led to identification of 3 studies relating to respiratory infection in childhood. Two of these also related to diarrhoeal infection.

3.1C.2 Description of studies

Three studies explored the association of infant size with respiratory infection during childhood and 2 of these also considered mortality or morbidity from diarrhoeal infection as an outcome (table 5).

Table 5: Summary of studies included in review of infant growth status and respiratory and diarrhoeal infection

Ref. no	Title	Author (year)	Design
24	Short-term benefits of catch-up growth for small-for-gestational age infants	Victora (2001)	Prospective cohort
25	The effect of malnutrition on the risk of diarrhoeal and respiratory mortality in children <2 years	Yoon (1997)	Prospective cohort
26	Risk factors for childhood pneumonia among urban poor in Fortaleza, Brazil: a case-control study	Fonseca (1996)	Case-control

Tables relating to respiratory and diarrhoeal infection in appendix 10 are tables 3.1C (i) and (ii).

All 3 were set in developing countries, 2 in Brazil[24;26] and the other in the Phillipines[25]. One of the Brazilian studies by Victora et al was based on the Pelotas cohort, a large-scale prospective study of children born in the urban area of Pelotar, Southern Brazil in 1982[24]. The second Brazilian by Fonseca et al was based in Fotaleza, a predominantly urban area with high levels of poverty where over half of residents were thought to live in shanty towns[26]. The study set in Philippines, by Yoon et al, was based on 33 administrative regions 7 of which were urban and the rest rural[25]. Two of the studies examined the relationship of infant growth with diarrhoea or respiratory disease[24;25]. In the Philippines study the outcome was mortality from either of these causes while the Pelotas study examined the relationship between infant growth and morbidity from either diarrhoea or lower respiratory infection. The Fortaleza study was a case-control study of childhood pneumonia[26]. Cases were selected from a single large public city hospital using a system which rotated the days of the week on which cases were recruited. Neighbourhood controls were recruited through door-to-door enquiries carried out by a fieldworker.

3.1C.3 Assessment of infant size/growth

The Philippines study by Yoon et al assessed growth at 3 intervals during infancy: 6-11 months, 12-23 months and 0-23 months. Weights were converted into weight-for-age z-scores[25]. The study by Victora et al set in Pelotas also assessed growth during infancy. Weights were converted into z-scores using NCHS references and fast growth was defined as change in weight z-score of >=0.66 from 0-20 months. The third study by Fonseca et al explored the relationship of infant size with pneumonia in infancy but the timing of infant size measurements varied for each subject[26]. The subjects, whose ages ranged between 0 and 23 months, had height and weight measured at admission to hospital. Weight-for height, height-for-age and weight-for-age z scores were calculated.

3.1C.4 Quality assessment

The prospective cohort studies by Yoon et al and Victora et al studied population-based samples of children and were considered to have low risk of bias in relation to study design[24;25]. Both studies were based on large samples

of over 1000 subjects in both cases. The study by Fonseca et al was a case-control study which used a sampling method equivalent to random sampling to select cases. Controls were matched with cases according to age and neighbourhood. All three studies were considered to have low risk of bias with respect to study design.

Standardised methods and equipment were used to measure infant size in all three studies but none gave details of observers or reliability and so were assessed as being medium risk of bias in this respect. Outcome assessment was based on information collected during interviews with mothers in both of the cohort studies. Given the potential for recall bias they were both assessed as medium risk of bias in this respect. The diagnosis of pneumonia in the case-control study by Fonseca et al was based on chest x-ray interpretation by a paediatric radiologist and was considered to have low risk of bias in this respect. Outcome assessment was not blind to exposure status in any of the studies.

The confounding factors we considered important in the relationship between infant growth status and respiratory or diarrhoeal infection were socio-economic status, housing conditions particularly overcrowding, parental smoking and infant feeding. Consideration and adjustment for the effects of confounding was thorough in studies by Yoon et al and Fonseca et al and they were both assessed as low risk of bias in this respect[25;26]. However the study by Victora et al made no adjustment for confounding factors and was assessed as having high risk of bias[24].

Compliance with follow-up was high in the case-control study by Fonseca et al with only 1 case refusing to take part. Levels of attrition were higher in the cohort studies with only 61% follow-up in the study by Victora et al and although 88.6% were followed-up in the Phillipines study by Yoon et al.

Overall the studies by Yoon et al and Fonseca et al were assessed as having low risk of bias and the study by Victora et al as medium risk of bias.

3.1C.5 Summary of findings
The cohort studies by Yoon et al and Victora et al both demonstrated an inverse association between infant growth and respiratory and gastrointestinal infection[24;25]. Yoon et al reported the relative and attributable risks for diarrhoea and acute lower respiratory infection in relation to wt-for-age z-scores for each of the 3 growth intervals[25]. For respiratory illness a 1-unit decrease in z-score was related to a significantly increased risk of infection (relative risk (RR)= 3.3, 95% CI 1.4, 7.8). The same trend was also seen in relation to growth from 0-23 months (RR=1.7, 95% CI 1.3, 2.3). However the findings in relation to growth from 6-11 months were not significant (RR=1.5, 95% CI 0.9, 2.5). Similar trends were observed in relation to diarrhoeal infection with statistically significant increase risk of infection at both 6-11 (RR=1.9, 95% CI 1.3, 2.7) and 0-23 months (RR=1.6, 95% CI 0.9, 3.2) and findings of borderline significance at 12-23 months (RR=1.6, 95% CI 0.9, 2.1). The Pelotas study by Victora et al examined the prevalence of hospital admission for diarrhoea and lower respiratory infection according to subjects'

growth from 0-20 months[24]. Infants who had been SGA and AGA were considered separately. The risk of morbidity from respiratory or diarrhoeal infection was related to growth between 0 and 20 months of age. Children who were SGA but had substantial weight gain up to 20 months (>=0.66 weight change z-score) had significantly fewer hospital admissions due to diarrhoeal and respiratory infection. The same trends were seen in the AGA infants.

The case-control study by Fonseca et al also demonstrated an inverse association between infant size and risk of pneumonia[26]. Cases had significantly increased odds of weight-for-height, height-for-age and weight-for-age z-scores of less than –3 after adjustment for important confounding factors. For example the odds ratio of weight-for-height z-score less than -3 compared with greater than or equal to 3 was 6.75 (95% 1.88, 24.27) after adjusting for family income, parental education and previous episode of pneumonia.

Although the evidence on the relationship of infant growth status and respiratory or diarrhoeal infection was relatively limited, the 3 studies reviewed were of high quality and 2 were based on large population-based samples. The measurements of infant size in the study by Fonseca et al were measured at the time of hospital admission raising the possibility that the outcome itself had led to reduced growth. However, this was not the case for studies by Yoon et al and Victora et al which measured size at several points during infancy. In summary the findings suggest that there is an inverse association between infant size and growth and childhood mortality and morbidity from respiratory and diarrhoeal infection.

3.1D Autism

3.1D.1 Selection of studies
A single study relating to autism was identified during searches relating to childhood development which are described in section 3.3.A.1

3.1D.2 Description of studies
The study, which is summarised in table 6, was set in San Diego USA. It was a retrospective cohort study comparing head size and growth in infancy in 48 children with autistic spectrum disorder (ASD) with national reference standards[27].

Table 6: Details of study included in review of infant growth status and autism

Ref. no	Title	Author (year)	Design
27	Evidence of brain overgrowth in the first year of life in autism	Courchesne (2003)	Retrospective cohort

The subjects, 92% of whom were white, were recruited from community advertisements and hospital clinics. All had been of normal birthweight and were born at full-term.

Tables in appendix 10 relating to autism are tables 3.1D (i) and (ii).

3.1D.3 Assessment of infant size/growth

Retrospective data on head size was obtained from medical records and was converted to z-scores using CDC standard growth charts. The head size z-score and head growth z-score at 3 intervals (3-5 months, 6-14 months and 15-28 months) in the children with ASD were compared with those of a normal reference population.

3.1D.4 Quality assessment

We considered the risk of selection bias to be high given the fact that a proportion of the subjects would have volunteered as a result of advertisements. The extent of bias was difficult to assess since characteristics of the subjects were only partially described; no information was given on sex or socio-economic characteristics of the group and there was no comparison with non-participants. Insufficient information was given to be certain of the percentage of potential subjects who were recruited and followed up but it approximately 25% of subjects. Another area in which the study was considered to have high risk of bias was in consideration of confounding since there was no adjustment for confounding factors within the analyses presented. Overall this study was considered to have high risk of bias.

3.1D.5 Summary of finding

Head size and growth were significantly greater in the ASD subjects than the reference population at both 6-14 and 15-28 months. Head growth was also significantly greater at 6-14 months but not at the other 2 intervals considered.

These findings suggest that children with ASD have larger head size in infancy and faster head growth. However, given the methodological weaknesses of the study and its small sample size, these findings need to be replicated in other studies and populations.

3.2 Significant health-related behaviours

The only eligible studies identified were related to childhood obesity, which can be considered to be a proxy for unhealthy patterns of eating and lack of physical activity. No studies were found in relation unhealthy eating or lack of physical activity themselves nor for smoking or substance misuse.

No studies were identified in relation to the non-health related quality of life outcomes for infancy and childhood listed in appendix 3.

3.2A Childhood obesity

Tables in appendix 10 relating to childhood obesity are tables 3.2A (i) and (II).

3.2A.1 Selection of studies

Searches of Medline and Embase identified 26,441 references, whilst 4 further references were identified through contact with experts. Screening of titles and abstracts led to identification of 31 studies that would potentially meet inclusion criteria for the review. Of these, 17 were subsequently ruled out. 12

of these were excluded due to obesity not being formally defined as an outcome – size or adiposity as a continuous variable was used instead. Four further studies did not include measurements during infancy, or relevant results relating to such measurements were not presented, and one final paper did not report any original data. The remaining 14 studies that met selection criteria for the review were scrutinized for references to earlier studies, which identified an additional 7 studies that met the inclusion criteria. Of the 21 studies identified, 9 related to childhood obesity (under 15 years), and these studies are summarised in Table 7. The other 12 studies related to adolescent and adult obesity. Synthesis of these studies is presented in section 5.2A.

Table 7: Summary of included studies in review of infant growth status and childhood obesity

Ref. no	Title	Author (year)	Design
28	Infant weight gain and childhood overweight status in a multicenter, cohort study	Stettler (2002a)	Retrospective cohort
29	Do fat babies stay fat?	Poskitt (1977)	Prospective cohort
30	Prevalence and risk factors for overweight and obesity in children from Seychelles…	Stettler (2002b)	Retrospective cohort
31	Physical development at 7 years of age in relation to velocity of weight gain in infancy	Mellbin (1963)	Retrospective cohort
32	Follow-up study of physical growth of children who had excessive weight gain in first six months of life	Eid (1970)	Retrospective cohort
33	Fat babies and fat children. The prognosis of obesity in the very young	Asher (1966)	Retrospective cohort & case-control
34	Does overweight in infancy persist through the preschool years? An analysis of CDC Pediatric Nutrition Surveillance System data	Mei (2003)	Longitudinal cohort
35	Obesity in urban black adolescents of high and low relative weight at 1 year of age	Johnston (1978)	Prospective cohort
36	Obesity in childhood: A community study in Newcastle upon Tyne	Wilkinson (1977)	Case-control

3.2A.2 Description of studies

The 9 studies explored the relationship between size or growth in infancy and overweight or obesity at ages ranging from 2 to 15 years. All but one had a cohort design; 4 were retrospective, 3 were prospective or longitudinal, one was a retrospective analysis of an originally prospectively designed study, and 1 was of case-control design. All but 1 of the studies were set in developed countries, although the source populations of 2 other studies were largely of low socio-economic status, and one study was restricted to black urban subjects. All the studies had roughly equal proportions of male and female subjects, although one study did not present a sex ratio. Two studies presented results for males and females separately.

3.2A.3 Assessment of infant size/growth

All the studies reported weight, and most height, at around either 6 months or 1 year. Obesity was defined in terms of infant size for 5 of the studies, in all cases using weight or weight-for-height. Two of the 5 used a definition of weight above the 90th percentile of the cohort[33;36] whilst three used varying definitions of increased weight-for-height (more than 1 standard deviation from the mean[35]; over the 95th percentile of the cohort[34]; over 120% of the mean[29]). 'Rapid' infant growth was defined for two further studies, both in terms of weight[31;32]. Of these, one defined rapidity as a total weight gain over the first year of over 7.5kg[31], whilst the other defined it as weight gain over the first 6 months above the 90th percentile[32]. The two remaining studies did not use cut-offs for increased infant size or growth[28;30].

3.2A.4 Quality assessment

For the outcome of childhood obesity the elapsed time between infancy and outcome was very small – a maximum of 15 years by our definition of childhood. This means that a retrospective cohort design may not be as susceptible to recall bias or missing records as for other outcomes. However a prospective cohort design was still considered to be the optimum. Only three studies used a prospective or longitudinal design[29;35;37], although a fourth was a retrospective analysis of an original prospective design[28].

Nearly all the studies were set in developed countries (4 in the UK, 3 in the US and 1 in Sweden), and did not seek to exclude subjects by race, sex or any other demographic variables. However, one study was set in the Seychelles[30], whilst a further study set in the US was restricted to black subjects[35]. In addition, the source population for the study by Poskitt was a 10% sample of the inhabitants of Dudley, West Midlands. However, all members of the study sample are white British, so the sample is unlikely to have been representative of an area in which many ethnic groups live[29].

The target population for most studies was all available subjects from one or more hospitals or schools. The population for the Seychelles study was four specific school grades in a single year, meaning that the age distribution of the subjects is not continuous[30]. Three studies used subjects taken from larger ongoing child health cohorts or databases[34-36]. Aside from the single case-control study, in which subjects were specifically selected to be cases or controls[36], two cohort studies also selected subjects falling into specific infant size or growth subgroups rather than studying the cohort as a whole[32;35]. Many of the older studies, published in the 1960s or 70s, either had deficiencies in design or had a lack of information given in the text, and all were considered to have either a high[33;35] or medium[29;31;32;36] risk of bias overall. For example, the study by Asher was based on child welfare clinic records from a single city, but it is not clear whether all such records were examined, nor is there any consistent information on exclusions or losses to follow-up, or demographic statistics. Two analyses relevant to the review are presented, but it is not clear how the subjects were drawn from the source population or to what degree the subjects in the two analyses overlap[33].

The most common sources of bias were losses to follow-up, a lack of information on non-participants, and lack of consideration of potential confounders. Thus it is hard to judge whether the study subjects are representative of the target population or whether the target population is representative of the general population. The analyses themselves were not generally a serious source of bias, mostly either presenting an odds ratio or relative risk or presenting raw data from which they could be calculated. Two studies were considered to have a high risk of bias overall – the Asher study and the study by Johnston[33;35]. This latter study suffered from a lack of baseline information on the participants, especially since only those with extreme weight-for-height in infancy were included. There was also no information on losses to follow-up from the target subgroup (though the proportion lost to follow-up was fairly low), nor any consideration for potential confounders.

Infant size or growth was in most cases either ascertained from child health records or measured by clinic staff or researchers. Two studies did not give any information on the source of the infancy measurements, but both used subjects from a large national cohort. This encourages faith in the validity of the measurements, but sufficient information was nevertheless not provided[28;35]. Specific details of procedure and reliability checks were rare: one study described the techniques and equipment used[31], whilst a further study also gave details of repeatability checks and the accuracy of the equipment, and stated that a nationally recognised protocol was followed[34]. However, all the studies except that by Asher were considered to be of low or medium risk of bias from infancy measurements.

Obesity at outcome was defined in all but two cases by using weight or weight-for-height. This is in contrast to studies of adult obesity where body mass index (BMI) is commonly used. One study defined childhood obesity using skinfold thickness as well as weight[35]. In all cases, cut-off values were defined by percentile (either the 90th, 95th or 97th) or by percentage of expected (110 or 120%), by comparison with standard reference values. The age at follow-up ranged from 3 to 15 years, although one study considered here included subjects aged between 4 and 17 years, with a median of 11.7 years[30]. Five studies fixed the outcome at a single specific age[28;29;31;32;36]; two studies presented separate results for each of a number of outcome ages[34;35]; whilst two studies measured the outcome over a range of ages that were considered together[30;38]. The procedures used in ascertaining outcome measurements were in all cases similar to those used in infancy – either measured at medical centres by nurses or researchers, or from school or clinic records. If insufficient information was given for infancy measurements, a similar lack of information was found for the outcome measurements. Thus, all the studies were considered to have either a medium or low risk of bias for outcome measurements.

The confounding factors we considered important were socio-economic status, parental size and infant feeding. Two studies adjusted for all or most of these in regression analyses[28;30], whilst two further studies did not adjust but presented separate analyses showing the relationship between potential

confounders and the outcome[32;36]. However, four studies were considered to present a high risk of bias from confounding[29;31;35;38], whilst only one study was considered to present a low risk[30].

3.2A.5 Summary of findings

Of the 9 studies of childhood obesity, 5 studies measured size in infancy, 3 studies measured infant growth only, and 1 study measured both size and growth.

Three of the studies examining infant size specifically explored the risks of childhood obesity (designed according to weight-for-height cut-offs) in children who had been obese as infants. These were the studies by Poskitt et al, Mei et al and Asher et al[29;33;34]. The studies by Poskitt et al and Mei et al demonstrated that infant obesity was associated with an increased risk of childhood obesity up to 5 years of age. The findings of the study by Asher et al were inconclusive. While weight above the 97th centile at 6 months of age was associated with a significant risk of obesity at 5 years of age (RR=6.56, 95% CI 2.90, 14.8) in the nested case-control analysis which was part of this study, the associations between weight above the 90th centile at 6 months and later obesity in the entire cohort was not significant. Of the other 3 studies of infant size, 2 demonstrated that being heavier in infancy was associated with increased likelihood of childhood obesity[28;35]. The remaining study relating to infant size was the case-control study by Wilkinson which reported outcome data at 10 years. The estimated odds ratios for obesity by infant obesity (defined according to weight) at 6 and 12 months were 2.0 and 1.6 respectively; neither were statistically significant[36]. This study also suggested that the prevalence of risk factors such as obese parents, an obese sibling, being an only child or having an older mother were higher among cases than controls, though no statistical analyses were presented.

The findings of the studies of infant size and the subset that assessed infant obesity are summarised in table 8. Although they are presented separately the definition of exposure status in the infant obesity studies overlaps with that of infant size. The studies of infant size also considered infants at the highest end of the distribution for weight or weight-for-height. The findings of the studies of infant size and obesity are consistent in demonstrating that infants at the highest end of the size distribution for weight or weight-for-height are at greater risk of subsequent obesity than other infants.

Four studies assessed the relationship of infant growth with later obesity[28;30-32]. Their findings are also summarised in table 8. The American study by Stettler et al suggested that rapid weight gain in the first 4 months of life was associated with increased risk of subsequent obesity. The estimated odds ratio for childhood obesity for each 100g increase in 1-year weight, after adjustment for birthweight, maternal BMI and education, infant feeding and others, was 1.5 (95% CI 1.4 to 1.6)[28]. The second study by Stettler et al set in the Seychelles also showed that greater weight gain in the first year of life was associated with increase risk of obesity at school age[30]. The findings of the study by Mellbin et al based in Uppsala, Sweden were inconclusive[31]. While they suggested that rapid weight gain over months 1-4 and 9-12 was

associated with increase risk of obesity at 7 years of age in boys, the findings for girls and for weight gain across the first year in both sexes were not significant. Similarly Eid et al failed to demonstrate a significant association between weight gain in the first 6 months and obesity at 8 years of age[32].

Table 8: Summary of findings relating infant size, obesity and growth to childhood obesity

Direction of relationship with infant size/growth	Number of studies (ref. nos. in parentheses)			
	Infant size	Infant size (obesity)	Infant growth	Total
Inverse	0	0	0	0
None	1 [36]	0	1 [32]	2
Positive	4 [28;29;34;35]	2 [29;34]	2 [28;30]	5
Inconclusive	1 [33]	1 [38]	1 [31]	3
Total	6	3	4	9*

*This figure is the total number of studies examined in the table, and may not equal the sum of the subtotals due to studies appearing more than once in certain cells of the table.

Four studies of infant size and childhood obesity either reported a relative risk of childhood obesity or allowed one to be calculated[29;33-35]. Three of these were the studies of infant obesity by Poskitt et al, Asher et al and Mei et al. The studies by Poskitt and Asher reported infancy data at around 6 months and outcome data at around 5 years, and both reported fairly large values (6.6 and 9.4 respectively), but both were small studies and the estimates had wide confidence intervals. The study by Johnston reported infancy data at 1 year and outcome data at 9 to 15 years, and was slightly larger (n = 798). The study by Mei reported data at various ages, but for comparison, the analysis relating obesity at 12-23 months to that at 48-59 months was considered. When combined in a meta-analysis, the effect of the Mei study, with a sample size of over 76,000, dwarfed the effects of the smaller studies. The combined estimate was for the risk of becoming an obese child among obese infants to be just under 3 times as large as among non-obese infants (estimate 2.9; 95% CI 2.8 to 3.1). As with the adulthood obesity analysis, any studies stratifying data by sex were entered into the meta-analysis as two studies. Despite the differences in age at outcome for the Johnston study, the estimates were similar (3.0 for males and 2.7 for females). The I^2 heterogeneity statistic was 51% (p = 0.09), indicating a significant amount of heterogeneity, probably due to the large imbalance in study sizes.

The UK and Swedish studies of infant growth reported their findings in a way that allowed calculation of a combined relative risk, although the other two studies of growth could not be included in this analysis. The UK and Swedish studies both related infant growth to childhood obesity and defined rapid growth' as an increase in weight from 6 months to 1 year of age over the 90th percentile. Combining the findings of these studies gave an estimated relative risk for becoming an obese child among rapidly-growing infants compared to normally-growing infants of 2.3 (95% CI 1.9 to 2.9)[31;32]. The UK study also suggested that, among rapidly-growing infants, the risk of later obesity is

increased for infants of higher birthweight and for those who were not breastfed[32].

It is notable that 6 of the 9 studies of childhood obesity were published before 1979. The findings of these studies were less conclusive than the studies published more recently with only 2 of these studies demonstrating a statistically significant positive association between infant growth status and adult obesity. As outlined in section 3.2, the quality of these studies was also lower than the more recent studies with 2 studies having high risk of bias and the remainder being of medium risk.

In summary, the risk of becoming obese in childhood was significantly greater among infants who were at the highest end of the size distribution for weight or weight-for-height and among rapidly-growing infants.

3.3 Childhood development

Childhood development was the area of the review where we identified the greatest number of studies; 52 in total. The majority of these (n=45) related to cognitive development while 10 related to motor development with 3 studies considering both cognitive and motor development. We did not identify any studies relating to sensory development that met our inclusion criteria. The same was true for social development, although a number of the studies summarised in the cognitive development section did address aspects of behaviour.

Given its large size, the review of cognitive development is divided into three sections based on the aspect of infant growth status assessed (head size or growth, growth faltering, or size/growth in length or weight).

3.3A Studies relating head size/growth to cognitive development

Tables in Appendix 10 relating to head circumference and cognitive development are tables 3.3A (i) and (ii).

3.3A.1 Selection of studies

Searches of 7 electronic databases were carried out: Medline, Embase, CINAHL, PsycInfo, ERIC and Sociological Abstracts. LILACS was also searched for foreign language studies although this did not lead to identification of any new studies. Studies exploring the relationship between infant growth status and childhood development (this includes motor, emotional and sensory development as well as cognition) were sought. Searches identified 25,575 abstracts. After independent screening by 2 reviewers 106 papers were sent for and assessed and a further 22 were found as a result of searching on other outcomes. Of the 128 assessed, 51 were considered to meet inclusion criteria for the review. Forty-four of these related to cognitive development. A further, as yet unpublished, study relating to cognitive development was identified through correspondence with first authors.

This section describes the findings of 16 studies relating head size or growth in infancy to later cognition. The studies relating head size or growth to cognitive development are summarised in Table 9.

Table 9: Summary of included studies in review of infant head size or growth and cognitive development

Ref. no	Title	Author (year)	Design
39	Patterns of cognitive development in very low birth weight children during the first six years of life	Koller (1997)	Prospective cohort
40	Relationship of growth and psychoneurologic status of 2-year-old children of birthweight 500-999 g	Ford (1986)	Longitudinal cohort
41	Children with superior intelligence at 7 years of age: a prospective study of the influence of perinatal, medical, and socioeconomic factors	Fisch (1976)	Prospective cohort
42	Intrauterine growth retardation and neurodevelopment at one year of age in Mexican children	Fernandez-Carrocera (2003)	Prospective cohort with nested case-control
43	Physical and behavioural correlates of 12-month development in small-for-gestational age and appropriately grown infants	Pryor (1996)	Longitudinal cohort
44	Growth parameters and attention to faces at 4 to 6 months of age	Camp (1990)	Cross-sectional
45	Head size at one year as a predictor of four-year IQ	Nelson (1970)	Prospective cohort
46	The relationship between physical growth and infant behavioral development in rural Guatemala	Lasky (1981)	Prospective cohort
47	Effect of very low birth weight on cognitive abilities at school age	Lucas (1992)	Prospective data from RCT
48	Size and cognitive development in the early preschool years	Ernhart (1987)	Prospective cohort
49	Head circumference in ELBW babies is associated with learning difficulties and cognition but not ADHD in the school-aged child	Stathis (1999)	Prospective cohort
50	Very low birth weight infants: effects of brain growth during infancy on intelligence quotient at 3 years of age	Hack (1986)	Prospective cohort
51	Relationship between head growth and neurodevelopmental outcome of Malaysian very low birthweight infants during the 1st year of life	Ong (1997)	Prospective cohort
52	Critical periods of brain growth and cognitive function in children	Gale (2004)	Prospective cohort
53	Developmental implications of head growth following intracranial hemorrhage	Bendersky (1998)	Prospective cohort
54	Effect of VLBW and subnormal head size on cognitive abilities at school age	Hack (1991)	Prospective cohort

3.3A.2 Description of studies

Of the 16 studies relating head size or growth to later cognition, 9 were based on populations of low birthweight (LBW) or very low birthweight (VLBW) preterm babies while 7 were based on normal populations of children. LBW is defined as birthweight below 2.5 kg and VLBW as below 1.5kg. The majority of studies (n=13) were prospective cohort studies of children followed up from infancy. Of the remaining three studies, the first by Fernandez-Carrocera et al was a nested case-control study based on a cohort of infants hospitalised in a single neonatal intensive care unit (NICU)[42]. The second by Lucas et al was based on all participants in a randomised trial of post-natal nutrition in preterm VLBW infants[47]. The third by Camp et al was a cross-sectional study of infants attending a single paediatric clinic to assess their visual response to stimuli[44].

All but 3 of the studies were based in developed countries: USA, UK, Australia and New Zealand. The 3 developing country studies were the study by Lasky et al based in 4 rural villages in Guatemala, the study by Ong et al based in Kuala Lumpar, Malaysia and the study by Fernandez-Carrocera et al based in Mexico. The Malaysian and Mexican studies were both based on infants who had been treated in NICUs whereas the Guatemalan study was based on a sample of all live-born infants in the villages studied.

Subjects in 2 of the 9 studies of VLBW infants were described as extremely low birthweight infants (ELBW) defined as mean birthweight <1000g. Mean birthweight in both of these studies by Stathis et al and Ford et al was 0.86kg and mean gestational age was 27.5 and 27.7 weeks respectively[40;49]. In 4 of the remaining 7 studies mean birthweight in 4 of the studies ranged between 1.17 and 1.45kg and gestational age between 29.1 and 31.8 weeks. The study by Fernandez-Carrocera et al was based on borderline preterm infants (mean gestational age 36 weeks) born in Mexico with a mean birthweight of 2.09kg[42]. Lucas et al did not report mean birthweight although subjects were said to be <32 weeks gestational age[47]. All of the subjects in these studies were selected from babies who had required admission to or treatment on NICUs. There is considerable scope for selection bias in these studies given that babies admitted to NICU may have experienced perinatal brain injury which is associated with slower head growth. Only two of the studies were based on infants selected from more than one hospital. The study by Lucas et al was a randomised multi-centre trial of preterm VLBW infants in 5 UK towns and cities[47]. The study by Koller et al was based on subjects recruited from 3 NICUs in New York[39]. Five of the studies were based on a single NICU or single hospital[40;42;49-51] and the 2 studies by Hack et al were based on the same cohort of VLBW babies followed up at 33 months and 8 years. The remaining study by Bendersky et al was based on VLBW infants with and without intracranial haemorrhage (ICH) though the setting for the study was not entirely clear[53].

Of the seven studies based on populations of normal birthweight children, all but one were set in developed countries (USA, New Zealand or the UK)[41;43-45;48;52]. Two of the American studies by Nelson et al and Fisch et al were based on populations of children born in the 1950s and 1960s and in fact both

used the same national perinatal study as the source for their target populations[41;45]. The other developed country studies by Gale et al, Pryor et al and Ernhart et al were all more recent and were based on children born in the 1980s and 90s[43;48;52]. The developing country study by Lasky et al was based on live births in four rural villages in Guatemala[46]. Like the studies based on LBW infants, a number of the study samples of normal birthweight children were selected from single hospitals or paediatric clinics. Pryor et al set their study in a single hospital in Auckland, New Zealand[43], Camp et al conducted their investigation in a single paediatric clinic in Denver, USA[44] and Gale et al looked at births at one UK maternity hospital[52]. However three of the studies were population-based. The 2 American studies by Nelson and Fisch used the same large-scale perinatal study as the source for subjects in their studies, although the study by Fisch et al was based on Minnesota births only[41;45]. The Guatemalan study by Lasky et al was based on all live births from 4 rural villages[46].

Cognitive development was assessed at a range of ages and consequently a variety of scales and measures are used across the studies. Ten of the studies considered cognitive development during infancy and of these, the majority used Bayley scales of infant development reporting mental development index (MDI) as their outcome measure. This applied to all but one of the 7 studies of cognition in infancy that were set in USA, New Zealand or the UK[39;40;43;47;48;53]. The remaining developed-country study by Camp et al assessed response to visual stimulation in infants of 4-6 months, a measure which is said to be correlated with intelligence in childhood, but they did not measure cognition directly and so cannot be compared with the other studies[44]. Three of the studies in infancy were set in developing countries and 2 of these used different approaches to assessment of cognition. The study by Fernandez-Carrocera et al which was based in Mexico used a number of different scales to evaluate communication, psychological and neuromotor aspects of development in LBW infants as well as assessment of any neurological abnormalities[42]. They report their findings in terms of the relationship of size with neurological alteration which appears to be defined as having one or more abnormal result on any of these scales, although this is not entirely clear. The study by Lasky et al based on infants living in rural Guatemala, assessed infants mental and motor development using a Composite Infant Scale (CIS)[46]. Although this represents a different approach to many of the other studies of infant cognition, there are some similarites in that the CIS is partly composed of elements from Bayley scales. The third developing country study by Ong et al based on Malaysian infants used Bayley scales of infant development[51].

All 4 studies that assessed cognition at ages 3 and 4 years used Stanford-Binet scales which measures IQ[39;45;48;50]. These studies were all set in the USA. However one of the studies, by Nelson et al which was based on children born between 1959 and 64 included over 50% subjects who were black or from ethnic minorities[45]. Given that Stanford-Binet was standardised for use in white populations this was a potential flaw in assessment of cognition in this study.

Six of the studies assessed cognition at 5 years and older[39-41;47;49;52]. Five of these, including the studies by Koller et al and Lucas et al[39;47] which also assessed cognition in infancy, used Wechsler intelligence scales either in their revised form for children (WISC-R) or in the adult form in the case of the study by Gale et al which assessed cognition at 9 years[52]. The sixth study by Stathis et al assessed cognition at 6 years using McCarthy scales which yield a general cognitive index (GCI) and this study also used an ANSER questionnaire (a well-validated questionnaire administered to teachers) to evaluate learning problems at school[49].

3.3A.3 Assessment of infant size/growth

All 16 studies explored the association between head size in infancy and cognition during infancy and/or childhood. However, only 3 of the studies also explored the relationship between head growth in infancy and later cognition. Hack et al explored the association between head size at 8 months, growth between 8 and 20 months and later cognitive function at 3 years in the first study and 8.6years in the second study[50;54] while the study by Gale et al examined head circumference (HC) at 9 months of age and also growth up to 9 months by adjusting for HC at birth in regression analysis[52]. Seven of the 13 remaining studies focused on head circumferences at age 1 and the remainder on HC at either 6 months, 8 months or 2 years with 3 studies looking at more than one of these. Lasky looked at HC at 6 months and 2 years[46], Lucas at 9 and 18 months[47] and Ernhart at 6 months and 2 years[48].

3.3A.4 Quality assessment

Although the majority of studies were prospective cohort studies, which would seem to be an optimal design for assessing the relationship between infant size or growth and later development, few samples were population-based, as discussed in section 3.3A.2. Only the studies by Lasky et al and Fisch et al were considered at low risk of bias in relation to design and setting. With the exception of one study, the remainder were assessed as having medium risk of bias due to selection procedures. The study by Ernhart et al was scored as having high risk of bias due to the fact that subjects were participants in a study exploring the relationship between maternal alcohol intake and birth defects[48].

Three of the studies gave poor baseline information on participants and were therefore considered as being at high risk of bias. This applied to the studies by Lasky et al and Ernhart et al[46;48]. The third study by Lucas et al was published in the form of a letter which was thought to contribute to an initial lack of baseline information[47]. However, even after two earlier references cited in the letter were obtained, there was still only an adequate description of a sub-set of the participants.

The description of measurements of head circumference (HC) in infancy was good in a number of studies and the 5 that gave descriptions both of how the measurements were carried out as well as information on observers and reliability were assessed as having low risk of bias[42;46;51-53]. Rounding was not considered a source of bias in these 16 studies since most reported measuring HC to the nearest 0.1cm.

Outcome measurement was generally carried out with well-validated scales appropriate to age at assessment and there was considerable consistency between studies as discussed in section 3.3A.2. However only the 5 studies that gave information on observers and reliability were rated as having low risk of bias[46;48;51-53] and only three studies stated that observers were blind to infant growth status[44;51;53].

The confounding factors we considered important in the relationship between infant growth status and cognitive development were socio-economic status, parental (particularly maternal) education, infant feeding and gender. Seven studies were poor in consideration of confounding and did not adjust for any of these factors in the analyses of interest to the review[39-44;47]. These studies were assessed as having high risk of bias in this respect. Only 3 studies were classified as at low risk of bias. These were the studies by Gale et al which was the only study to consider all of the confounding factors[52], and the studies by Ernhart et al which considered all but infant feeding[48] and Lasky et al[46].

There were high levels of attrition in many of the studies and 9 studies were assessed as having high risk of bias due to loss to follow-up[45-53], although 2 of these studies gave sufficient information on non-participants to suggest that those lost to follow-up had similar patterns of infant growth to subjects thus reducing the risk of bias[51;52].

Assessment of overall risk of bias in these 16 studies is summarised in appendix 10, table C. Only the studies by Ong et al and Gale et al were considered to have low risk of bias overall[51;52]. Whilst the majority of studies were assessed as having medium risk of bias, 3 studies were seen as having high risk of bias. In the case of the studies by Ernhart et al and Lucas et al this was primarily due to potential for selection bias and lack of information about certain aspects of the study[47;48]. In the case of the study by Nelson et al the main flaws were seen as the lack of statistical analysis and the inappropriate use of cognitive assessment scales in a predominantly black study population[45].

3.3A.5 Summary of findings

All 16 studies explored the association between head size during infancy and later cognition. Seven focused on HC at 1 year and 5 of these showed a significant positive association with cognition assessed both in infancy and childhood. Of the 5 showing a positive association, the studies by Koller et al and Fisch et al focused on cognition up to 6 and 7 years respectively[39;41] whereas Bendersky et al, Ong et al and Pryor et al looked at the association with Bayley MDI and PDI at 1 year[43;51;53]. Koller et al demonstrated a significant trend across clusters defined by IQ where the percentage of infants with HC<10th centile was significantly greater in the groups with poor cognitive function[39]. Fisch et al demonstrated a 0.5cm differences in mean HC at 1 year when a group with superior intelligence at 7 years (defined as IQ≥120 points) was compared with an average group (IQ 80-119)[41]. Bendersky et al present their findings in the form of standardised beta weights which are achieved by converting all independent variables to z scores prior to regression. As stated

in appendix 10, table 3.3A (i) the closer the value to 1 the stronger the ability of that variable to predict the outcome[53]. Values for the association between 1 year HC and MDI and PDI were 0.23 and 0.27 respectively both of which were statistically significant (p<0.01). Ong et al presented their findings in the form of regression coefficients which corresponded to the increase in MDI or PDI score for every unit increased in the value of 1-year HC[51]. The coefficients for MDI and PDI were 241.6 and 179.2 respectively. Pryor et al demonstrated a positive association in small for gestational age (SGA) infants but this was absent in appropriate-for-gestational-age infants[43]. Of the remaining 2 studies, Fernandez-Carrocera et al showed only a borderline effect of low HC at 1 year on neurological alteration (defined as one or more abnormal results across a range of measures) with odds ratio 2.58 (95% CI 1.00 to 6.64)[42]. Nelson et al did not conduct any statistical analysis of their data although the trends in the raw data do suggest a positive association between HC at 1 year and 4-year IQ[45].

Eight studies explored the association of head size between 6 and 9 months and later cognition[44;46-50;52;54]. All but one of these demonstrated a positive association between head size and cognitive development at ages ranging up to 8.6 years. Four of the studies give some indication of size of effect through reporting regression coefficients for the relationship between HC and cognitive development: Gale et al reported that 9-year IQ increases by 1.98 for every unit increase in 9 month HC (95% CI 0.34 to 3.62)[52]. Hack et al reported a 9.96 reduction in 3-year IQ for HC more than two standard deviations below the mean (p=0.01). In the second study by Hack et al, verbal IQ increased by 7.78 (SE 3.04) for every unit increase in 8-month HC (p=0.011) and the corresponding values for performance IQ were 7.70 (2.84), p=0.007. Stathis et al report a decrease of 8.30 in mean GCI score at 6 years for the presence of HC ≤10th centile at 8 months (p=0.04)[49]. Three studies reported the correlation of HC with later IQ. Lucas et al show that 9-month HC is significantly correlated with both 9-month (r=0.21) and 8-year (r=0.29) cognition[47]. Lasky reported a correlation of 0.12 between HC at 6 months and 6-month (r=0.12) and 2-year CIS (r=0.15)[46]. Ernhart reported a correlation of 0.13 between 6-month HC and 3-year IQ although no association was observed with cognition at 6 months[48].

Four studies explored the association between HC at 2 years and cognition up to 3 years of age[40;46-48]. Three of these showed a significant positive association and all report correlation coefficients. In the study by Lucas et al, 18 month HC and 18 month developmental quotient were significantly correlated (r=0.17, p<0.0005)[47]. In the study by Lasky et al the correlation between 2-year HC and 2-years CIS was 0.18[46] and in the study by Ernhart et al the correlation between 2-year HC and 3 year MDI was 0.22 (p<0.001) although the same study did not find a significant correlation with MDI at 2 years[48]. Ford et al failed to show an association between 2-year HC and 2-year MDI[40]. However the sample size in this study was small and so it may have been statistically underpowered.

Only 3 studies assessed the relationship of head growth in infancy and cognitive development. Gale et al demonstrated a positive association

between head growth up to 9 months and IQ at 9 years with a regression coefficient of 2.30 (95% CI 0.56 to 4.03)[52]. However the other two studies failed to demonstrate significant associations. Hack et al explored the association between head growth from 8 to 20 months with 3 year IQ and Stathis et al between head growth from 4 to 8 months of age and learning problems at school[49;50].

Given the differing approaches to analysis and presentation of findings across studies coupled with the differences in outcome measure and age at follow-up, meta-analysis was not feasible. Nevertheless 12 of the 16 studies demonstrated a significant positive association between head size in infancy and cognitive development at ages ranging up to 9 years, as shown in Table 9 below. Of the 4 studies that failed to show an association 2 (the studies by Pryor et al and Ford et al) were based on small sample sizes of less than 100 subjects and so may have been underpowered[40;43]. The findings of the study by Nelson et al, which was based on over 9000 subjects, strongly suggest a positive association but in this case no statistical analysis was carried out[45]. In the fourth study by Fernandez-Carrocera et al, the unadjusted analyses suggest an association between low weight at 1 year and neurological abnormality[42]. However logistic regression adjusting for IUGR and social factors suggested that this effect may be largely attributed to IUGR.

In light of the potential differences between LBW and normal birthweight influences, particularly in relation to perinatal influences on growth, we compared the findings of the 9 studies of LBW or VLBW infants with those of normal birthweight infants. This revealed that the positive association between head size and cognition was present in both groups (table 10).

Table 10: Summary of cognitive development findings

Direction of relationship with infant head size	No. studies of LBW or VLBW infants	No. studies of normal birthweight infants	Overall no of studies
Inverse	0		0
Positive	7	5	12 [39;41;44;46-54]
Inconclusive	2	2	4 [40;42;43;45]
Total	9	7	15

Overall the findings of these studies strongly support a positive association between head size and cognitive development during childhood up to 9 years of age. However the findings in relation to head growth are inconsistent and are based on fewer studies and so no firm conclusions can be reached.

3.3B Studies relating growth faltering to cognitive development

3.3B.1 Selection of studies

This section describes the findings of 19 studies that explored the relationship of growth faltering with later cognition (table 11). These 19 studies specifically set out to look at growth faltering (usually described as failure to thrive) and we have therefore grouped them in this section. However there is some overlap with studies included in section 3.3C which described studies exploring the relationship between weight and height in infancy and cognitive development.

Table 11: Summary of included studies in review of infant growth faltering and cognitive development

Ref. no	Title	Author (year)	Design
55	Postnatal growth and mental development: evidence for a "sensitive period"	Skuse (1994)	Longitudinal cohort with nested case-control
56	The developmental sequelae of nonorganic failure to thrive	Boddy (2000)	Prospective cohort
57	Does a fall down a centile chart matter? The growth and developmental sequelae of mild failure to thrive	Corbett (1996)	Prospective FU of earlier case-control
58	Cognitive and educational attainments at school age of children who failed to thrive in infancy: A population-based study	Drewett (1999)	Prospective cohort with nested case-control
59	Early psychological outcomes in failure to thrive: Predictions from an interactional model	Drotar (1985)	Prospective cohort with intervention
60	Prediction of intellectual development in young children with early histories of nonorganic failure-to-thrive	Drotar (1988)	Prospective cohort with intervention
61	The cumulative effect of neglect and failure to thrive on cognitive functioning	Mackner (1997)	Cross-sectional
62	Cognitive development of children in poverty with failure to thrive: a prospective study through age 6	Mackner (2003)	Prospective cohort with intervention
63	Failure-to-thrive, maltreatment and the behavior and development of 6-year-old children from low-income, urban families: a cumulative risk model	Kerr (2000)	Prospective cohort with nested case-control
64	Risk factors and outcomes for failure to thrive in low birth weight preterm infants	Kelleher (1993)	Prospective data from RCT
65	Concurrent sensorimotor and visual-perceptual functioning in failure to thrive infants	Singer (1984)	Longitudinal cohort
66	Physical and psychological development of children with early failure to thrive	Glaser (1968)	Retrospective cohort
67	Long-term follow-up and outcome of infants with non-organic failure to thrive	Reif (1995)	Retrospective cohort with nested case-control
68	Follow-up developmental status of infants hospitalized for nonorganic failure to thrive	Field (1984)	Longitudinal cohort
69	Prognostic significance of postnatal catch-up growth in VLBW infants	Hack (1982)	Prospective cohort
70	Failure to thrive: a study in a primary care setting	Mitchell (1980)	Prospective cohort
71	Relationship between postnatal growth on psychomotor development in VLBW infants	Tudehope (1983)	Prospective cohort
72	Facial expressivity in failure to thrive and normal infants	Abramson (1991)	Prospective cohort
73	Behavioral and cognitive status in school-aged children with a history of failure to thrive	Dykman (2001)	Prospective cohort

Tables in appendix 10 relating to growth faltering and cognition are tables 3.3B (i) and (ii).

3.3B.2 Description of studies

Nineteen studies explored the association between growth faltering and cognition during infancy and/or childhood. Seventeen of the studies were set in either the UK or USA, 1 was set in Australia and the remaining study in Israel. According to the GBDS Israel is regarded as a developing country.

Nearly all studies refered to failure to thrive (FTT) rather than growth faltering. We therefore use the term FTT in order to be consistent with how study characteristics and findings were described.

The 4 UK-based studies were all population based prospective cohort studies comparing FTT infants with a comparison group using a nested case control design. Skuse et al studied a cohort of infants selected from a survey of all births in 1986 in a socially disadvantaged district of South East London[55]. Subjects were followed up until 15 months of age and cognitive development of 47 FTT infants was compared with that of 47 non-FTT controls. Boddy et al followed the same cohort of children up at age 6 years comparing cognition of 42 cases of FTT and 41 comparison infants[56]. Similarly, the study Drewett et al was based on 107 cases of growth faltering and 117 comparison infants selected from a cohort of births during 1986 in Newcastle-upon-Tyne who were followed up until 8 years of age[58]. The study by Corbett et al, which was also based in Newcastle-upon-Tyne, was a follow-up of FTT and control infants born during 1985-86 who took part in a previous case control study to assess their cognitive and behavioural development at 6-7years of age[57]. Although the setting and timing of births in the two Newcastle studies are similar they do not appear to be based on the same populations of children. The study by Corbett et al was based on infant registered at 2 city child health clinics while the study by Drewett used all births in the study region in 1986 as the source population.

Eleven of the 13 USA studies were based on FTT cases admitted to or treated in particular hospitals or clinics. The 1985 study by Drotar et al was based on 68 infants hospitalised for FTT in 7 hospitals in Cleveland, Ohio whose cognition was assessed at 12 and 18 months of age[59]. The 1988 study by the same authors was a later follow-up at 36 months of age of 59 of the infants from the 1985 study[60]. In both of these studies the subjects had been recruited to the study and had received post-discharge family support interventions to treat the FTT. In the 1997 study by Mackner et al, 177 subjects from an inner-city paediatric clinic in Baltimore were followed up between 3 and 30 months of age to assess the association of FTT and neglect with cognitive development[61].   Subjects fell into 4 groups: those with FTT only (n=64), those with maltreatment only (n=21), those with both (n=28) and a control group with neither risk factor (n=80). The 2000 study by Kerr et al was a follow up of the same infants at 6 years of age[63]. The later study by Mackner et al in 2003 was an RCT of home intervention for FTT based on subjects recruited from the same clinic as the 1997 study[62]. The 128 FTT cases were compared with 98 non-FTT controls at yearly intervals from 3 to 6 years of

age.   The 1993 study by Kelleher et al was also a prospective follow up of subjects taking part in an RCT[64]. In this case the interventions were targeted at LBW preterm infants and included stimulation and home visits. 180 FTT and 591 non-FTT infants were followed up until 36 months of age to assess cognition and behaviour. The study by Mitchell et al was based on children aged 2 to 5 years of age attending 3 clinics which were part of a federally funded project in a deprived area. No information on mean birthweigth is reported but 10% were preterm (<37 weeks). The studies by Abramson et al and Dykman et al were both small-scale cohort studies of FTT cases. In the study by Abramson et al the emotional development of 12 FTT infants from 2 children hospitals in Michigan were compared with 12 normal infants[72]. In the study by Dykman et al 27 FTT cases form a single children's clinic were compared with 17 comparison children from the local community to assess cognition between the ages of 8 and 12 years[73].

The American study by Hack et al and the Australian study by Tudehope et al were cohort studies of VLBW (<1.5kg) infants who were admitted to single NICUs in 1977-78 and 1978-80 respectively[69;71]. In both studies the developmental outcome of the AGA and SGA infants was related to their size in infancy – at 8 months in the study by Hack et al and at up to 4 years of age in the study by Tudehope et al.

In the 1984 study by Singer et al subjects were recruited from 3 hospitals in Cleveland, Ohio and this is the only study that included organic cases of FTT[65]. Three groups were compared each including 13 subjects. The groups were infants with organic FTT, non-organic FTT and a group of normal control infants. All 3 groups were followed up at 2, 20 and 36 months of age. In order to be consistent with the other studies in the review we are presenting the comparison of the non-organic FTT and control group only. The study by Glaser et al was based on 40 infants with FTT discharged from a single Boston hospital between 1958-65[66]. Subjects were followed up to assess cognition at a wide range of ages from 8 months to 11 years of age (mean 4.5 years). Similarly, Field et al studied 17 infants admitted to a single Philadelphia hospital with growth failure[68]. The age at which cognition was assessed is not entirely clear although it is below 2 years of age since infants were followed up between 6 and 13 months post-discharge and mean age at discharge was 8.5 months.

The Israeli study by Reif et al, set in a single medical centre in Israel, followed up cases of FTT diagnosed between 1982-90 until 8 years of age and compared them with a control group. The control infants were selected from admissions to the same medical centre[67].

As stated above, the studies by Hack et al and Tudehope et al were based on cohorts of VLBW babies with mean birthweight less than 1.5kg[69;71]. While the majority of other studies state that participants were of normal birthweight and were not preterm, only 3 of the studies of normal birthweight infants reported mean birthweight or gestational age[55;56;65]. Two studies did include preterm LBW infants. In the first of these by Field et al, a third of the 17 subjects were described as preterm with mean birthweight of 2.0kg[68]. The second by

Kelleher et al which was an RCT of interventions for LBW preterm infants, mean birthweight was 1.8kg and mean gestational age 33 weeks[64]. Information relating to potential confounding factors of socio-economic status and parental education was more consistently reported with the majority of studies giving baseline information on one or both of these.

As in the studies of head circumference, cognition was assessed at a range of ages with a variety of scales and measures. Eleven of the studies consider cognitive development in infancy and of these 8 use Bayley scales of infant development reporting MDI at a range of ages up to 24 months as the outcome[55;59-62;65;68;69]. The study by Glaser et al assessed development at a range of ages from 8 months up to 11 years (mean 4.5 years) using a variety of scales including Cattell IQ scales in infancy, Stanford-Binet and Wechsler WISC at later ages[66]. The remaining study by Abramson et al assessed expressivity and emotional development during an 80-minute videotaped session in which the infants were variously stimulated, tested and allowed to play[72].

Two of the studies assessed cognition at around 3-4 years of age. In the study by Tudehope et al cognition was assessed up to 4 years of age with children below 3 years being assessed with Griffiths test and those aged 3-4 years with McCarthy scales[71]. Mitchell et al assessed children between 3 and 6 years of age using McCarthy scales[70].

Eight of the studies assessed cognition at 6 years of age or above[56-58;62;63;66;67;73]. Six of these including 2 of the UK studies by Drewett et al and Corbett et al used Wechsler intelligence scales either in their revised form for children or Wechsler pre-school or primary scale of intelligence-revised (WIPPSI-R). Another of the UK studies by Boddy et al assessed cognition using McCarthy scales[56]. The final study by Reif et al, which was based in Israel examined educational and developmental achievements at 6 years of age[67]. However the outcomes are not clearly defined and validated scales for assessing cognition do not appear to have been used in this study.

As well as measuring cognition, a number of the studies also reported behavioural outcomes at follow-up. The study by Corbett et al based in Newcastle-upon-Tyne assessed behaviour at 6 years of age using the Child Behaviour Checklist (CBCL)[57]. Kerr et al also used the CBCL which was administered to mothers when the subjects were 6 years of age together with the Teachers report form which allowed teachers to report on behaviour and adaptive functioning at school[63]. Kelleher et al used CBCL at 24 and 30 months, Bates temperament scale at 12 months and a prosocial behaviour scale at 36 months[64].

3.3B.3 Assessment of infant size/growth
In general studies used the term 'failure to thrive' (FTT) rather than growth faltering. Definitions of FTT varied between studies although all but one study based their definition on the body weight of subjects at one or more point during infancy.

The 4 UK studies defined FTT in terms of growth in body weight during infancy. The studies by Skuse et al and Boddy et al used the same definition[55;56]. FTT was diagnosed when the weight-for-age z-score was below the 3rd centile at 12 months of age and that this was maintained for at least 3 months. Drewett et al defined FTT according to a thrive index measured at 3-monthly intervals, where the index at a particular point was calculated by regression of current weight on a baseline weight[58]. Infants were considered to be failing to thrive when the thrive index was below the 5th centile on at least 2 occasions between 3 and 18 months. Corbett et al defined FTT as weight declining from a maximum achieved at 4-8 weeks, crossing 2 or more centiles downwards and remaining at this low level for at least 1 month[57]. Thrive indices were also calculated for cases in this study. The American study by Dykman et al based the definition of FTT on three parameters: weight below the 5th centile according to NCHS standards, low weight-for-height ratio and abnormal weight velocity.

The majority of studies defined FTT according to body weight less than a certain size at one or more points in time during infancy. Kelleher et al and Reif et al both defined FTT as weight less than the 5th centile on at least 2 occasions during infancy and in the case of the Reif study this had to be over a 6 month period[64;67]. The 2003 study by Mackner defined FTT as decline in weight-for-age or weight-for-height from AGA to below the 5th centile at intake to the study[62]. The 1988 study by Drotar et al defined FTT according to decrease in weight gain from birth to below the 5th centile at hospital admission[60]. In the 1997 study by Mackner et al FTT was diagnosed when weight-for-age was less than the 5th centile at a point before 24 months using NCHS standards[61]. The same definition was used in the study by Kerr et al[63]. Similarly, Glaser et al defined FTT as weight below the 3rd centile on admission to hospital at a mean age of 12.5 months[66] and Abramson et al defined FTT as weight below the 3rd centile according to NCHS standards between 6 and 25 months of age[72]. The 1985 study by Drotar et al grouped subjects according to their percentage expected weight-for-height and height-for-age at 12 and 18 months[59]. Mitchell et al defined FTT as weight <80% of normal for age between 0 and 24 months using Harvard growth charts[70].

The studies by Hack et al and Tudehope et al both took a similar approach to assessing growth faltering in VLBW subjects[69;71]. The AGA and SGA subjects were examined separately and in each group a comparison was made between those who had adequate or normal size at 8 months or 1 year and those whose size was subnormal or retarded. In the case of the Hack study, subnormal growth was defined as weight <-2SD using growth charts of Babson and Benda. In the Tudehope study retarded growth at 1 year was defined as weight less than the 3rd centile.

One of the studies did not give a precise definition of FTT. Field et al stated that one of their aims was to define the limits of FTT and so they did not use a definition for FTT which was based on size or growth at particular points in time[68]. Subjects were those that had been admitted to hospital for FTT who had measurements indicating poor weight gain.

3.3B.4 Quality assessment

The majority of studies were prospective cohort studies and a number used a nested case control design where cases were selected on the basis of exposure status (i.e. FTT) rather than outcome. Only three of the studies were population based. The fact that these were all UK studies and that none of the USA studies had a population basis is likely to be due to differing organisation of health services within these settings. There is therefore considerable potential for selection bias in all 6 US studies particularly since many of the subjects were admitted to hospital and therefore represent the more severe end of the spectrum of FTT[59-63;65].

The majority of studies gave adequate information on socio-economic and family characteristics when describing study participants. However mean birthweight and gestational age of subjects were less consistently reported and only 6 studies were considered to have given sufficient description of study participants[55;56;58;64;69;70;72].

In most of the studies measurements of infant size were ascertained from child health clinic or hospital records. Only one study, by Skuse et al, gave a detailed description of how measurements were taken[55]. The information on measurements of the outcome was much more detailed. As stated in section 3.3B.2, all but one of the studies used well-validated scales or measures to assess cognition. However only 7 studies gave sufficient information about observers, their training or inter-observer variability and only these studies were scored as having low risk of bias in this respect[56-59;62;70;72].

The confounding factors we considered important in the relationship between growth faltering and cognitive development were socio-economic status, parental (particularly maternal) education, infant feeding and sex. Only two studies by Drewett et al and Mackner et al were considered to have low risk of bias, both assessing 3 of the 4 factors[58;62]. Eight of the studies were assessed as having medium risk of bias and the majority of these adjusted for socio-economic status and maternal IQ or educational status. Five studies were assessed as having high risk of bias[57;65-67;71]. This included the studies by Glaser et al, Reif et al and Tudehope et al which did not considering any confounding factors in their analyses. Of the 4 studies including an intervention[59;60;62;64], only the 2003 study by Mackner et al adjusted analyses for intervention status of subjects[62].

A number of studies, including the 1977 study by Mackner et al and the later follow up of the same subjects by Kerr et al, did not give sufficient information about the target population to allow adequate assessment of loss to follow up[61;63]. The majority of studies were assessed as having medium risk of bias in this respect with only the study by Skuse et al considered to have low risk of bias[55]. Six studies were considered to have high risk of bias due to loss to follow up although 3 of these by Kelleher et al, Drotar et al and Field et al gave some information on non-participants[60;64;68].

Assessment of overall risk of bias in these 17 studies is summarised in appendix 10, table C. Only the studies by Skuse et al and Drewett et al were

considered to have low risk of bias overall[55;58]. The majority of studies were considered to have medium risk of bias. However the studies by Glaser et al, Tudehope et al and Field al had high risk of bias[66;68;71]. In the case of the study by Glaser et al this was primarily due to poor consideration of confounding and lack of any statistical analysis[66]. This study also included assessment of cognition at ages ranging from 8 months to 11 years which considerably limits the generalisability of findings. The study by Field et al which was based on only 17 subjects did not give a clear definition of FTT and was considered to have high potential for selection bias[68]. The study by Tudehope et al had high attrition rates and poor consideration of confounding.

3.3B.5 Summary of findings

There is considerable consistency in the findings of the 19 studies about the relationship between FTT or growth faltering and cognitive development during infancy and childhood.

Of the 11 studies that examined the association between FTT and cognition up to 3 years of age, 7 demonstrated that FTT was associated with lower levels of cognition. This applied to the UK study by Skuse et al and to the American studies by Singer et al, Kelleher et al, the 1985 study by Drotar et al and the 2003 study by Mackner et al[55;59;62;64;65]. The study of VLBW infants by Hack et al demonstrated that growth faltering was associated with poorer cognition in the AGA infants but not in the SGA. In the AGA infants absence of FTT was associated with significantly higher developmental quotient (DQ) at 8 months.

Of the remaining 4 studies assessing cognition up to 3 years of age, 3 suggested that growth faltering was associated with poorer cognitive development but were inconclusive overall. The 1988 study by Drotar et al showed that age at onset and duration of FTT were associated with cognition, with earlier age at onset and longer duration associated with lower levels of cognition, but that weight-for-height in infancy was not associated with cognition[60]. In the study of 17 FTT infants by Field et al there was a significant correlation between change in weight and change in MDI[68]. However MDI only increased several months after weight increase and given the lack of a clear FTT definition it is difficult to integrate these findings with those of other studies in the review. Abramson et al failed to show an association between FTT and emotional expressivity although FTT infants showed significantly more negative emotions and lower face actions than controls. The remaining study by Mackner et al failed to show a significant difference in MDI in FTT and comparison infants[61].

The studies by Tudehope et al and Mitchell et al, which assessed cognition up to 4 and 6 years respectively, had conflicting findings[70;71]. The study by Mitchell et al, which was based on a normal population of children, failed to show an association between FTT and cognition. However the study by Tudehope et al of VLBW infants revealed that growth faltering was associated with poorer cognition in AGA infants but not SGA infants and in this respect the findings are consistent with those of Hack et al.

Seven studies explored the association between FTT and cognition between 6 and 8 years. Six of these were consistent in showing lower mean IQ in FTT cases than controls at follow-up[56-58;62;63;73] although these findings were only statistically significant in 1 of the studies. The findings of the remaining study by Reif et al were inconclusive; while higher levels of learning difficulties and developmental delay were seen in the FTT children, the findings in relation to other measures of cognition were not significant[67].

A meta-analysis was carried out for the 6 studies for which a mean difference was given between cases and controls for cognitive development in childhood (6 to 12 years). The combined estimate, using a fixed-effects model and inverse-variance weighting, was -3.4 (95% CI -5.3 to -1.6; $p < 0.001$). The forest plot, shown in Figure 1, suggests that the studies are homogeneous apart from that by Dykman[73]. This could be due to the fact that the subjects in this study were aged 8 to 12 years (mean 10 years), rather than 6 to 8 years for the other studies. It also has a small sample size (27 cases and 17 controls), so could potentially be less reliable. However, Figure 5 suggests that the Dykman study does not significantly affect the combined estimate (removing it adjusts the combined estimate to -2.9, with a 95% CI of -4.9 to -1.0). The I^2 statistic, interpretable as the proportion of total variation in the estimate of size of effect due to heterogeneity between studies, was 3.5% ($p = 0.959$), so heterogeneity was negligible. Thus, children whose growth faltered in infancy can expect to score an average of 3.4 IQ points lower than those who did not suffer growth faltering in infancy, before adjustment for potential confounders.

Figure 1: Forest plot of mean differences in childhood IQ between cases of infant growth faltering and control infants, with combined estimate
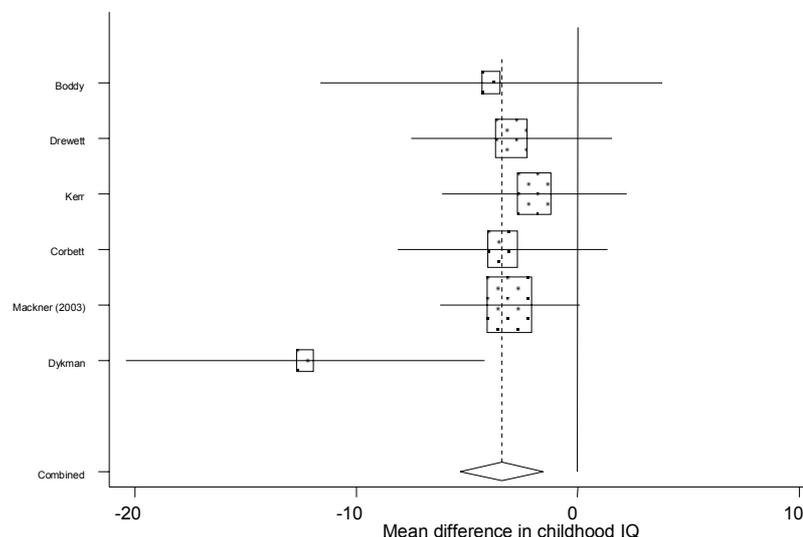
In summary, the findings of this review suggest that FTT is associated with poorer levels of cognition up to the age of 3 years. The findings of individual studies of later cognition between 6 and 8 years suggest that this effect reduces with age but that mean IQ of FTT infants remains lower. This is confirmed by the findings of a meta-analysis which demonstrated a modest reduction in IQ in the FTT children compared with those who did not suffer growth faltering during infancy.

3.3C Studies relating infant size/growth in weight or height to cognitive development

3.3C.1 Selection of studies

This section describes the findings of 9 studies that explore the relationship between size /growth in infant height or weight and later cognitive development. The studies are summarised in table 12.

Table 12: Summary of included studies in review of infant size/growth in height or weight and cognitive development

Ref. no	Title	Author (year)	Design
74	Effects of stunting on cognition in late childhood	Berkman (2002)	Prospective cohort
75	Malnutrition and mental development	Monckeberg (1972)	Cross-sectional
76	Postnatal growth and psychomotor development in small for gestational age Brazilian infants	Paine (1984)	Prospective cohort
77	Neurological follow-up of small-for gestational age newborn infants	Gherpelli (1993)	Prospective cohort
78	Growth achievement in LBW premature infants: relationship to neurobehavioral outcome at one year	Ross (1983)	Prospective cohort
79	Relation between physical growth and information processing in infants born in India	Rose (1994)	Cross-sectional
80	Weight gain in infancy and educational attainment at ten	Corbett (2004) (unpublished)	Prospective cohort
81	Infant predictors of cognitive development in an undernourished Kenyan population	Whaley (1998)	Longitudinal cohort
82	Postnatal growth in VLBW infants: significant association with neurodevelopmental outcome	Latal-Hajnal (2003)	Prospective cohort
83	Growth and behaviour development in rural infants in relation to malnutrition and environment	Upadhyay (1992)	Prospective cohort

Tables in appendix 10 relating to height and weight and cognitive development are tables 3.3C (i) and (ii).

3.3C.2 Description of studies

Seven of the studies were set in developing countries: 2 in Brazil, 1 in Chile, 2 in India, 1 in Peru and 1 in Kenya[74-77;79;81;83]. The other 3 studies had developed country settings in the UK, USA and Switzerland[78;80;82]. Eight of the studies were prospective cohort studies and 2 were cross-sectional.

Paine et al followed up 80 infants born between 1975-78 in a single hospital in Brasilia, 36% of which were SGA with mean birthweight below 2.5kg[76]. The other study set in Brazil by Gherpelli et al followed up 37 SGA subjects with mean birthweight 2.02 kg born between 1984-86 until 1 year of age[77]. The Indian study by Rose et al recruited 183 infants with mean birthweight of 2.76 kg from a well-baby clinic and assessed the relationship between size and cognition when the infants were between 5 and 12 months old[79]. The second

Indian study by Upadyhyay et al assessed the relationship between degree of malnutrition (graded from mild to severe) and cognitive development in 224 infants born between 1981 and 1983 in 10 rural Indian villages[83]. Monckeberg et al carried out a cross-sectional study of 110 infants randomly sampled from a slum community and assessed the relationship between level of nourishment (assessed by weight-for-age) and deficits in height and HC from 1-3 years and simultaneous levels of development[75]. The mean birthweight of subjects in this study was not reported. The study set in Peru by Berkman et al was a prospective follow up of children, some of whom had been stunted in infancy, to assess their cognitive development at 9 years of age[74]. Whaley et al assessed the cognitive development at 30 months of a cohort of 132 infants born in a defined geographical region of Kenya who had been measured at monthly intervals until 6 months of age[81]. The mean birthweight z-score for the study population was –0.23 (SD 0.98).

In the UK study by Corbett et al, 1724 infants born during 1987-88 in Newcastle-upon-Tyne were followed up until 10 years of age to explore the relationship between infant size and growth and later educational achievement[80]. The 2 other developed country studies were both based on cohorts of VLBW (<1.5kg) babies. Latal-Hajnal et al followed up 219 infants treated in a single maternity hospital who were born in 3 phases between 1983 and 1994 and assessed the relationship between size and cognition at 2 years of age[82]. Ross et al studied 86 preterm (<29 weeks) infants admitted to a single perinatology department between 1978 and 1979. Their size at 3-monthly intervals throughout the first year of life was related to cognition at 1 year of age[78].

Cognition was assessed at a range of ages with a variety of scales. Six of the studies assessed cognitive development at 1 year of age. The studies by Monckeberg et al and Paine et al both used Gesell scales[75;76]. In the study by Paine et al infants were also assessed at 4, 8, and 18 months and in the study by Moncekberg infants were assessed between 1 and 3 years[75;76]. The study by Gherpelli et al used Denver developmental scales and Ross et al used Bayley scales of infant development[77;78]. Rose et al assessed visual recognition memory between 5 and 12 months[79]. In the other Indian study, DQ was measured at 4, 16, 28, 40 and 52 weeks[83]. The studies by Latal-Hajnal et al and Whaley et al both used Bayley scales at 24 and 30 months of age respectively[81;82]. Two of the studies explored cognitive development in later childhood: Berkman et al used WISC-R to derive IQ at 9 years of age and Corbett et al looked at level of educational attainment at 10 years of age using 4 well-validated tests (picture vocabulary, problems of position, reading and maths)[74;80].

3.3C.3 Assessment of infant size/growth

Nine of the studies explored the association between infant size at various points during the first 2 years of life and cognitive development. Only the study by Corbett et al assessed infant growth.

Latal-Hajnal et al and Berkman et al focused on size (height, weight and HC) at 2 years, and in the study by Berkman et al these measurements were used

to assess which infants were wasted or stunted. Wasting was defined as weight-for-age z score less than -2 and stunting as height-for-age z score less than -2 using WHO reference data. Weight at 1 year was assessed in the Brazilian study by Gherpelli et al and also the American study by Ross et al which measured infants at 3-monthly intervals up to the age of 1 years. The other Brazilian study by Paine et al focused on height at 4, 8, 12 and 18 months. In the study by Monckeberg et al, subjects' height, weight and HC were measured between 1 and 3 years of age but no mean age was reported. Upadhyay et al assessed weight for age at regular intervals during the first 52 weeks of life[83]. In the Kenyan study by Whaley et al infants' height, weight and arm circumferences were measured at monthly intervals up to 6 months[81]. Rose et al in the Indian study measured height, weight and HC between 5 and 12 months of age[79].

In the UK study by Corbett et al, weight measurements throughout the first 2 years of life were converted to standard deviation (SD) scores using 1990 UK growth standards[80]. Weight gain in late infancy was assessed where the late infancy measurement was the latest available weight SDS between 9 and 24 months.

3.3C.4 Quality assessment

The majority of studies were prospective cohort studies though few were population based. The studies by Paine et al, Gherpelli et al, Ross et al and Latal-Hajnal were all based on subjects recruited from single hospitals[76-78;82]. Similarly Rose et al recruited subjects from a single well-baby clinic[79]. The study by Upadhyay et al was a cross sectional analysis of malnutrition and development in infants taking part in a prospective cohort study. All 6 studies were assessed as having medium risk of bias in relation to study design. The Kenyan study by Whaley et al recruited subjects from a defined geographical area but insufficient information is given selection of subjects to be sure that the sample was population based[81]. Monckeberg et al randomly sampled subjects from a slum community in Santiago but gave little information about how the middle-class comparison group were recruited[75]. Only the studies by Berkman et al and Corbett et al which were both population based were assessed as having low risk of bias in relation to study design[74;80].

Good baseline information about age, sex, mean birthweight and gestational age of subjects was given in 5 of the studies, consequently these were assessed as having low risk of bias in this respect[76-79;81]. Of the other studies, Berkman et al gave detailed information on socio-economic variables but did not report mean birthweight or gestational age. The same was true for Monckeberg et al and Whaley et al[75;81]. Upadhyay et al reported socio-economic status (SES) and birthweight[83]. Corbett et al gave information on birthweight, gestational age and SES but nothing on sex or race. These studies were assessed as having medium risk of bias in this respect[80].

Two studies gave detailed information on observers, techniques and reliability of infancy measurements and both were assessed as having low risk of bias in this respect[76;78]. All but one of the other 7 studies were assessed as having medium risk of bias since most gave information on techniques of observers

but not both. The study by Monckeberg was assessed as having high risk of bias because it was not clear how infant size measurements had been ascertained[75.]

Seven of the studies had low risk of bias with respect to outcome assessment since they gave sufficient information about observers and used well-validated scales and measures to assess cognition[74;76-79;81;82]. In addition observers were blind to infant growth status in 5 of these studies[74;76;78;79;82]. Only the studies by Monckeberg et al, Upadhyay et al and Corbett el al were assessed as having medium risk of bias in relation to outcome assessment and in all cases this was due to lack of information relating to observers and reliability[75;80;83].

As before, the confounding factors we considered important were socio-economic status, parental (especially maternal) education, sex and infant feeding. Only the study by Rose et al based in India adjusted for all of these factors and was assessed as having low risk of bias[79]. Two studies by Monckeberg et al and Gherpelli et al did not adequately consider any of these factors and were assessed as high risk of bias[75;77]. The remaining studies assessed some of these factors and most adjusted for socio-economic status.

There were high levels of attrition in many of the studies with 6 assessed as having high risk of bias in this respect although this was due to lack of denominator information in the studies by Paine et al and Monckeberg et al[75-78;80;81]. Two of these studies by Paine et al and Gherpelli et al also gave insufficient information relating to non-participants. Only 47% of the study population were followed up in the study by Corbett et al but the authors demonstrated that subjects and non-participants had similar size in infancy. Only the studies by Rose et al and Latal-Hajnal et al were assessed as having low risk of bias. In the case of Rose et al the author states that virtually all subjects were followed up although the exact percentage follow-up is not given[79]. In the study by Latal-Hajnal et al 94% of subjects were followed up at 2 years of age[82].

Assessment of overall risk of bias is summarised in appendix 10, table C. Only the studies by Rose et al and Latal-Hajnal et al were considered to have low risk of bias[79;82]. All but two of the other studies were considered to have medium risk of bias mostly due to weaknesses in consideration of confounding and losses to follow-up. The studies by Monckeberg et al and Upadhyay et al were assessed as having high risk of bias overall. In the case of Monckeberg et al, this was primarily due to poor reporting of methods and lack of consideration of confounding[75]. The study by Upadhyay et al had high loss to follow-up[83].

3.3C.5 Summary of findings

The findings of the majority of studies suggested a positive association between height and weight in infancy and later cognition. Monckeberg et al and Paine et al which both used Gesell development scales demonstrated a positive association between developmental quotient (DQ) and infant size although in the study by Paine et al a significant association was only observed in the SGA infants who accounted for 36% of the study population and not in the AGA infants[75;76]. Gherpelli et al demonstrated that infants with weight <2.5th centile at 1 year of age were more likely to experience abnormal development[77]. In the Indian study by Rose et al infants of normal weight had significantly better visual recognition and cross-modal transfer scores between 5 and 12 months than those of low weight (defined as <-1SD below the mean)[79]. Ross et al in the study of VLBW American infants demonstrated a positive association between height, weight and HC at 3, 6, 9 and 12 months and 1-year cognition[78]. This was observed in both boys and girls.

The findings of the studies by Whaley et al and Latal-Hajnal et al were less conclusive[81;82]. There was no association between weight-for-age and height-for-age and 30 month cognition in the study by Whaley et al but arm circumference between 0-6 and 3-4 months was positively associated with cognition. In the study by Latal-Hajnal et al, mean MDI was significantly higher in the normal than low weight group but this was only significant in the subjects who were normal birthweight for gestational age[82]. Similar trends in MDI were observed for height but only in the low birthweight (SGA). In a regression analysis which adjusted for sex, gestational age and SES, there was no significant association. Similarly the study by Upadhyay et al showed no association between degree of malnutrition up to 52 weeks of age and cognition after adjustment for socio-economic factors including parental occupation and education[83].

The two studies of cognition in later childhood also suggested a positive association between infant size and cognition. Berkman et al demonstrated poorer cognition was associated with severe stunting compared with less severe and no stunting, and with longer-term stunting (>12 months)[74]. Corbett et al demonstrated that weight gain in infancy was associated with one aspect of educational attainment at 10 years. However findings in relation to the other 3 measures were not significant[80].

Only two of the studies that demonstrated a positive association between size in infancy and later cognition reported estimates of the size of effect. In the study by Berkman et al, in an adjusted regression analysis the coefficient (95% CI) for change in 9-year WISC score for severe stunting of 12 months or over duration was –10.0 (-17.5, -2.4)[74]. In the study by Corbett et al, the regression coefficient for change in the picture vocabulary test scores according to a unit change in late infancy weight SD score was 0.68 (0.1, 1.1)[80].

Overall, all but 2 studies reported a positive association between infant size and cognitive development up to 9 years of age. However, although the findings of one of the studies suggest that weight gain in late infancy is positively associated with cognition, there is insufficient evidence to draw any

conclusions relating to infant growth. As outlined in section 3.3B.1, there is some overlap between these studies and those relating to FTT. It is therefore reassuring that findings are consistent across studies. Lower weight gain in the normal range and growth faltering were consistently associated with poorer cognitive development.

3.3D   Motor development
Tables in appendix 10 relating to motor development are tables 3.3D (i) and (ii).

3.3D.1   Selection of studies
The same abstracts screened for cognitive development were also assessed for motor development (page 36, section 3.3A.1). This section describes the findings and quality of 10 studies exploring the association between infant size or growth and motor development during childhood. The titles and authors of the studies are summarised in table 13 below.

Table 13: Summary of included studies in review of infant size/growth and motor development

Ref no	Title	Author (year)	Design
38	Sequelae of growth failure in appropriate for gestational age, VLBW infants	Astbury (1986)	Prospective cohort
84	Influence of growth on development outcome in extremely LBW infants at 2 years of age	Connors (1999)	Prospective cohort
85	The motor development of fat babies	Jaffe (1982)	Cross-sectional
86	Outcome of VLBW infants at 1 and 2 years of age	Kohlhauser (2000)	Prospective cohort
81	Infant predictors of cognitive development in an undernourished Kenyan population	Whaley (1998)	Longitudinal cohort
82	Postnatal growth in VLBW infants: significant association with neurodevelopmental outcome	Latal-Hajnal (2003)	Prospective cohort
37	Developmental outcome of small for gestational age children at 2 years	Tenovuo (1988)	Prospective cohort
54	Effect of VLBW and subnormal head size on cognitive abilities at school age	Hack (1991)	Prospective cohort
87	Catch-up head growth and motor performance in VLBW infants	Simon (1993)	Prospective cohort
88	Fetal growth, early postnatal growth and motor development in Pakistani infants	Cheung (2001)	Prospective cohort

3.3D.2   Description of studies
Of the 10 studies of motor development, 6 were based on populations of LBW or VLBW babies[38;54;82;84;86;87] and 1 on SGA infants[37] while 3 were based on normal populations of children[81;85;88]. All but one were prospective cohort studies while the remaining study by Jaffe et al was a cross-sectional study of 135 infants aged 6-18 months who were attendees at infant welfare clinics in Haifa, Israel[85].

The 6 studies of LBW or VLBW infants all took place in developed countries. The studies by Astbury et al and Connors et al were based in Australia. Astbury et al studied the 2-year motor development of 235 AGA infants with birthweights of less than 1.5kg who were discharged from NICU between 1977 and 1980[38]. Connors et al followed up 226 ELBW preterm infants with birthweight under1kg and gestational age less than 28 weeks discharged between 1987-92 from a single NICU[84]. The American study by Hack et al was a later follow-up at 8.6 years of the cohort of 249 VLBW (<1.5kg) infants described in the section on cognitive development[54]. Kohlhauser et al assessed the motor development between 1 and 2 years of 76 VLBW preterm infants treated in a single NICU in Vienna, Austria[86]. In the most recent study by Latal-Hajnal et al 219 VLBW infants were recruited from a single maternity hospital in 3 phases between 1983 and 1994[82]. The study by Simon et al was based on 48 VLBW infants discharged from a single NICU in Alabama, USA[87].

The study of SGA infants by Tenuvuo et al was based on 463 infants with birthweight below the 10th centile for gestational age who were born in Turku, Finland[37]. A minority (6%) of subjects in this study were preterm.

The 3 studies based on normal populations of children were set in developing countries. One was the Israeli study by Jaffe et al which is described above[85]. The study by Whaley et al was based on 132 infants born during a 17-month period in a defined geographical region of Kenya[81]. The third developing country study by Cheung et al was based on 1476 infants born between 1984 and 1987 in 4 areas of Lahore, Pakistan[88].

Five of the studies assessed motor development at 2 years of age[37;38;82;84;86]. All used well-validated methods although the choice of measurement scale varied between studies. The studies by Astbury et al and Latal-Hajnal et al used Bayley scales of infant development. Connors et al used the Neurosensory Motor Development Assessment (NSMDA). Kohlhauser et al used Griffiths scales and Tenuvuo et al used Denver scales of infant development which includes items relating to gross and fine motor development (walking, eating, dressing) as well as aspects of sensory and cognitive development. The other developed country study by Simon et al also assessed motor development at 1 year of age using the Denver developmental test.

Two of the developing country studies assessed motor development from 6 months of age[81;85]. The Israeli infants were followed up between 6 and 18 months and assessed with the Sheridan Stycar test. The Kenyan infants were followed up twice at 6 months and 30 months using Bayley scales. The study by Cheung et al assessed the gross and fine motor developmental milestones of age at independent walking and age at building a 3-brick tower[88].

Hack et al assessed the visual motor and fine motor skills of their subjects at a mean age of 8.6 years using Bender Gestalt and Pergue Pegboard tests[54].

3.3D.3 Assessment of infant size/growth

Six of the studies assessed the relationship between head size in infancy and motor development[37;38;54;85-87]. The studies by Tenuvuo et al, Astbury et al and Latal-Hajnal et al focused on head circumference (HC) at 2 years of age and Kolhauser at 1 year of age. Hack et al measured HC at both 8 months and 20 months. Simon et al at 3, 6 and 12 months after discharge from NICU and catch-up head growth was defined as HC recovering to 5th centile according to NCHS growth charts[87]. Astbury et al also assessed growth in height and weight up to 2 years of age as well as head growth. Latal-Hajnal et al also examine the relationship between height and weight at 2 years and motor development.

Five of the studies explored the association between weight or height and later motor development[38;81;84-86]. Two of these are the studies by Astbury et al and Latal-Hajnal et al which both focused on 2-year height and weight as well as 2-year HC. In addition, Connors et al looked at the relationship between 2-year weight and motor development. Whaley et al, in their study of Kenyan infants, measured weight, height and mid-upper arm circumference (MUAC) at monthly intervals up to the age of six months and then explored the relationship between growth at a number of intervals within this 6-month period to motor development at 6 and 30 months of age. The study of Israeli infants by Jaffe et al assessed the size (weight and height) of infants between 6 and 18 months. These were then used to derive an index of overweight and obesity (Sveger's Index which is based on expected weight-for-age /expected height-for-age) where overweight was defined as an index of 111-120 and obesity as greater than120. The study by Cheung et al assessed growth in height and weight which were standardised according to WHO international reference standards[88]. Weight-for-length SDS was used as a measure of thinness (or wasting) and changes in length SDS and weight-for length SDS from 1.5 to 6 months of age were used as indicators of postnatal growth.

3.3D.4 Quality assessment

All but one of the studies were prospective cohort studies. However, only the studies by Tenuvuo et al and Cheung et al included population-based sample of subjects[37;88]. The studies of VLBW infants all selected subjects from single hospitals or NICU. Whaley et al selected subjects from 3 sub-areas of a particular geographical region in Kenya. However the sampling approach was unclear.

The cross-sectional study by Jaffe et al was population based[85]. Subjects were randomly sampled from infant welfare clinics in Haifa. Only the studies by Jaffe et al and Tenuvuo were assessed as having low risk of bias in relation to study design. The other 7 studies were assessed as medium risk of bias in this respect.

Four of the studies gave good baseline information relating to participants including birthweight, gestational age, gender and socio-economic status[82;84;86;88]. These studies by Connors et al, Kohlhauser et al, Cheung et al and Latal-Hajnal et al were assessed as having low risk of bias in this respect.

In the other 6 studies information was less complete and they were assessed as having medium risk of bias.

Information on infancy measurements was only sufficient in the study by Cheung et al[88]. Most of the other studies gave information on either procedures used or on observers but none on both and all were scored as medium risk of bias whereas the study by Cheung et al was scored low risk of bias. None of the studies gave information on reliability of measurements. Rounding was not considered to be a source of bias in the majority of studies.

Outcome measurement was described in more detail. As stated in section 3.3D.2 all of the studies used well-validated methods to assess motor development. Six of the studies also included sufficient information in relation to observers and reliability to be assessed as having low risk of bias[38;81;82;85;87;88]. However the information on observers in the studies by Connors et al, Kohlhauser et al, Tenuvuo et al and Hack et al was insufficient and so they were assessed as having medium risk of bias in relation to outcome assessment. Observers were blind to infant size/growth in only 3 of the studies.

The confounding factors we considered important in the relationship between infant size/growth and motor development were socio-economic status, parental (particularly maternal) education, infant feeding and sex. Four studies by Astbury et al, Jaffe et al, Simon et al and Tenuvuo et al were poor in consideration of confounding and did not adjust for any of these confounding factors in their analyses[37;38;85;87]. The other 6 studies adjusted for some but not all factors and all were assessed as having medium risk of bias in this respect.

Response rates were above 90% in the studies by Astbury et al, Jaffe et al, Whaley et al and Latal-Hajnal et al and these studies were therefore assessed as having low risk of selection bias[38;81;82;85]. The other 6 studies had follow-up rates between 70 and 90% and were assessed as having medium risk of bias. However the study by Hack et al demonstrated that non-participants were similar to subjects in terms of infant size and growth thus reducing the level of bias due to attrition[54].

Assessment of overall risk of bias in these 10 studies is summarised in appendix 10, table C. Only the studies by Latal-Hajnal et al, Cheung et al and Hack et al were assessed as having low risk of bias overall[54;82;88]. The other 7 studies were all assessed as having medium risk of bias.

3.3D.5 Summary of findings

Four of the 5 studies that considered head size in infancy demonstrated a positive relationship between HC at 1 or 2 years of age and motor development. Kohlhauser et al demonstrated that HC <3rd centile at 1 year was a risk factor for poor developmental outcomes at 1-2 years of age[86]. In the study by Tenuvuo et al, children found to have abnormal development at 2 years were significantly more likely to have HC <25th centile at 2 years than those with normal development[37]. No estimates of effect size are given in the analyses by Kolhauser et al and Tenuvuo et al. Latal-Hajnal et al

demonstrated that mean PDI was significantly lower in infants whose HC was <10th centile[82]. However these findings were only significant in the infants who had had normal birthweight for gestational age. In an adjusted logistic regression analysis the odds of having 2-year PDI<84 was significantly greater in subjects with low 2-years weight compared with normal 2-year weight (odds ratio=5.77, 95% CI1.15, 29.0). The fourth study by Simon et al assessed the relationship of catch-up head growth by 6 months of age with motor performance at 1 year of age[87]. Mean catch-up time and % catch-up by 6 months were both positively associated with motor performance at 1 year of age. The fifth study considering head size in infancy by Hack et al did not show any association between 8-month HC and visual motor and fine motor development at 8.6 years[54].

The findings in relation to height and weight in infancy were less conclusive. The study by Connors et al demonstrated that lower weight at 2 years was associated with poorer motor development : odds rato for poor development when 2-years weight was below the 10th centile was 3.4 (95% CI 1.1, 11.2)[84]. However this association became non-significant after adjustment for birthweight. In the study by Whaley et al weight-for-age and height-for-age were not associated with motor development at 6 months of age, although the findings in relation to MUAC suggested a positive association between growth in upper arm circumference between 5 and 6 months and 6-month outcome[81]. A positive association between 2-year weight and 2-year fine and gross motor development was demonstrated in the study of VLBW Australian infants by Astbury et al with significantly poorer developmental scores in the infants with 2-years weight less than the 10th centile[38]. The study by Jaffe et al focused on infants who were overweight or obese[85]. A significantly higher percentage of overweight or obese infants had delayed motor development at 6-18 months compared with non-obese infants. The study by Cheung et al demonstrated that postnatal wasting and stunting both led to poorer motor development[88]. These findings were seen across the range of infant size and were not just restricted to babies who were small at birth.

In summary, smaller head size (below 25th centile) during infancy was associated was associated with poorer motor development in studies that were predominantly of LBW or VLBW infants. However the evidence relating to height or weight in infancy was inconclusive.

4. Results relating to outcomes in adolescence

4.1 Mortality and health-related quality of life
Of the 14 outcomes relating to mortality and morbidity in adolescence, eligible studies were identified for only 2. These were insulin dependent diabetes (IDDM) and mental illness. The studies of IDDM were based on adolescent subjects although we recognise that IDDM also presents a burden of disease in both childhood and adulthood. Only 3 studies, relating to obesity, were identified for outcomes in the category health-related behaviours. These are considered alongside the studies in adulthood in section 5.2A.

No studies were identified in relation to non-health related quality of life in adolescence.

4.1A Insulin dependent diabetes

Tables in appendix 10 relating to insulin dependent diabetes mellitus (IDDM) are tables 4.1A (i) and (ii).

4.1A.1 Selection of studies

Searches of Medline and Embase identified 6524 references, whilst 3 further references were identified through contact with experts. Screening of titles and abstracts led to identification of 13 studies that would potentially meet inclusion criteria for the review, and one study was identified from the reference list of another study. Two further studies were identified through correspondence with experts. Of these, 6 were subsequently ruled out. Four of these did not include measurements during infancy, or measurements taken between 3 months and 2 years could not be isolated from those of a larger age group. A further study only considered a single case, and another had no variability in the outcome. Of the 10 papers thus meeting the selection criteria for the review, seven concerned IDDM, and these are summarised in Table 14. All studies related to adolescent subjects (below the age of 18 years).

Table 14: Summary of included studies in review of infant growth status and IDDM

Ref. no	Title	Author (year)	Design
89	A high weight gain early in life is associated with an increased risk of type 1 (insulin-dependent) diabetes mellitus	Johansson (1994)	Case-control
90	Infant feeding, early weight gain, and risk of type 1 diabetes. Childhood Diabetes in Finland (DiMe) Study Group	Hypponen (1999)	Case-control
91	Obesity, increased linear growth, and risk of type 1 diabetes in children	Hypponen (2000)	Case-control
92	Weight gain in infancy and subsequent development of diabetes mellitus in childhood	Baum (1975)	Case-control
93	Stature at time of diagnosis of type 1 diabetes mellitus	DiLiberti (2002)	Cross-sectional
94	Height at onset of insulin-dependent diabetes in children in southern India	Ramachandran (1994)	Cross-sectional
95	Association between infant growth before onset of juvenile type-1 diabetes and autoantibodies to IA-2. Netherlands Kolibrie study group of childhood diabetes	Bruining (2000)	Case-control

4.1A.2 Description of studies

Five of the studies used a case-control design and assessed the relationship between size or growth in infancy and risk of developing IDDM at some point in time[89-92;95]. Of these, one was set in a single region of south-east Sweden and targeted cases of diabetes diagnosed between 1974 and 1988[89], whilst two used data from the same Finnish source population, which targeted cases diagnosed between late 1986 and early 1989[90;91]. The remaining study

targeted cases diagnosed between 1995 and 1998 in the south-west of the Netherlands[95]. All four of these only included cases diagnosed before the age of 15 years. The final case-control study involved cases attending a single paediatric clinic in Oxford, but little further information is provided on the study setting or participants[92].

A further 2 studies that were cross-sectional in design, determined the mean size of subjects at onset of diabetes and comparing it with that of a control population at the same age[93;94]. In both studies, the cases were stratified by age, meaning that those subjects whose onset was before 2 years of age, and who were therefore of relevance to the review, could be isolated. One study involved cases attending a single paediatric clinic in Wisconsin, USA[93], whilst the other involved cases from a diabetes research centre in Madras, India[94]. Both used national reference populations unrelated to the study (though of the same age and sex) for the control population. The findings of these studies will be influenced by their cross-sectional nature since infant size may well have been influenced by the development of diabetes in subjects.

All of the studies included males and female subjects in roughly equal proportions.

4.1A.3   Assessment of infant size/growth

Three of the 7 studies reported size in terms of weight at various ages between 3 months and 2 years[90-92], whilst 2 further studies reported infant size in terms of height[93;94]. Two studies reported findings relating to infant growth. One such study measured growth in height and BMI over the first and second years of life[95], whilst the other reported growth in weight and height between birth and various ages ranging from 3 to 30 months[89]. In both cases, size measurements were converted to standard deviation (SD) scores using growth charts, and growth was expressed as the change in SD score over the period of interest.

4.1A.4   Quality assessment

Only the two Finnish studies had a design that was considered wholly appropriate. They targeted all cases of diabetes in the country, and obtained age- and sex-matched controls from the national population register, thus ensuring that the source population of cases and controls was as similar as possible[90;91]. The Swedish study also made use of the national population register to select controls, and furthermore matched them on geographical region to ensure similarity. However, the cases were only selected from 5 paediatric clinics in one part of the country, potentially introducing bias[89]. The infant heights or weights of both cases and controls were ascertained from child health records in all three studies, which should be a reliable source[89-91].

The two cross-sectional studies used cases from single institutions whilst using national reference populations as the control group[93;94]. Since the sources of cases were so specific, and the source population of cases and controls markedly different, there was a high risk of bias. However, infant size in both studies was assessed either by the research team or by trained staff. The controls for the Wisconsin study were drawn from participants in a large-scale

national survey (NHANES), and thus it is reasonable to assume that their measurements were reliably ascertained[93]. The Madras study used data from height charts available from the Indian Council of Medical Research, making them slightly more open to risk of bias[94]. A further potential source of bias with both these studies arises from the fact that, since they are cross-sectional, they only consider cases with extremely early onset of disease and thus may not wait long enough for the disease to occur.

The Oxford and Netherlands case-control studies were considered to have a high risk of bias, principally for not providing information demonstrating whether the study subjects were representative of their source populations, of having small sample sizes and high attrition rates, and not considering the effect of potential confounding factors[92;95].

None of the studies gave an explicit definition of their outcome except to state that it was diagnosed. However, it can reasonably be assumed that the diagnosis was made by a medical professional and is accurate. Two studies stated that the diagnosis was confirmed by medical tests carried out by the research team[91;94], and 2 further studies defined age at onset of disease by the timing of the subject's first insulin injection[89;93].

The confounding factors we considered important in the relationship between infant growth and IDDM were socio-economic status, infant feeding and sex. In assessing quality, we considered matching of cases and controls in relation to confounding factors (with appropriate analysis) or adjustment in statistical analyses to be appropriate means of controlling for confounders. All of the studies either matched for sex or reported results for each sex separately. Only one study adjusted for infant feeding, and then only for the time at which formula milk feeding was introduced[90]; however, one other study compared growth in breastfed compared with non-breastfed subjects without adjustment[89]. A further study adjusted for socio-economic status *post hoc*[91]. In general, the level of consideration of the potential for confounding was poor among these studies.

4.1A.5   Summary of findings

The two Finnish studies, which had the lowest risk of bias and were also among the largest of the diabetes studies considered here, both suggested a positive association between infant size and risk of IDDM. The 1999 study suggested that the association was strongest at 3 months (odds ratio for later diabetes among those with weight in the highest quartile compared to the others: 1.48; 95% CI 1.1 to 2.1) – the association is barely significant at 6 months and non-significant at 9 months. The authors state that there was no evidence of a difference in risk by sex, and the results were adjusted for time of introduction of formula feeding[90]. The 2000 study suggests that the size of the association increases from birth to peak at around 12-18 months before decreasing (for example, the odds ratio for risk of later diabetes associated with a 10-unit increase in weight relative to the mean for age and sex at 12 months, adjusted for sex, was 1.67; 95% CI 1.37, 2.03). The authors state that *post hoc* adjustment for socio-economic status did not alter the result[91].

The two Finnish studies described above considered both sexes together. The studies that give results for males and females separately were less conclusive. The Oxford study, for example, reported that the weight difference between cases and controls was significant only among males at 6 months (cases weighed 0.61kg more than controls; $p < 0.05$), but was significant only among females at 12 months (cases weighed 0.82kg more than controls; $p < 0.05$)[92]. By contrast, the Wisconsin study suggested that, when both sexes are considered together, infant size was *inversely* related to prevalence of diabetes at 1 year (mean *z*-score of cases is 1.00 units less than that of controls; $p=0.001$), but that the effect became non-significant by age 2 years and reversed itself by mid-to-late childhood. This effect, however, seemed to be limited to males – female cases at 1 year weighed *more* than controls, but the difference was non-significant[93]. This anomalous result may be due to the cross-sectional design, since it is otherwise a good study, and its analogous results in later childhood suggest the positive association also seen in the Finnish studies. The study by Ramachandran suggested that cases weigh more than controls at both 1 and 2 years, with the difference being greater among females, but no statistical analysis was carried out[94].

The 2 studies of infant growth both suggested that later diabetics grow faster in infancy than controls. When both sexes are considered together, the Swedish study reported that the cases had significantly increased infant growth compared to the controls over nearly all time periods (for example, from birth to 18 months the growth in weight was 0.24 SDs greater among cases than controls, with a 95% CI of 0.05 to 0.43), but these differences became non-significant among males[89]. The Netherlands study suggested that the BMI of cases increased more quickly than that of controls during the first year, and was higher at 1 year, but that size and growth in height was similar in cases and controls. However, the results suggested that the height of the cases and controls increased more quickly after 1 year than would be expected from a reference population, suggesting that the study sample was not representative of the wider population.

In conclusion, taking into account the relative sizes and risks of bias posed by the studies, infant size and growth appear to be positively associated with risk of IDDM. This association may be stronger among females than males. Studies from the UK, Europe, the US and a developing nation (India) are included, and only one study (the Wisconsin study[93]) reported a result that was inconsistent with those of the others. The generalisability of these findings, summarised in Table 15, should therefore be fairly good.

Table 15: Summary of IDDM findings

Direction of relationship with infant size	No of studies (ref. nos. in parentheses)			
	Males	Females	Both	Total
Inverse	1 [93]	--	--	1
None	1 [89]	1 [93]	--	2
Positive	--	1 [89]	3 [90;91;95]	4
Inconclusive	--	--	2 [92;94]	2
Total	2	2	5	7*

*This figure is the total number of studies examined in the table, and may not equal the sum of the subtotals due to studies appearing more than once in certain cells of the table.

4.1B   Mental Illness

4.1B.1   Selection of studies
Electronic searches identified 13,101 references. Screening of titles and abstracts led to the identification of 15 studies that were considered to potentially meet inclusion criteria for the review, although 13 of these related to outcomes other than mental illness. Following assessment only 1 study which related to adult mental illness was found to meet inclusion criteria.   The only study relating to adolescent mental health is as yet unpublished and was identified through correspondence with first authors of included studies.

4.1B.2   Description of studies
The study presented in table 16 was based on a cohort of births in Newcastle-upon-Tyne in 1987-88.

Table 16: Summary of included studies in review of infant growth status and mental illness

Ref. no	Title	Author (year)	Design
[96]	Physical and emotional development, appetite and body image in adolescents who failed to thrive as infants	Drewett (2004) (unpublished)	Prospective cohort

The study sample was all cohort members who suffered growth faltering in infancy and an equal number of comparison subjects with normal infant growth. The aim was to explore the relationship between growth faltering in infancy and aspects of self-image and emotional development in early adolescence (12 years of age), including appetite and eating style, body satisfaction, self-perception, mood and anxiety.

Tables in appendix 10 relating to mental illness are tables 4.1B (i) and (ii).

4.1B.3   Assessment of infant size/growth
The study assessed growth faltering in infancy, although the phrase 'failure to thrive' is used instead[96]. This was defined according to a conditional weight gain criterion described in detail in another study included in the review[58]. A 'thrive index' was measured at 3-monthly intervals, where the index at a particular point was calculated by regression of current weight on a baseline

weight. Infants were considered to be failing to thrive when the index was below the 5th centile on at least 2 occasions between 3 and 18 months.

4.1B.4 Quality assessment

The study was based on a large, city-wide birth cohort, but only those cohort members suffering from growth faltering in infancy, along with a comparison group of normally-growing infants, were selected. Thus, data on the remaining cohort members was discarded, which could introduce selection bias. This study was not considered to be of case-control design since the 'cases' and 'controls' were defined in infancy on the basis of exposure and followed up prospectively, rather than being defined on the basis of outcome status and looked at retrospectively.

Infant size measurements were obtained from baby clinic records but no information is given relating to observers. Outcome status was reliably ascertained through use of well-validated scales administered by trained nurses. Losses to follow-up were high (65% for cases of growth faltering; 68% for comparison subjects).

The confounding factor we considered important in relation to mental health was socio-economic status, and alcohol/drug intake was judged to be relevant in terms of effect modification. Regression analyses were employed for some outcomes, but not all. Growth faltering status, sex and BMI were considered as covariates.

Overall, the study was assessed as having medium risk of bias.

4.1B.5 Summary of findings

The Newcastle study showed no significant differences in outcome measurement between cases of growth faltering in infancy and comparison subjects for any of the main scales. Significant differences were found for some subscales, specifically rating of actual body shape on a pictorial scale, body shape dissatisfaction, and eating restraint. However, in each of these cases the difference became non-significant when entered into a regression analysis controlling for sex and current BMI.

In summary, no significant associations were found between infant growth faltering and mental illness in early adolescence.

5. Results relating to outcomes in adulthood

5.1 Mortality and health-related quality of life

Eligible studies were found in relation to just over half (n=7) of the 12 adult outcomes relating to mortality and morbidity (listed in full in appendix 3). These were ischaemic heart disease (IHD), stroke, non-insulin dependent diabetes (NIDDM), cancer of lung and breast, osteoarthritis and mental illness (depression and suicide).

5.1A Ischaemic heart disease

Tables in appendix 10 relating to IHD are tables 5.1A (i) and (ii).

5.1A.1 Selection of studies

Searches of Medline and Embase identified 1445 references. Screening of titles and abstracts led to identification of 9 studies that would potentially meet inclusion criteria for the review. Four of these were subsequently ruled out as they did not include measurements during infancy. Screening of reference lists of included studies did not lead to identification of any further studies of relevance. The 5 included papers are summarised in Table 17.

Table 17: Summary of included studies in review of infant growth status and IHD

Ref. no	Title	Author (year)	Design
4	Weight in infancy and death from ischaemic heart disease	Barker (1989)	Retrospective cohort
5	Early growth and coronary heart disease in later life: longitudinal study	Eriksson (2001)	Retrospective cohort
97	Weight in infancy and prevalence of coronary heart disease in adult life	Fall (1995)	Retrospective cohort
98	Early growth and death from cardiovascular disease in women	Osmond (1993)	Retrospective cohort
99	Growth of girls who later develop coronary heart disease	Forsen (2004)	Retrospective cohort

5.1A.2 Description of studies

All 5 were retrospective cohort studies[4;5;97-99]. Three were based on the cohort of adults born in Hertfordshire UK between 1911 and 1930[4;97;98], and the other two on men and women born in Helsinki, Finland between 1933 and 1944[5;99].

Three of the studies were based on men[4;5;97]. One concentrated on women[99] and the fifth study on both sexes[98]. Mean age at follow-up is not reported in any of the studies. However years of birth were reported and, based on these data, the men studied in Hertfordshire were between 60 and 80 years of age at the time of follow-up whereas the Hertfordshire women were aged between 60 and 70 years. The Helsinki studies were based on a younger cohort of men and women aged between 53 and 63 years of age[5]. Ischaemic heart disease (IHD) was defined according to the International Classification of Disease (ICD9 410 - 414) in all of the studies although 4 used the term coronary heart disease in preference to IHD. Two of the Hertfordshire studies examined the relationship between infant size/growth and mortality from IHD[4;98]. The Finland studies examined mortality and morbidity expressed as a hazard ratio for IHD[5;99]. The third Hertfordshire study examined the prevalence of IHD based on Rose angina questionnaire and ECG findings[97].

5.1A.3 Assessment of infant size/growth

All 5 studies reported birthweight and weight at 1 year of age which were ascertained from historical birth and child health records. In each study the

measurements had been taken by a midwife or health visitor. The Finland studies also reported body mass index (BMI) and height at 1 year[5;99].

All 5 studies considered the relationship between infant size and IHD, each reporting the relationship between weight at 1 year and the IHD outcome. Three of the studies also considered the relationship of infant growth with IHD[4;5;97]. However, only three of the studies by Fall et al, Eriksson et al and Forsen et al reported their findings in a form that could be abstracted. The study by Fall et al describes the simultaneous analysis of the relationship of birthweight and weight at 1 year with prevalence of IHD and the Finland studies report the findings of a Cox's proportional hazards regression analysis of the simultaneous effect of birthweight and weight at 1 year on risk of IHD[5;97;99]. A similar analysis was carried out in the study by Barker et al but the findings were reported in the form of a figure displaying lines joining points of equal risk for IHD according to birthweight and weight at 1 year. The Finland study of men reports an analysis of the effect of weight at 1 year independent of birthweight for 'all cardiovascular diseases' but not for IHD alone[5].

5.1A.4 Quality assessment

Losses to follow-up in the three Hertfordshire studies were greater than 50% and all were considered to have high risk of bias in this respect, although non-participants appear to have been comparable to subjects with regard to exposure status[4;97;98]. Follow-up rates in the Finland studies were much higher at 80% and 75%[5;99]. A particular strength of the 4 mortality studies was their large sample size and all were considered to have low risk of bias in this respect[4;5;98;99].   However in the Hertfordshire study of prevalence, which required subjects to attend a research clinic and have ECGs and blood tests, only a small proportion (25%) of the target population were followed up and sample size was small at 290[97].

Measurements of size at birth and at 1 year were obtained from birth/child health records in all studies but since the studies are retrospective, the reliability of the methods used to measure infants is unknown. All five reported that rounding of weight measurements up to a 0.25lb (113g) had occurred although this clearly could not have been biased by any knowledge of later outcome. Outcome status was ascertained using valid techniques in all 5 studies. The 4 based on mortality used death certificates and any misclassification of cause of death will not have been influenced by knowledge of exposure status. However, the Finland studies also relied on morbidity data obtained from hospital records, which are likely to have been a less complete and less reliable source of outcome data. The study by Fall et al employed well-validated methods to collect information on prevalence of IHD. These were the WHO Rose angina questionnaire and ECGs administered to subjects in a clinic setting. Observers interpreting the ECGs were blind to exposure status.

The confounding factors we considered important in this review of IHD were: current size (body mass index), socio-economic status, ethnicity, smoking, alcohol intake and physical activity. Since the studies were all retrospective,

few data were available to allow adequate adjustment for the effects of important confounding variables and we considered 4 of the 5 studies to have high risk of bias in relation to confounding[4;5;98;99]. In the study by Barker et al birth register data on infant feeding allowed consideration of the relationship between weight at 1 year and SMR from IHD according to pattern of infant feeding[4]. The association between weight at 1 year and IHD was absent in the bottle-fed group but the group represented only 7.6% of the total sample. The study by Osmond et al considered infant feeding and social class in the analysis relating to cardiovascular disease but not to IHD[98]. The Finland studies did not make any adjustments for confounding factors[5;99]. The study by Fall et al included interviews with subjects in adulthood and so it was possible to collect information on current lifestyle allowing consideration of the effect of smoking[97]. The relationship between infant size and prevalence of IHD are described according to current social class and smoking but none of the other confounding factors were adjusted for. None of the studies reported analyses adjusted for adult body mass index (BMI).

5.1A.5   Summary of findings

Two of the Hertfordshire studies reported the relationship between weight at 1 year and SMR for IHD across subgroups of subjects (categorised according to weight at 1 year) and showed a significant negative relationship[4;98].

The study by Barker et al reported an SMR in the lowest 1-year weight (≤18lb) category of 104 compared with 81 in the highest 1-year (≥27lb) category and this corresponded to a 17% change in risk across categories[4]. The findings from a Cox proportional hazards regression analysis were reported in the form of a figure and the general trend was for greatest risk of death from IHD among the men with lowest birthweight and lowest weight at 1 year. The study by Osmond, which used the same 1-year weight groupings, reported SMR in the lowest 1-year weight group of men of 105 compared with 42 in the highest group and this corresponded to a 16% change across groups. The corresponding SMR values for women were 91 and 76 but this trend was not statistically significant[98].

The study of men born in Helsinki also reported a significant inverse relationship between size at 1 year and IHD[5]. The hazard ratio for risk of death from CHD for subjects in lowest 1-years weight category (≤18lb) compared with the highest category (>26lb) was 1.82 (95% CI 1.25 to 2.64). Corresponding hazard ratios for 1-year BMI and 1-year height were 1.83 (1.28 to 2.60) and 1.55 (1.11 to 2.18) respectively. The effect of infant growth was assessed in a Cox's proportional hazards regression analysis which demonstrated that the inverse association between size at 1 year (weight and body mass index) was independent of size at birth. Patterns of infant and childhood growth in the 357 men who developed IHD were also demonstrated graphically with z scores for weight, height and BMI according to age, where the z-score for the cohort was set at zero. The general trend was for low birthweight and low weight at 1 year followed by accelerated weight gain during childhood. The raw data used to produce these graphs was not presented. The Finland study of women failed to show an association between weight at 1 years and hazard ration for CHD[99]. In the study of men

born in Hertfordshire by Fall et al, prevalence fell progressively with increasing weight at 1 year[97]. The difference in 1-year weight between those with CHD and those without was −1.0lb (95% CI −1.8 to -0.1) and these findings remained statistically significant after adjustment for birthweight. The approximate risk reduction across each 1-year weight category was 13%. The confounding effects of current social class and smoking were explored. Men in lower social classes had higher IHD prevalence. However the inverse relationship between infant growth and IHD was observed within each social class and was also present in current, former and non-smokers.

The differences between the statistical analyses reported make it difficult to produce any quantitative summary of the relationship between infant size/growth and IHD across the four studies. Two of the studies report different summary measures of IHD risk according to weight at one year (hazard ratios and odds ratios)[5;97]. The other studies describe the significance of trends across subgroups of subjects but do not report any summary statistics describing the association between infant size or growth and IHD.

A summary of the direction of the relationship between infant size (weight at 1 year) and IHD was feasible and is shown in table 18. Four of the 5 studies demonstrated an inverse relationship between weight at 1 year and IHD although in the study by Osmond et al this applied to the men alone – the findings in the Hertfordshire women were not statistically significant. This is consistent with the findings of the study by Forsen et al of the women born in Helsinki, Finland. The Hertfordshire study by Fall et al and the Finland study by Eriksson et al also showed an inverse association between infant growth and IHD. No non-linear relationships between infant size or growth and IHD were reported in any of the studies.

Table 18: Summary of IHD findings

Direction of relationship with infant size	*Number of studies (ref. nos. in parentheses)*		
	Males	Females	Total
Inverse	4 [4;5;97;98]	0	4 [4;5;97;98]
None	0	2 [98;99]	1 [98]
Positive	0	0	0
Inconclusive	0	0	0
Total	4 [4;5;97;98]	1 [98]99	5*

*This figure is the total number of studies examined in the table, and may not equal the sum of the subtotals due to studies appearing more than once in certain cells of the table.

In summary, the findings were consistent across studies; men who had smaller size in infancy experienced greater mortality and morbidity from IHD as adults. No association was demonstrated in women. However the fact that these findings are based on just two cohorts and on studies assessed as having medium risk of bias somewhat limits their external validity.

5.1B Cerebrovascular disease

Tables in appendix 10 relating to cerebrovascular disease are tables 5.1B (i) and (ii).

5.1B.1 Selection of studies

Electronic searches identified 4746 abstracts. Screening of titles and abstracts led to identification of only one reference that met the inclusion criteria for the review, which is summarised in Table 19.

Table 19: Summary of included studies in review of infant growth status and cerebrovascular disease

Ref. no	Title	Author (year)	Design
[100]	Mothers' pelvic size, fetal growth, and death from stroke and coronary heart disease in men in the UK	Martyn (1996)	Retrospective cohort

5.1B.2 Description of studies

The single included study, which was based on the Hertfordshire cohort, explored the relationship between infant growth and stroke[100]. The subjects were 10141 men born between 1911 and 1930.

5.1B.3 Assessment of infant size/growth

The relationship between infant size (weight at 1 year) and risk of stroke in later life was assessed.

5.1B.4 Quality assessment

This study was based on the same cohort of Hertfordshire males as that used in the IHD study by Osmond[98]. Like the Osmond study, the reliability with which measurements of infant size were taken is unknown and measurements were rounded by up to 0.25lb, although death was reliably ascertained from death certificates. Only 44% of the target subjects were followed-up, and no information was given on those lost to follow-up, resulting in a high risk of bias. No confounding factors were considered in this study. However, since the design, setting and measurements were considered to be fairly good, the study findings were rated overall as having medium risk of bias.

5.1B.5 Summary of findings

SMRs for stroke decreased from 105 in the lowest 1-year weight category (≤ 18lb or 8kg) to 52 in the highest 1-year weight category (> 26lb or 12kg). The fitted percentage reduction in SMR between successive 1-year weight categories was 14% (95% CI 2 to 25%), using log-linear regression.

No further studies relating to the relationship between infant growth status and cerebrovascular disease have been identified during review searches.

In summary only one study examined the association between infant growth status and cerebrovascular disease. The findings of this study suggest that infant size is inversely associated with risk of death from stroke in men. However the generalisability of these findings, based on only one population in the south of England, is limited.

5.1C   Non-insulin dependent diabetes

Tables in appendix 10 relating to non-insulin dependent diabetes (NIDDM) are tables 5.1C (i) and (ii).

5.1C.1   Selection of studies

Electronic searches of Medline and Embase identified 6524 abstracts. Screening of titles and abstracts led to identification of 1 study. Two further studies published during the review were identified through correspondence with experts.

5.1C.2   Description of studies

All three NIDDM studies had a cohort design as seen in table 20.

Table 20: Summary of included studies in review of infant growth status and NIDDM

Ref. no	Title	Author (year)	Design
[101]	Fetal and infant growth and impaired glucose tolerance at age 64	Hales (1991)	Retrospective cohort
[102]	Early adiposity rebound in childhood and risk of Type 2 diabetes in adult life	Eriksson (2003)	Retrospective cohort
[103]	Relation of serial changes in childhood body-mass index to impaired glucose tolerance in young adulthood	Bhargava (2004)	Prospective cohort

Of the two retrospective studies, both were set in Europe and were based on births in the early half of the 20th century. One was based on males born between 1920 and 1930 from the same Hertfordshire-based cohort as previously discussed[101], whilst the other was set in Finland and included subjects of both sexes born in a single city hospital between 1934 and 1944[102]. The remaining study was based on a prospective cohort of subjects born between 1969 and 1972 in a geographically defined region of India[103].

5.1C.3   Assessment of infant size/growth

The Hertfordshire study reported size in terms of weight at 1 year, whilst the other two studies measured weight, height and BMI longitudinally throughout childhood. These studies also compared the size and growth of those subjects that later developed NIDDM with the cohort as a whole using z-scores.

5.1C.4   Quality assessment

The two retrospectively designed studies were considered to have a medium risk of bias overall, whilst the Indian prospective cohort was considered to be low risk. In the Hertfordshire and Indian studies subjects with NIDDM were diagnosed following a glucose-tolerance test at a clinic which was carried out by the researchers[101;103]. Full descriptions of the procedures and definitions used were given. The Finland study, however, obtained cases from a national

register of persons receiving diabetes medication[102]. This is a less reliable method, since it is likely to underestimate the true number of cases, and there may be significant differences (socio-economic, for instance) between those identified and those not. The national register also did not differentiate between insulin and non-insulin dependent diabetes, and so the study excluded all those under 40 years in an effort to compensate. The Hertfordshire study stated that some of its researchers at follow-up were blind to infant size/growth data, although it was unclear whether full blinding had been achieved. No blinding information was given for the other 2 studies. All three studies, as befits their cohort design, had large sample sizes but high rates of attrition.

We considered socio-economic status, current size and ethnicity to be important confounders for NIDDM. The Indian study adjusted for sex and current size but not infant feeding or socio-economic status. The cohort was fairly ethnically homogeneous (84% were Hindu)[103]. Neither of the other 2 studies considered any confounding factors in their analyses.

5.1C.5   Summary of findings

The Hertfordshire study reported that prevalence of newly-diagnosed NIDDM between 59 and 70 years of age was 17% among those of lowest weight at 1 year (at most 18lb or 8.16kg), falling to 2% among those of highest weight at 1 year (at least 27lb or 12.25kg)[101]. The approximate risk reduction across each infant weight category was 28%. However, no statistical analysis was carried out. The Finland study suggested that cases of later NIDDM were slightly smaller in infancy than the cohort as a whole, as measured by weight, height or BMI, but these differences were not significant. Furthermore, the Indian study, after adjustment for potential confounders, presented odds ratios that were almost indistinguishable from one, and non-significant[103].

In summary, there is some slight evidence that adults with NIDDM were of smaller size in infancy than comparison groups, but statistically the results were not significant. Only 3 studies were found, of which only one made any consideration for potential confounders. Having adjusted for sex, age and current size, the effect sizes were negligible, suggesting that the slight evidence presented in the other 2 studies can probably be accounted for by these confounders. The overall conclusion is that there is no association between infant size or growth and NIDDM in adulthood.

5.1D   Cancer

Tables in appendix 10 relating to cancer are tables 5.1D (i) and (ii).

5.1D.1   Selection of studies

Electronic searches identified 12,678 references. Screening of titles and abstracts led to identification of 14 studies that were considered to potentially meet inclusion criteria for the review. Following assessment, all but one of the studies were excluded. This study relates to childhood cancer and is described in section 3.1B. Two further studies of lung cancer had already been identified through the work on IHD. These also met the inclusion criteria for the cancer review. Screening of reference lists did not lead to identification

of any additional studies. A third study, of breast cancer, was identified through contact with experts. The 3 included papers are summarised in Table 21.

Table 21: Summary of included studies in review of infant growth status and adult cancer

Ref. no	Title	Author (year)	Design
4	Weight in infancy and death from ischaemic heart disease	Barker (1989)	Retrospective cohort
5	Early growth and death from cardiovascular disease in women	Osmond (1993)	Retrospective cohort
104	Childhood growth and breast cancer	De Stavola (2004)	Retrospective cohort

5.1D.2 Description of studies

All 3 studies were retrospective population-based cohort studies. Two were based on cancer deaths in the Hertfordshire cohort also reported in the section on IHD. The first was a study of lung cancer in men born between 1911 and 1930[4]. The second was also a study of lung cancer, this time in a larger cohort of males born between 1911-30 and a new cohort of females born between 1923-30[5]. The third study was based on the MRC National Survey of Health and Development which is a socially stratified birth cohort of men and women born in 1946. The study was a follow-up of 2187 women to examine breast cancer morbidity between the ages of 36.4 and 53.8 years[104].

5.1D.3 Assessment of infant size/growth

Both studies of lung cancer assessed the relationship between infant size (weight at 1 year) and death from lung cancer. The breast cancer study assessed the relationship of height and BMI at 2 years of age and later breast cancer risk.

5.1D.4 Quality assessment

All 3 studies were population-based and included internal comparison groups and we considered their setting and design to be appropriate for their stated objectives. Follow-up rates were generally poor in the Herfordshire studies at 36% and 43% leading to a high risk of bias[4;5]. Comparison of infant size showed that subjects had 'slightly higher' weights at 1 year than those lost to follow-up but no statistical analyses were presented. There was less attrition in the breast cancer study with 86% of the target population followed up[104].

Measurements of size at 1 year were obtained from child health records and so the reliability of methods used to take the measurements is unknown. The measurements of infant size in Hertfordshire were all rounded as in the IHD and mental illness studies and insufficient information on rounding was given in the study by De Stavola et al. Outcome status in the lung cancer studies was ascertained from death certificates by tagging of subjects at NHSCR and any misclassification of cause of death would not have been biased by knowledge of exposure status. Breast cancer outcome was ascertained through a postal questionnaire completed by subjects and so there is potential

for reporting bias although verification from the NHS cancer register was also sought[104].

The confounding factors that we considered important in studies of cancer were current size (body mass index), smoking, physical activity, alcohol intake and socio-economic status. Consideration of confounding factors was poor in all 3 studies and they were assessed as having a high risk of bias in this respect. None of the studies adjusted for any of the confounding factors in the analyses of interest to the review.

5.1D.5   Summary of findings
The two studies of lung cancer failed to show an association between weight at 1 year and SMR from lung cancer in either men or women[4;5]. While both studies were based on large population-based samples, losses to follow-up were high and the generalisability of findings is also limited by the fact that both were based in one geographical area in southern England. We have not identified any further studies of the relationship between infant growth status and lung cancer.

The study of breast cancer also failed to show an association between 2-year size and growth and odds ratio for breast cancer[104]. Although this might be partly attributed to the lack of cases of breast cancer in a relatively young cohort of women, significant associations with later growth (age 4-7 years) were demonstrated suggesting that the study was sufficiently powered.

In summary there is very limited research evidence of the relationship between infant growth status and adult cancer. What evidence there is suggests that size in infancy is not related to risk of breast or lung cancer.

5.1E   Osteoarthritis
Tables in Appendix 10 relating to osteoarthritis are tables 5.1E (i) and (ii).

5.1E.1   Selection of studies
Electronic searches identified 319 abstracts. Screening of titles and abstracts led to identification of only one study that met the inclusion criteria for the review. The included paper is summarised in Table 22 below.

Table 22: Summary of included studies in review of infant growth status and osteoarthritis

Ref. no	Title	Author (year)	Design
105	Weight from birth to 53 years: a longitudinal study of the influence on clinical hand osteoarthritis	Aihie Sayer (2003)	Retrospective cohort

5.1E.2   Description of studies
The single included study was a retrospective cohort study of 2986 men and women in the UK 1946 national survey of health and development. They were followed up at a mean age of 53 years to look for the presence of hand osteoarthritis (OA)[105].

5.1E.3 Assessment of infant size/growth

The relationship between infant size (weight at 2 years) and risk of OA was assessed.

5.1E.4 Quality assessment

Study design was considered appropriate for exploring the association between infant size and adult OA. Measurement of weight at 2 years was carried out by community nurses and subjects were weighed to the nearest 0.1kg. These measurements were not carried out under experimental conditions but as part of normal child health surveillance and the study was assessed as having medium risk of bias in this respect. Outcome assessment was through hand examination by trained nurses using well-validated clinical criteria with good reproducibility and risk of bias was considered low. Almost a third (31%) of subjects were lost to follow-up but the authors state that study subjects were representative of the national cohort population in terms of socio-economic status and birthweight. The confounding factors we considered important were socio-economic status, sex and current size. However only sex was adjusted for in analyses. Overall the study findings were rated as having medium risk of bias.

5.1E.5 Summary of findings

Prevalence of hand OA in the men was 19% and in the women was 30%. Mean 2-year weights for subjects with and without hand OA were compared. The differences were not statistically significant in either the men or the women. There was no association between weight at 2 years and prevalence of hand OA.

In summary, in the single study identified, there was no evidence that infant size was related to later risk of OA.

5.1F Mental illness

Tables in appendix 10 relating to mental illness are tables 5.1 F (i) and (ii).

5.1F.1 Selection of studies

Electronic searches identified 9,876 references. Screening of titles and abstracts led to the identification of 15 studies that were considered to potentially meet inclusion criteria for the review, although 13 of these related to outcomes other than mental illness. Following assessment only 1 study relating to mental illness was found to meet inclusion criteria. A further study was identified through screening reference lists. The two included papers are summarised in Table 23.

Table 23: Summary of included studies in review of infant growth status and mental illness

Ref. no	Title	Author (year)	Design
106	Low weight gain in infancy and suicide in adult life	Barker (1995)	Retrospective cohort
107	Birth weight and the risk of depressive disorder in late life	Thompson (2001)	Retrospective cohort

5.1F.2 Description of studies

Both studies were based on the Hertfordshire cohort. The first explored the relationship between infant growth and suicide[106]. The subjects comprised 10141 men born from 1911-30 and 5585 women born 1923-30. The second, which was based on 542 men and 340 women who represented a sub-set of births from 1920-30, explored the relationship between infant size and growth and adult depression[107].

5.1F.3 Assessment of infant size/growth

The relationship between infant size (weight at 1 year) and infant growth (weight gain from birth to 1 year) and later outcome were assessed in both studies.

5.1F.4 Quality assessment

Both were large-scale population based study where outcome status was reliably ascertained either through death certification of through administration of well-validated depression scales by trained nurses. Like the other analyses based on the Hertfordshire cohort, the reliability with which measurements of infant size were taken was unknown and measurements were rounded by up to 0.25lb. Losses to follow-up were high in the study of suicide with only 55% of men and 40% of women followed-up resulting in a high risk of bias[106]. However non-participants were reported to have similar weights at birth and at 1 year to subjects. Follow-up was more complete in the study by Thompson et al at 86% in men and 87% in women. However information on non-participants was insufficient to be sure that their patterns of infant growth did not differ from those of participants.

The confounding factor we considered important in relation to mental illness was socio-economic status and we considered alcohol/drug intake to be relevant in terms of effect modification. In the study of suicide, socio-economic status at death was considered but not adjusted for in analyses. Consideration of confounding in the study by Thompson et al was thorough and socio-economic status was adjusted for in analyses as were birthweight and presence of CHD.

Overall the study findings of both of these studies were rated as having medium risk of bias.

5.1F.5 Summary of findings

There was a significant inverse association between infant size and death from suicide in all subjects in the study by Barker et al with SMR in the lowest 1-year weight group of 61 compared with 34 in the highest group. In the men there was also a significant increased risk of suicide associated with poorer infant growth with a 45% increase in risk with every kg decrease in weight gain (95% CI 7 to 98%). The corresponding figure in the women was 31% but this was not statistically significant. The study of depression did not demonstrate an association between infant size or growth and depression in adulthood. These results are shown in Figure 2.

Figure 2: Plots of results from studies by Barker[106] and Thompson[107] on infant growth and mental illness
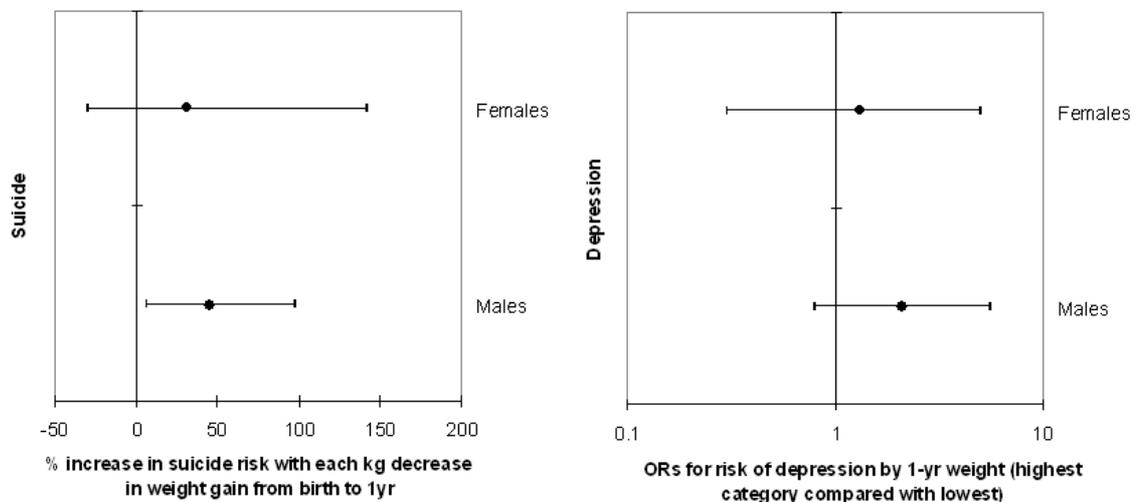
No further studies relating to the relationship between infant growth status and later mental illness or death from suicide have been identified during review searches. The generalisability of these findings based on only one population in the south of England is therefore limited.

5.2 Significant health-related behaviours

As outlined below, we identified a number of eligible studies relating to adult obesity, a proxy for unhealthy patterns of eating and physical activity. However, although searches in Medline and Embase identified 12,179 references, no eligible studies were found in relation to any of the other outcomes which included unhealthy eating and lack of physical activity themselves, smoking, drug and alcohol misuse and unsafe sex.

5.2A Adult obesity

Tables in appendix 10 relating to adult obesity are tables 5.2A (i) and (ii).

5.2A.1 Selection of studies

The selection of papers relating to obesity at any stage of the life course is described in section 3.2A.1. Of the 21 studies relating to obesity at any age, 12 concerned obesity in adulthood or late adolescence (15 years or over), and

it is these studies that are discussed here. They are summarised in Table 24. The remaining 9 studies concerned obesity in childhood, and are presented in section 3.2A.1.

Table 24: Summary of included studies in review of infant growth status and adult obesity

Ref. no	Title	Author (year)	Design
108	Prediction of adult overweight during the pediatric years	He (1999)	Retrospective cohort
109	Predicting obesity in young adulthood from childhood and parental obesity	Whitaker (1997)	Retrospective cohort
110	Tracking the development of adiposity from one month of age to adulthood	Rolland-Cachera (1987)	Longitudinal cohort
111	The predictive value of childhood body mass index values for overweight at age 35 y	Guo (1994)	Longitudinal cohort
112	Obesity from cradle to grave	Eriksson (2003)	Retrospective cohort
113	40-year follow-up of overweight children	Mossberg (1989)	Longitudinal cohort
114	Childhood antecedents of adult obesity. Do chubby infants become obese adults?	Charney (1976)	Retrospective cohort
115	Rapid weight gain during infancy and obesity in young adulthood in a cohort of African Americans	Stettler (2003)	Retrospective analysis of prospective cohort
116	Two-decade follow-up of fatness in early childhood	Garn (1985)	Longitudinal cohort
117	A prospective study of weight and height going from infancy to adolescence	Tienboon (2002)	Prospective cohort
118	Birth size, early childhood growth, and adolescent obesity in a Brazilian birth cohort	Monteiro (2003)	Prospective cohort
119	The relationship between obesity in adolescence and early growth	Heald (1965)	Case-control

5.2A.2 Description of studies

The 12 studies explored the relationship between size or growth in infancy and overweight or obesity in adulthood. Of 11 cohort studies, 6 were prospective, 4 retrospective, and one was a retrospective analysis of an originally prospectively designed study. One study was of case-control design. All were set in developed countries, except the study by Monteiro, which was set in a city in Brazil[118], and none had inclusion criteria defined by social or ethnic groups except Stettler (2003), which was restricted to African-Americans[115]. Seven of the studies either reported odds ratios for the risk of later obesity among obese or rapidly growing infants compared to that among 'normal' infants, or gave sufficient information for such an odds ratio to be calculated[108-111;114;117;118].

All but one of the studies included subjects of both sexes; the exception was the case-control study by Heald, which was restricted to females[119]. Three studies reported results separately for each sex, although one of these also reported combined results[108;111;112].

Age at follow-up ranged from 14 to 34 years. Two studies calculated an average adult size over a period of 8 or 9 years, whilst the remainder used a single size measurement at a particular age.

5.2A.3 Assessment of infant size/growth

Six of the studies reported infant size at either 1 or 2 years of age or both using body mass index (BMI)[108-112;117], of which all but one defined obesity as having a BMI above a certain point. Such 'cut-off' points were defined using percentiles (the 75th or 85th)[109-111], standard deviations above the mean[117], or by a raw value (18kg/m^2)[108]. The Eriksson study split the infant BMI distribution into 5 strata but did not explicitly define obesity in infancy[112]. Three further studies defined infant obesity using measures other than BMI: two used weight or weight-for-height[114;118] and the other used subscapular or triceps skinfolds[116]. 'Cut-offs' for obesity were defined in similar ways to those used for BMI. The Mossberg study reported size in terms of weight-for-height z-score at admission to hospital, which was before the age of 2 years[113]. It is suggested that such a z-score of +1.0 corresponds to approximately 15% excess weight, but obesity in infancy is not formally defined. The case-control study by Heald reported weight and height measures at 6 months and 1 year, and also growth in weight and height between each of birth, 6 months and 1 year. Two other studies reported a measure of growth in infancy: one defined 'rapid' growth as a weight-for-age z-score change from birth to 2 years of over 0.67, whilst the other defined it as an increase in weight-for-age z-score over the first 4 months above one standard deviation from the mean[115].

5.2A.4 Quality assessment

A prospective cohort study was considered to be the optimal design. Four of the studies used a retrospective design, which risks bias from missing or incomplete data[108;109;112;114].

Setting was also considered to be a potential source of bias in many of the studies – it was often a single institution or a small region, which is unlikely to be representative of the general population. One study retrospectively combined data from 4 different studies, including subjects born anywhere between 1929 and 1991, making interpretation difficult[111]. Only one study (by Monteiro) was considered to have a low risk of bias with respect to its design and setting. It was a prospective birth cohort consisting of all hospital births in a city (accounting for 99% of the total) and designed for the collection of relevant growth data[118].

Three studies were considered to have a high risk of bias[113;114;119]. Common sources of bias were attrition, a lack of information on the source population and the method of selection used for the study subjects. Thus, it is hard to judge whether the study subjects are truly representative of the general population. The analyses themselves were generally not a serious source of

bias. Often either a complete regression analysis was carried out, and full results and errors presented, or cell counts were presented, allowing estimates of size of effect and standard errors to be calculated manually. The exceptions were the studies by Mossberg and Heald, and both were judged to have a high risk of bias overall.

The details of ascertainment of infant size or growth were not always entirely clear. Four studies abstracted such measurements from child health records, which can be assumed to be fairly reliable[108;109;112;114]. Three studies used measurements taken as part of a formal longitudinal study, but the procedures are not sufficiently explained[110;115;116]. The studies by Monteiro and Mossberg, whilst stating that study subjects were measured either by the research team or in hospital, did not give sufficient methodological or repeatability information[113;118]. The study by Guo stated that the measurement procedures used were "closely similar" to those recommended by a reference manual, and that their reliability was "excellent", but since data from several different studies was used the methods are likely to have differed, potentially introducing bias[111].

Eight of the studies defined obesity in adulthood using BMI; 2 others used weight or weight-for-height, and a further study used subscapular or triceps skinfold. The age at which obesity was defined varied widely, from 18 years to middle-age or a lifetime maximum, making general conclusions difficult. For 7 studies, the method of ascertainment of size at follow-up was the same as that used in infancy – either child health or other records, or measurement by the research team as part of the study[108-111;115;116;118]. One study abstracted size in infancy from child health records but had size at follow-up measured by a single researcher[117], whilst 3 further studies ascertained size in infancy either by direct measurement or from child health records but ascertained size at follow-up by self-response questionnaire[112-114]. This latter approach is especially vulnerable to bias, as the responses are likely to be prone to misreporting; 2 of the studies reported efforts to ensure such bias was minimised[112-114]. The case-control study by Heald is even more vulnerable – the adolescent cases are defined only by their attendance at a "fat camp", with no formal definition of obesity at all[119]. None of the studies gave information as to whether observers measuring or abstracting the follow-up data were blind to subjects' infant size or growth data.

Two studies calculated an average adult size over a period of 8 or 9 years[109;111], whilst the remainder used a single size measurement at a particular age. The former approach possibly has a lower risk of bias, since subjects' weights can fluctuate. Five studies defined obesity in adulthood by the raw BMI value being above a certain threshold, varying between 25 and 30 kg/m^2[108;109;111;112;115]. A further 2 studies used BMI percentiles of a reference population to define adult obesity: one used the 75th percentile[110], one the 85[th][118]. Two studies used weight-for-height above a certain distance from the mean (≥ 1 standard deviation)[11] or median (≥ 20%)[101], whilst a further study used the 85th percentile of a reference population for skinfold thickness[116].

The confounding factors we considered important in the relationship between infant growth and obesity were socio-economic status, parental size and infant feeding. With the exception of 4 studies, which either adjusted for all or most of these in a regression analysis or presented separate analyses in which individual confounders were controlled for[109;114;115;118], none of the studies considered any of these important confounders either in regression or otherwise. This is a major flaw, and together with high attrition and lack of information on the source population, means that there is no guarantee that the study subjects would be representative of the general population.

5.2A.5 Summary of findings

Of the 12 studies of adult obesity in total, 9 studies measured size in infancy, 2 measured both size and growth, and 1 study measured infant growth only. Eight of the studies (7 of size and 1 of growth) specifically considered infant obesity.

Of the studies that defined infant size in terms of obesity, 4 studies presented significant odds ratios or relative risks indicating that the risk of becoming obese in adulthood is greater among heavier or fatter infants than among other infants[108;110;114;118]. However, it should be noted that definitions of infant obesity varied between studies, as described in section 5.2A.4, so that a particular infant deemed obese in one study might not have been deemed obese if they had taken part in a different study. The study by Charney et al suggested that the positive association between infant and adult obesity was stronger among subjects with at least one overweight parent[114]. Of the other 3 studies considering infant obesity, Whitaker et al suggested that the association remains non-significant after the addition of parental obesity to a regression model[109]. This study and those by Guo et al[111] and Tienboon et al[117] all presented odds ratios that were either non-significant or inconclusive. The study by Garn et al reported that infants defined as obese in terms of skinfold thickness remained obese in adulthood than would be expected by chance[116].

Of the 3 remaining studies that considered infant size, the study by Eriksson reported that, for both sexes, the prevalence of obesity in adulthood was significantly lower among infants with lower BMI at 6 months than among those with higher BMI[112]. The case-control study by Heald reported that cases (obese adolescents) had significantly higher weight at 1 year than controls, although at 6 months the difference was not significant. At neither 6 months nor 1 year was height significantly different between cases and controls, suggesting that infant weight-for-height, and therefore presumably fatness, was greater among cases than controls[119]. Finally, the Mossberg study reported that, whilst the study subjects had a weight-for-height standard deviation score of +2.3 in infancy, this had dropped to +1.8 by late childhood and to +0.2 by mid adulthood (40 to 50 years). This cohort of significantly overweight infants became, on average, only slightly overweight adults. This suggests that infant obesity is not associated with adult obesity – though this study was very small, did not give any statistical analysis, and did not make consideration for confounding factors. Its risk of bias was high and its findings are not generalisable[113].

The findings of these studies relating to infant size and infant obesity are summarised in table 25.

Three studies reported findings relating growth in infancy to adult obesity (table 24). The Brazilian study by Monteiro et al reported an odds ratio of 1.66 (95% CI 1.20 to 2.31) for infants growing more rapidly (> 0.67 z-score change) from birth to 2 years, after adjustment for all potential confounding factors. The study of African-American subjects by Stettler et al reported an odds ratio of 5.22 (95% CI 1.55 to 17.6) for rapid growth from birth to 4 months (weight-for-age z-score change ≥1 SD from the mean), after adjustment for birthweight, sex and maternal size and education but not infant feeding. The case-control study by Heald et al reported that growth in weight from birth to 1 year was significantly greater among cases (obese adolescents) than controls, but that growth in height over the same period was not significantly different. Splitting this time period into 2 six-month periods did not yield any significant differences in growth in either weight or height.

Table 25: Summary of findings relating infant size and growth to adult obesity

Direction of relationship with infant size	*Number of studies (ref. nos. in parentheses)*			
	Infant size	Infant size (obesity)	Infant growth	Total
Inverse	0	0		0
None	0	3 [109;116;117]		3
Positive	1 [112]	4 ({He, 1999 10103 /id;Rolland-Cachera, 1987 10106 /id;Charney, 1976 10135 /id;Monteiro, 2003 10133 /id})	3 [115;118;119]	5
Inconclusive	2 [113;119]	1 [111]		3
Total	3	8	3	12

Of the 11 studies relating infant size (or obesity) to adult obesity whose findings are described above, seven either reported an effect size measure (relative risk or odds ratio) or enabled one to be calculated. A meta-analysis was therefore carried out using relative risks (of obesity in adulthood for those obese in infancy compared to those not obese in infancy), since this is the favoured measure of effect size for cohort studies. All studies reported infancy data at 1 year of age except those by Whitaker (1 to 2 years)[109], Charney (6 months)[114] and Monteiro (2 years)[118]. Two studies stratified data by sex, and in both cases estimates for each sex were entered into the meta-analysis as separate studies. Since, as discussed previously, adjustment for confounders was rare, unadjusted estimates were used to improve homogeneity.

Entering 7 estimates from 6 studies gave a combined estimate for the relative risk of 1.99 (95% CI 1.74 to 2.29), using a fixed-effects model and inverse-variance weights (Figure 3). The I^2 heterogeneity statistic, interpretable as the proportion of total variation in the estimate of size of effect due to heterogeneity between studies, was 39% ($p = 0.27$). This statistic is given in preference to the usual Q-statistic as Q is dependent on the number of studies

in the analysis, and thus can underestimate the amount of heterogeneity when the number of studies is small.

Figure 3: Forest plot of relative risks for adult obesity among obese infants compared to non-obese infants, with combined estimate.
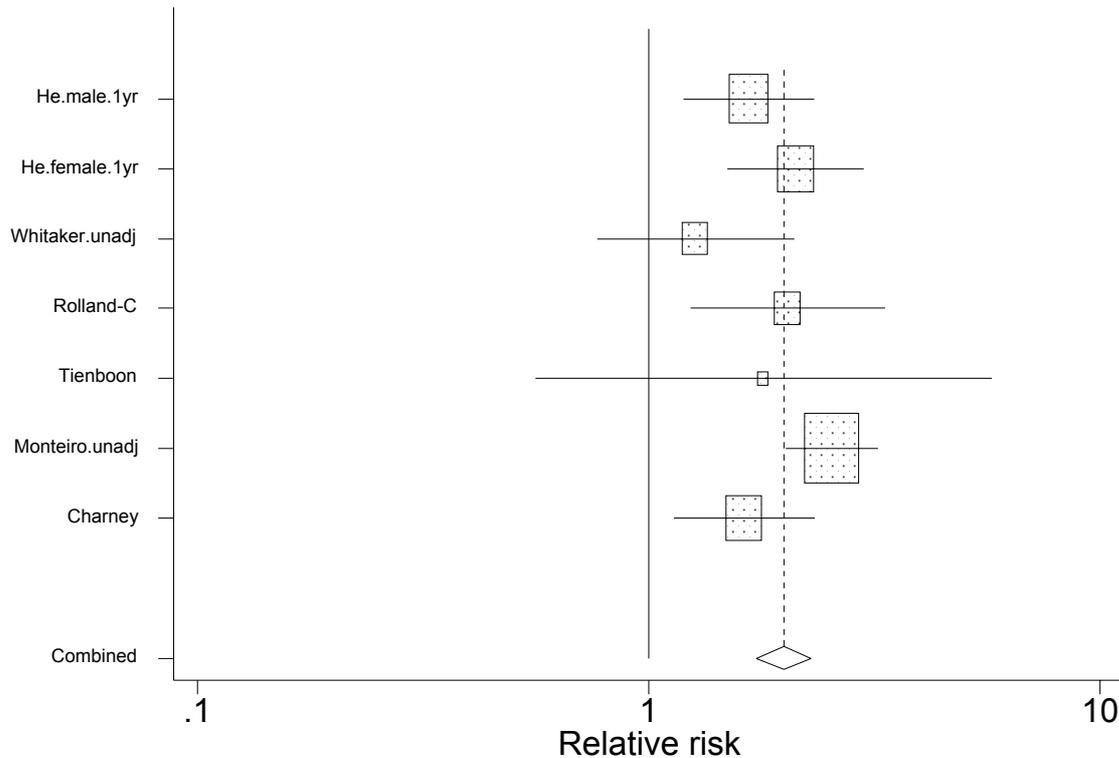
Although the heterogeneity was non-significant, a proportion of nearly 40% should still provoke some concern particularly given the differing definitions and cut-offs. The influence of each study was analysed by re-calculating the estimate omitting one study at a time, and the Monteiro study was shown to have a greater influence than any other study, reducing the effect size estimate from 1.99 to 1.75 (95% CI 1.47 to 2.08). Omitting this study reduced the heterogeneity to near zero ($I^2 \approx 0\%$; $p \approx 1.00$). This may be due to the fact that this study is unique among those considered here in being set in a developing country[118].

One study, by Guo, presented odds ratios instead of relative risks. To investigate whether the above analysis would be altered if this study could be included, the pooled estimate for the odds ratio was calculated with and without the data from the Guo study, using a fixed-effects model and inverse-variance weights as before (see table 26). A z-test for the difference in effect size between groups showed that including the Guo study in the analysis would not significantly alter the pooled estimate ($z = 1.07$; $p = 0.29$). This result also held if the Monteiro study was omitted ($z = 0.72$; $p = 0.47$).

Next, the effect of the length of follow-up between infancy and the later outcome on the pooled estimate was assessed. Three studies had follow-ups of less than 18 years (range 14 to 17 years)[108;117;118], and three studies had

follow-ups of over 18 years (range 20 to 34 years)[109;110;114]. The pooled estimate for those studies with shorter follow-ups was greater (2.18 versus 1.62) – in other words, the association between infant obesity and adult obesity was stronger for shorter follow-up times. However, a z-test for the difference in effect size between subgroups (z = 1.94; p = 0.053) and a χ^2-test for the proportion of heterogeneity explained by differences between subgroups (Q_B = 3.76 on 1 df; p = 0.053) both suggested that the effect did not quite reach statistical significance. However, if the Monteiro study is omitted, the difference is no longer anywhere near significant (z = 1.21; p = 0.23 and Q_B = 0.14 on 1 df; p = 0.70), although the pooled estimate for shorter follow-up times is still greater (2.03 versus 1.62). It may be that the near-significant result is due mostly to the presence of this one study rather than an actual effect due to follow-up length. Overall, the number of studies is probably too small to be certain. The effect of sex was also stratified for, and was found not to significantly affect the estimate, whether or not the Monteiro study was omitted.

Table 26: Sensitivity analysis to assess effects on meta-analysis of inclusion/exclusion of certain studies

Studies involved here: He, Whitaker, Rolland-Cachera, Tienboon, Monteiro, Charney, Guo.			
	RR (95% CI)	z-test	Q-test
With Monteiro	1.99 (1.74, 2.29)	z = 1.17; p = 0.24	N/A
Without Monteiro	1.75 (1.47, 2.07)		
With Guo	2.09 (1.79, 2.44)*	z = 1.07; p = 0.29	N/A
Without Guo	2.38 (1.98, 2.87)*		
< 18yrs	2.18 (1.85, 2.58)	z = 1.94; p = 0.05	Q_{diff} = 3.76; p = 0.05
> 18yrs	1.62 (1.26, 2.08)		
Males	1.64 (1.24, 2.16)*	z = 0.79; p = 0.43	Q_{diff} = 0.62; p = 0.43
Females	1.92 (1.44, 2.55)*		

* ORs, not RRs. Including Guo.

In summary, the risk of becoming an obese adult was significantly greater among heavier infants; in particular those infants with increased weight for their height. Although few studies adjusted for potential confounders, the effect appears to be independent of sex and parental size. Without adjustment for confounders, infants at the highest end of weight or weight-for-height distribution are approximately twice as likely to become an obese adult as a non-obese infant. The risk of becoming an obese adult is also significantly greater among rapidly-growing infants even after adjustment for all important confounding factors, although only 3 studies reported such data, 2 of which were set in the developing world

5.3 Non-health related quality of life
Only one eligible study was identified in this category. This study related to income and educational status.

5.3A Income and educational status

5.3A.1 Selection of studies
Over 19,000 abstracts were identified through searches of Medline and Embase. Screening of abstracts did not lead to identification of any studies relating to non-health related quality of life that met inclusion critieria. One unpublished study was identified through correspondence with first authors in the review.

5.3A.2 Description of studies
The study, listed in table 27, was based on the Finland cohort described in the section on IHD with 4630 men born between 1933-44 being followed up in adulthood to assess their income, educational achievement and occupation[120].

Table 27: Details of study relating infant growth status to adult income and educational status

Ref. no	Title	Author (year)	Design
[120]	Infant growth and income fifty years later	Barker (2004) (unpublished)	Retrospective cohort

Tables in appendix 10 relating to adult income are tables 5.3A (i) and (ii).

5.3A.3 Assessment of infant size/growth
Height at 1 year and growth in height between birth and 1 year were related to later income, educational attainment and occupation.

5.3A.4 Quality assessment
The study was a retrospective cohort study set in a single university hospital. Subjects were all survivors with relevant records.

Infant height was ascertained from city child welfare records but no information on observers or reliability was reported. The outcome, personal taxable income and occupation, were both obtained from 1990 census information. However no information was given as to how the third outcome (achieved level of education) was ascertained but previous studies of the Finland cohort had access to school records so it is likely to have been via this route.

Follow-up varied between 63 and 77% and no information is reported comparing those lost to follow-up with subjects. However the social class distribution is said to be typical of Helsinki as a whole.

Analyses were adjusted for father's occupation but not for maternal educational status which we considered to also be an important confounding factor.

Overall the study was assessed as having medium risk of bias.

5.3A.5 Summary of findings

Percentages reaching high levels of education were reported for 6 height groups at 1 year. The same analyses were carried out for mean taxable income and occupation (% becoming labourers). A significant trend was seen for those with greater height to have higher levels of education and income and to be less likely to be labourers.

A regression analysis for mean income by growth in height up to 1 year is also reported. The percentage increase in income associated with a 2-cm increase in 1-year height after controlling for father's occupation was 3.5% (1.6%, 5.4%).

In summary only one study was identified that related infant growth status to any measure of non-health related quality of life. While the findings suggest that size and growth in height are positively associated with improved outcomes in relation to income and educational status, these findings are based on only one cohort of males and therefore have limited generalisability.

Review of Lay Perspectives and Focus Groups

6 Methods

6.1 Objectives
The objectives of this part of the review were to understand:

- lay (particularly parental) perspectives on infant growth
- issues that parents, and young people think are important in relation to infant growth
- where infant growth lies among priorities for those responsible for creating and maintaining child health
- the extent to which infant growth is a salient issue for parents and children
- to generate hypotheses to explain lay views and behaviours

and to pursue these objectives through:

- a systematic review of the literature on lay views and experiences of, and attitudes toward size and growth in infancy
- 4 focus groups relating the findings of the review to local contexts, addressing remaining questions and examining the acceptability of possible interventions in infant growth.

The review of lay perspectives on infant size and infant growth which follows describes:

a) the methodology and findings of a systematic review of the literature relating to views of early growth and
b) the methodology and findings of interviews on infant size and growth with families living in East London conducted along side the systematic review.

The findings of the systematic review are synthesised using two methods: thematic and narrative synthesis, described in sections 7 and 8 respectively.

6.2 Searching, Identifying and retrieving relevant studies

The guidance from CRD on carrying out reviews[9] and the review criteria formulated by the Social, Psychological, Educational and Criminological Trials Register (SPECTR)[121] were used to develop the search strategy. In the medical field a single electronic database predominates (Medline) and there are resources for identifying studies from other sources (e.g. Cochrane database, national register of clinical trials). In the fields of social policy, sociology, psychology and anthropology the literature is distributed across a range of electronic databases, and unpublished work is seldom registered. Mapping work confirmed that searches for diverse study types, that are neither restricted to RCTs nor to single exposures or outcomes, are complex. In the light of this, two approaches to searching were adopted incorporating both focussed and diffuse searching techniques.

The field was scoped to identify electronic databases where relevant literature might be included. Subsequently, the following 12 databases were searched: Medline, PsychInfo, CINHAL, Sociological abstracts, IBSS, ASSIA, BNI (British Nursing Index) ChildData, Caredata, SIGLE, Dissertation Abstracts (US), and Index to Theses.

Search terms were developed in the initial mapping exercise. In the first instance, a list of words which might refer to growth and size in infants was created by the research team. These terms were used in the mapping searches, and search results were checked for inclusion of potentially relevant studies already known to the research team.   Search terms for papers on views or opinions were developed in a similar way. The mapping in this case was supplemented with searches for literature on breastfeeding. This topic was selected as being better developed than views on infant growth, and it was thought likely that the terminology used would overlap. The databases were searched for papers which included terms relevant to infancy, growth or size and attitudinal or opinion data published since 1978. This date limit was selected in order to seek literature on infant growth within one generation (defined for the purpose of this work as 25 years).   Searched terms included attitudinal or opinion data (including attitud*, focus group$, phenomenolog*). For full search terms used for each database see appendix 12.

Registers of studies and search techniques for qualitative and views studies are not well developed, and there was a potential risk that by using search terms for attitudes and views, studies would be missed.   To broaden the scope, all searches conducted for the review of outcomes were also reviewed for inclusion in the reviews of views of growth. A similar, but no identical, approach has been taken in other systematic reviews combining qualitative and quantitative findings[122-124], although in these cases a single search was conducted to identify both outcomes and views literature using search terms for intervention and population only. In our case searches were separated by outcome to aid screening of abstracts, but we judged that the sum of all the searches was similar to a single search including a range of outcomes (where some 80,000 abstracts were screened) and views literature (where 2694 abstracts were screened). This approach was thorough but time consuming. There also remained the possibility that within the literature for other outcomes, studies of views may have been missed. By considering studies located within in each search 'arm' we hoped to assess the efficacy of this approach. Screening the outcomes searches identified 35 potential papers and from these one paper was included in the final review of lay perspectives. However, this paper also appeared in the Medline search for qualitative and opinions data. In the light of this we consider that our method may have been overcautious, and was not an efficient method for extending our searches for qualitative studies. This highlights the importance of tailoring methods of research reviewing to specific research questions, particularly when reviewing non-trial literature.

Two reviewers screened all abstracts. Any paper selected by either reviewer was screened in full (again by two reviewers) using the following eligibility criteria:

- Infant age: birth until twenty-four months of age, including pre-term babies. Data collected pre-natally was included only when opinions concerning the growth/size after birth or in terms of consequences after birth are included.
- Study focus :opinions, views and attitudes explicitly referring to the body size or growth of infants in any context (e.g. health, feeding)
- Methodology: focus groups, interviews, qualitative analysis of non-interview data, psychometric tools (e.g. attitudinal surveys). Questionnaires and surveys were included only when meeting other criteria. Excluded items were nutritional intake surveys such as food diaries, surveys of child-care behaviour and studies of feeding practices that did not use infant size or growth as a variable. The review sought data on views of growth including those represented in quantitative studies.
- Country – no countries were excluded *a priori*. Reviewers were asked to consider the likely relevance of each study to a UK context, taking into consideration geographically similarity, cultural similarity and countries which have contributed significantly to the UK population. Reviewers were provided with ethnicity data recorded in the 2001 census[125], a list of the most common languages spoken in London schools in 2000[126] and a copy of the designation of developed or developing world countries used in the review of outcomes associated with early growth[10].

Quality criteria were not used to determine eligibility since reviews of qualitative studies[122-124] have found applying existing quality appraisal tools over-restrictive. There are no universally accepted qualitative quality criteria, and qualitative studies are rarely written up with such criteria in mind, unless written for a journal with specific guidelines such as the British Medical Journal. Rigidly applying quality criteria retrospectively is likely to exclude relevant studies that do not fit a particular 'recipe.' Rather than lose data due to the relative novelty of the use of quality guidelines in this field, all studies meeting the inclusion criteria listed above were included and quality appraisal was used at the data synthesis stage. The development of an appropriate method for quality appraisal is described further below.

In total, 2694 abstracts retrieved from searches of electronic databases were screened (36 of which were from searches for the systematic review of outcomes). Of these, 76 full papers were reviewed for inclusion. All papers where reviewers disagreed (n=11) were considered by a third reviewer who independently assessed the paper. Nineteen papers were included in the review.

In order to attempt to locate unpublished work, and studies not appearing in database searches, field experts were contacted to ask for studies they knew of which might be relevant to the review. Additionally, all first authors of included studies were contacted and asked for details of any further relevant studies (published or unpublished). No additional papers were located from these sources.

6.3   Data extraction
Each study was independently data extracted by PL and one of two researchers (LA, TL), one working in another institution. Differences in assessment of study design, methods and quality were compared and resolved by consensus.

Data extraction for this piece of work presented challenges. Since diverse study designs were to be included, but the total number of studies was small it was not considered useful to have different data extraction forms for each study type. Instead the questions asked needed to be flexible enough to allow completion across study types. Seven questions were developed to interrogate the studies for findings relevant to this review:

1. What is healthy growth/size?
2. How important is growth/size to participants?
3. What concepts are used to define healthy growth/size?
4. How do participants assess normal growth/size author?
5. Where does growth lie in priorities for child health?
6. What kind of information influences views/behaviour?
7. Who influences views/behaviour?

An example of a completed data extraction form for one of the included studies is reproduced in appendix 13.

The development of an appropriate methodology for quality appraisal was an important step in this review, but also posed challenges. There is vigorous debate on whether qualitative research (QR) can be subjected to scrutiny using standard quality criteria, or whether this process is contrary to the very nature of qualitative enquiry[127;128]. There has been a proliferation of tools, criteria and guidelines produced to assess quality in QR, for instance Dixon-Woods and colleagues have found more than 100[129]. Following a review of the literature and work by colleagues[130;131] a revised version of a quality checklist produced by Popay[132] was used, supplemented with items suggested by other researchers. During the course of this review a framework for quality appraising qualitative research was produced for the Cabinet Office by the National Centre for Social Research[127]. This is more comprehensive and detailed than previous guidelines, but the utility of this and other tools are still being considered[133]. The quality criteria adopted for this review needed to be flexible enough to enable different study designs (including quantitative designs) to be assessed. Thus studies were scrutinised for the extent to which they enabled an exploration of subjective experience. This constitutes an assessment of fitness for purpose for *this* review rather than a measure of study quality as such.

In qualitative research the richness (or 'thickness') of the data is a crucial measure of the quality of the design. The following criteria were used to assess quality:

1. **Design** (extent to which design is fit for purpose of exploring subjective meanings or experience)
2. **Setting** (is the research design sensitive to the context in which study takes place?)
3. **Participants – sampling strategy** (is the sampling strategy described appropriate for research question?)
4. **Participants – participation/response rate and loss to follow up where appropriate**
5. **Analysis - triangulation** (are different sources of knowledge/understanding compared?)

6. **Analysis - potential for assessing typicality** (what claims are being made about generalisibility to other populations or other bodies of knowledge?)
7. **Relevance to policy** (is the relevance to stakeholders indicated?)

Systematic reviews can identify gaps in primary research. In order to think about why we considered studies poor for our purposes, we considered what the ideal study might have been. For the purpose of this review, an ideal study might have asked about lay (particularly parental) views of the importance and meaning of early growth, with a view to understanding whether growth adaptation is likely to be acceptable or even welcome to parents. This would have been researched using qualitative methods exploring subjective experience, with a robust design and sound execution. Our search strategy did not identify such a study. Instead the topic of growth/size emerged from data from studies with diverse study aims. Thus the first step in extracting study findings was to define the research questions with which to interrogate the data. These questions were derived from the original study protocol.

The study findings were extracted in two forms; the interpretations of the study authors (2nd order findings) and direct data from participants i.e. verbatim quotations (1st order findings). These were extracted separately so pooling of participant data was possible. This also allowed a test of study quality. The highest quality studies presented both kinds of data, in other words they provided both an interpretation of the data and the data themselves to support this interpretation. Thus the data were represented twice in the extraction, once as data (first order), and once as the authors' interpretation of the data (second order). Where authors did not provide sufficient interpretation of the data, or did not provide data to support their interpretations, the findings would only appear once. Moreover the method showed where author interpretations were unsupported by participant data.

The relevance of study findings to each of the review questions was 'read' from the data by the reviewers. The statement *"She will start to gain weight . . . Like a little kid ought to with fat in their cheeks"*[134] was used by one reviewer in answer to the question to "what is healthy growth/size" and by another to "how do participants assess normal growth/size" and it has relevance to both these questions. Where reviewers' views on relevance of findings to review questions differed, we took an inclusive approach rather than settling on a single interpretation.

6.4 Methods for synthesis of review findings

The purpose of the review of studies is to undertake a secondary analysis of findings presented in existing studies with the aim of contributing to both a description of the existing literature and the development of new concepts and hypotheses arising from these findings. Methods for reviewing qualitative studies, and for integrating qualitative with quantitative studies are currently under consideration[129]. The important work by Dixon-Woods and colleagues will aid researchers by developing a common vocabulary, by opening discussion about the strengths and weaknesses of different approaches and by spelling out questions remaining in this field. One question identified for future study is to "compare[s] the results of applying the different methods of synthesis . . . distinguishing the trivial and non-trivial points of divergence between the methods." p.31

There is no consensus on which approach may be appropriate in any particular instance. This review represents the synthesis of results from a relatively small number of studies (n=19) and thus a decision was made that it was feasible to attempt two methods of synthesis; thematic and narrative analyses. This enables a comparison of conclusions drawn by each method, alongside an evaluation of methods.

The strength of narrative approaches is that researchers can explicitly account for heterogeneity in study design, context and quality. These approaches have been particularly successful in synthesising different types of research evidence (e.g. qualitative, quantitative, economic). Examples include Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) review[122-124] or the Hopkins et al[135] review of tobacco use and exposure to tobacco smoke (in each of these, the synthesis was constructed as a series of questions defined by the literature), reviews of ultrasound in pregnancy[136] and communication between health care professionals and patients about prescribing[137]. In the case of the EPPI-Centre review, this involved using barriers and levers for change derived from the review of qualitative literature to structure the sub-groups used for synthesis of quantitative, effectiveness findings. In the case of the study of strategies to reduce exposure to tobacco smoke, the findings from each type of study were grouped by research question – those that aimed to reduce tobacco smoke in the environment, and those that aimed to reduce tobacco use. A potential weakness of these approaches is that they leave the methods by which findings are grouped, described and synthesised to the judgement of the researcher. We rely, therefore, on the quality of the reporting to maintain the principles of transparency that are key to systematic reviewing. In order to carry out this kind of synthesis, the researcher needs to have a broad overview of the area. While this may be considered a potential strength of this approach, it means that replication may be difficult.

Thematic syntheses also enable researchers to review findings from diverse study types, but it may be challenging to make the heterogeneity explicit in the interpretation of the findings. There is an additional drawback in that studies of variable quality are subsumed into new findings, and poorer quality studies given equal weight to studies of better quality. However, this sort of aggregation can also be used to strengthen findings. In quantitative meta-analysis, aggregating studies with small sample sizes or broad confidence intervals can demonstrate significance. When considering qualitative data the equivalent concepts to statistical significance in quantitative studies might be considered to be the richness or 'thickness'[130;131] of the data. Using methods such as thematic synthesis when 'thinner' and 'thicker' studies agree, it may be that we can have more confidence in the representation presented by the thinner studies. In addition, by pooling findings from different participant groups we can consider the extent to which what is reported is a universal, or at least common, experience or one which is specific to those questioned in a particular piece of research. Thematic synthesis is more appropriate to the model of hypothesis generation familiar to qualitative researchers. Thematic syntheses have also been criticised for lacking transparency in their execution, relying as they do on the interpretations of the researchers involved to identify themes.

It is important to point out that the synthesis of qualitative work is controversial; some qualitative researchers see this process as 'context stripping' and thus anathema to the philosophy of qualitative research.

6.4.1 Methods for thematic synthesis of studies

The findings from the studies were synthesised using thematic analysis to develop emerging themes across studies. In order to synthesise studies thematically, findings were collated under the questions used in data extraction (e.g. How important is growth to participants?). Each researcher independently conducted a thematic analysis of the data set. An initial summary was converted into emerging themes, and these themes were then considered alongside each other and the data set to examine relevance, strength and duplication of themes. This process was repeated until each researcher was satisfied with the themes defined. The aim was to construct a number of themes within each question under which all data could be interpreted. These interpretations were then compared by two researchers (LA/PL). Points of agreement and disagreement were explored and resolved by consensus. Although data was extracted from studies using the 7 questions set out above, the questions were combined on the basis of the themes identified to give three broad areas addressed in the thematic synthesis of research findings:

1. Understanding of healthy growth/size (including assessment of growth/size)
2. Concerns about growth/size
3. Influences on views, behaviour, interpretations of growth/size

6.4.2 Methods for narrative synthesis of review findings

A total of 2694 abstracts were screened from searches specific to the review of lay perspectives on early growth, and more than 80,000 abstracts from the review of outcomes associated with early growth. Nineteen papers met the inclusion criteria. Appendix 11, table 1 gives details of designs for all included studies, and summary quality assessments are provided in Table 2.

Given the diversity of studies that contribute to this review, the findings of the studies were grouped according to the characteristics of participant and infant sub-groups. These groupings were derived prior to data extraction from factors considered likely to affect views on infant growth in the research literature. These were:

<u>Participant characteristics</u>

1. Relationship to baby (mother, father, other family member, health professional, no relation)

It was hypothesised that parents will have information other than size on which to base judgements. Those who have less contact with the infant, including health professionals, may be more reliant on physical appearance to judge health.

2. Weight status of participant

It was hypothesised that the weight status of participants (for example over- or underweight) might affect the acceptability of under or overweight in infants and the way in which the meaning of size is interpreted.

3. Ethnicity

It was hypothesised that views of healthy size and growth in infancy were likely to have a cultural component.

Infant characteristics

1. By weight/growth status during infancy (except those born too small or too early)

It was hypothesised that the size of the infant was likely to affect the views of those around them. For example, infants who are very over- or underweight are more likely to be receiving medical care and this may heighten awareness of the importance of their weight.

2. For infants born too small or too early, or who were placed in neonatal intensive care unit (NICU)

It was hypothesised that expectations would differ importantly from expectations of infants who were born at term, appropriately grown and healthy.

3. Feeding practice of infant (breast fed, bottle fed, weaned)

Given the different rates of growth exhibited by breast-fed and bottle fed infants it was hypothesised that type of feeding might affect views on appropriate growth. This relationship is less clear once infants become weaned, but any data on decision to wean with regard to weight would have relevance to this review.

4. Age of infant.

It was hypothesised that views of healthy size would vary according to infant age.

7 Findings of thematic synthesis of studies

7.1 Whose views are included?
In total the views of 3590 individuals from the UK, Canada, Finland and the USA are represented in this review. The sample includes 16 dieticians and 263 public health nurses. The majority of the sample was mothers (n=1948) and only 10 were other family members. There was a substantial sample of other members of the community, including 303 school children, 427 adolescents (both child and adolescent samples from a single study[138;139]) and 816 adults, although of course

some of the latter may have been parents. A total of 276 mothers were recruited from WIC clinics which provide supplemental nutrition for women, infants and children for low income families in the USA. A number of mothers were recruited because their children were showing poor growth, or were high risk because of prematurity or admission to neonatal intensive care (n=212). The remainder of mothers were from general population samples (n=1469).

The age of infants across the studies ranged from newborn to 9.6 years, the high upper ages were included where children across a wide range were sampled, but data were not separated by age. For 598 of the participants the age of the children was not stipulated, but context led us to believe that the relevant infant age group was covered. A further 1517 participants were asked their views regarding babies, but the age of these babies were open to interpretation (e.g. pictures of sitting babies). Six-hundred and eleven of the participants were parents of infants between birth and 3 months of age, and 228 were parents of 6 month old infants. The remaining studies considered growth across a wide age range, 452 participants considered the growth of their infants in the first year of their life, and another 22 in the first 2 years. Forty were parents of 9-13 month olds, 14 of 4-12 month olds, 82 of 2-12.5 month olds and 26 were parents of children aged between 9 months and 9.6 years. Parents were always asked to consider one child in their family. In no instances were the number or age of siblings reported.

7.2   Understanding of healthy size/growth

Notions of what constituted healthy growth/size for participants were dominated by the question of normality. Data considering concepts or definitions of healthy growth/size fell into two themes; the creation of norms and the explanation of difference. The explanation of difference enabled participants to construct a different set of norms when considering infants they knew, for example a family norm. This explains difference from the population as a whole, but crucially enables an interpretation of infant size as "normal" or adaptive within the context of their family. The use of the word "normal" in this discussion is deliberate. It is the term that emerged from the literature and that participants used in our interviews.

There is a second dimension to the consideration of concepts concerning size, and that is the difference between observable and non-observable causes in relation to growth. Observable categories relate to themes where we might easily "see" a cause, for example we can observe food eaten. In contrast non-observable causes cannot be easily seen, for example we have to infer nutritional status from food intake (see table 29, page 115.   In some cases, such as family tendency, there are both observable (we can see the size of family members) and non-observable elements (e.g. genetic heritability of size).

A matrix of themes was created from these two dimensions. These are shown in Table 28 below, with themes relating to each dimension. Data relating to each construct are described below.

Data on assessment of growth largely overlapped with data on observable norm creation themes. In other words ways of *assessing* growth are also ways of *defining* appropriate growth. For example, growth charts are both an assessment

measure, and a definition of norms. Therefore data relating to assessment of growth is incorporated into the observable norm creation category.

Table 28: Understanding of healthy size/growth

	Norm Creation	Explanation of Difference
Observable	Medical comparative (monitoring growth charts) Material comparative (Fitting manufacturer clothing sizes, changes in clothing sizes) Social comparative (looking "normal") Familial comparative	Relation between eating and growth
Non-observable	Inherited/familial characteristics	Parental preference Inherited/familial characteristics Fatalism Medical cause of growth rates Quality of care

Theme: Norm creation/Observable

Six studies contained data directly relating to assessment of size or growth[134;138-142]. Four themes emerged.

Medical comparative (including use of growth charts)

The use of growth charts to assess individual growth and to decide what growth was appropriate was extensively reported. *"I take her to clinic where they measure her height and her weight. They show me a sheet where they record her weight and height and what is the normal height for children her age" (*WIC mother)[134].

In one survey, 81% of parents rated weighing children at home visits as very important or important and 2% as of little importance. In the same survey, 62% of public health nurses rated weighing as very important or important, and 31% of little importance[143]. In another study 85% of mothers with a high risk baby and 92% of

those with a low risk baby gave weighing the baby as the most common reason to attend the health clinic[142].

Material comparative (Fitting manufacturer clothing sizes, changes in clothing sizes)

In one study participants spoke about the use of clothing as a measure of average size and to assess growth in their own child. One mother said; *"I tend to go by clothes size.. . .if they are not fitting in the clothes they should be fitting in, they're not average" (*WIC mother)[134].

Social comparative (looking normal)

Comparison with other children was commonly used to assess growth and define appropriate size. One mother said, *"you just want him to be normal, like everyone else."*[140] and authors report: *"Comparisons made by others between the child's current growth pattern and what was expected of a child of that age"*[140]. The concept of "normal" size and development was key for many parents, particularly for parents of low birth-weight infants[141]. Mothers were looking for signs of 'normalcy' in their pre-term infants, comparing weight and size with normal birth weight infants[141]. Generally there was a preference for infants of mid-range body sizes[138;139].

Familial comparative

As well as comparing with other children in their community, comparisons with family members was an important way to define what was expected of growth and whether growth was "normal". For example one mother said; *"She's just a little below average as far as the children in the family"*[134].

In one study of mothers of children who had previously had poor growth mothers *"compared their children to each other,* [and] *to other children in the extended family"*[134].

Theme: Explanation of difference/Observable

Relation between eating and growth

There was a relationship between feeding and growth for both study authors and participants. Good growth represented good nutrition. Good feeding resulted in good growth, thus if children were appropriately fed participants were more likely to think any size was acceptable. Parents said: *"If they [my children] are overweight, at least I know they're eating"* (low income mother) and *"parents believed that the bigger and faster their children grew, the better their food intake must be, and the better their health"*[144].

Theme: Norm creation/Non-observable

Inherited/familial characteristics

As well as comparing within families, parents felt that size was an inherited characteristics. For example, study authors wrote: *"among mothers who were overweight, there was the assumption that their children were genetically predisposed to be heavy"*[144] and parents that: *"I really do believe it is genes"* or *"I think it's in the blood, in the family"*[134].

Theme: Explanation of difference/Non-observable

Parental preference

One author explicitly looked for differences in ratings of attractiveness of babies of different size and showed that fat infants were rated less positively than average size infants[145].

In a study of the views of mothers who had previously had a low birthweight baby some mothers were reported as having a preference for small size *"either because it meant an easier delivery, or that the child . . . remained a baby for longer"*.[146]

In one study dieticians felt that a preference for bigger babies resulted in over-feeding and over-weight among mothers[144].

Inherited/familial characteristics

Linked to the concept of familial norms, participants also used familial comparisons to explain differences between their child and others: *"you look at me and his father, so he's not gonna be little either."* (low income mother)[144].

Fatalism

Some parents spoke in a way that implied that growth or size was pre-determined and they had little power to change it. One WIC mother said of her child that *"he's finally taking the form he's supposed to have."*[134] This could be interpreted to mean the size her child was pre-determined to have, or simply the 'normal' size for a child of his age. And another that *"I think that her size is out of my hands."* (WIC mother)[134].

Medical cause of growth rates

In the studies that considered children with poor growth[134;140;147] a medical explanation of growth difference was often implicit in findings, for example in the fact that parents wanted medical attention for children who were not growing well[140], and medical explanations of growth were offered by participants[148]. However it is also the case that a cause of growth difference was often assumed within the study aims and designs, making it difficult to know the extent to which this reflects the nature of the studies rather than the views of the participants. This focus on a

medical model may reflect bias in our search strategy, which was more likely to identify health care literature.

Quality of care

The care given to babies was seen as a key factor in determining size *"a heavier baby proved to others that they were effective parents"*[144], *"having a heavy infant or toddler as a mark of both their child's health and their competence as a parent"*. One parent said: *"The care, the diet, parents having a lot of love toward their children makes them grow."* (WIC mother)[134].

7.3   Concerns about size/growth

Two points are important to consider under this heading. Firstly, data were extracted under two headings, the importance of growth and the relative importance of growth within a health context. It was not assumed before data were examined that the meaning of growth would be confined to considerations of health, but rather that it might include other factors such as development or attractiveness. In fact most data reported here discusses growth in relation to health. Even so, we would not assume that this is necessarily the primary meaning attached to growth, rather it is possible that this is an artefact of the body of research reviewed. Searches were probably most efficient in medical databases where search facilities are more sophisticated. Given that the review did not restrict itself to medical databases the fact that the included studies come exclusively from a health context may well mean that growth is little studied outside of this context.

Secondly, growth was largely seen as unimportant in it's own right. Even parents of low-birthweight infants gave low priority to this[139;142]. Instead unhealthy growth was an indication something else was going wrong: *"we were panicked, we knew something was wrong".*[140]. The findings in this section are tentative. If growth is of secondary importance to other factors, then given the health context of the studies the prominence of health concerns and the absence of other types of concern within this category may not be reliable.

One study exploring the experience of families where a child was not thriving showed concern with size among families. We would argue that the findings of this study vary from others because in this case parents could not interpret growth as normal, nor explain difference in the ways described above. This study and the way it fits within themes described are considered in the discussion of findings 12.5 below.

Despite the lack of concern about growth per se, a high value was placed on growth monitoring. As stated in section Noa, one Finnish survey found that most parents (81%) rated weighing children at home visits as important and that this was true for a smaller proportion of public health nurses (62%)[143]. In another study 85% of mothers with a high risk baby and 92% with a low risk baby gave weighing the baby as the most common reason to attend the clinic[142]. These authors suggest that weighing babies is a way to gain legitimacy to attend the clinic[142]. For parents in the included studies, this does not seem to be the case, and parents valued growth monitoring. One reason that this may not be recognised by health professionals relates to the way parents use this information. If as argued here, growth rate or

size can be a response to underlying health status, but is not in itself healthy or unhealthy then a lack of concern for growth might be observed. As before, these can be further categorised into observable and non-observable characteristics, but in this case can also be categorised into those relating to the concerns of the parent, or to the well being of the child. Themes falling into these categories are shown in Table 29.

Table 29: Concerns about size/growth

	Relating to infant	**Relating to parent**
Observable	Medical context Development context (prematurity) Infant behaviour Food Intake	Attractiveness Ease of birth
Non-observable	Nutritional status	Feelings of confidence and competence Perceptions of child needs

Theme: Relating to infant/Observable

Medical context

The importance of growth was often determined in relation to its medical context. There was an association between size and health. If growth was poor, parents were concerned that medical causes were investigated. Parents of infants who were not growing as expected were relieved when this was acknowledged. The longer they had to wait for a referral, the more anxious they became[147]. Eighty-four per cent of mothers referred for poor growth in their children in one study said that the greatest benefit of referral was reassurance that there was no underlying organic cause[147]. For parents whose children were not growing well but where no cause had been established, there was considerable anxiety. The fear of what might be causing the poor growth was making parents *"exhausted and . . . physically ill"*.[140] Given small size, growth itself was not the most important concern for parents; parents of low-birth-weight babies were more worried about keeping babies warm and protecting them from infections than they were about weight[141;146]. Parents felt that this same strategy should be applied by health professionals: *"As my daughter was healthy and full-term, I felt too much was made of her weight. If she was having trouble with her breathing, I could understand the concern. (Mother, dispenser in pharmacy)."*[146] In one study parents agreed most with statements giving medical explanations of poor growth[148].

Developmental context (prematurity)

Parents of preterm infants in one study had more physical than behavioural concerns, including growth, although parents of full term infants were more concerned about growth than those of pre-term infants[149].

Infant behaviour

Parents of low birth-weight babies paid close attention to the results of tests and diagnoses, watching for movement to average or expected results[141]. Parents also used perceptions of child hunger determined by crying to decide upon action (or inaction).

Food intake

Participants and study authors viewed feeding and growth as complementary; good feeding results in good growth. Conversely if infants are appropriately fed parents are less likely to be concerned about size. Thus one low income mother said: *"If they [my children] are overweight, at least I know they're eating"*[144].   There was a general perception that underweight infants were also 'picky' eaters[134;147].

Theme: Relating to infant/Non-observable

Nutritional status

For many mothers, the perceived nutritional value of food was more important than weight gain that might be associated with eating. Authors of one study argue that the fact that half of mothers of children with failure to thrive (FTT) children restricted intake of sweet foods and 30% restricted food they considered unhealthy demonstrates that healthy diet may be considered more important than healthy growth, although they also say this may be a result of food restricting habits in the mother herself[147]. In another study authors point out that mothers *"who breastfeed for longer did not readily associate the rate of infant weight gain with their ability to nourish their infant adequately; they continued to breastfeed, disregarding many of the disparaging remarks made to them."*[150]

Theme: Relation to parent/Observable

Parental preference – attractiveness

Parental preferences for different sizes reported in above (p. 13) is also relevant here, as it is likely that this is related to level of concern relating to different sizes.

Theme: Relating to parent/Non-observable

Feelings of confidence & competence

Poor or unhealthy growth had a significant impact on parents' feelings of confidence and competence. In one study *"61% (16) mothers reported blaming themselves for their child's poor weight gain, feeling they had not done sufficient to ensure adequate weight gain."*[147] This was not always the case however, for example in one study those who breastfed for longer did not associate weight gain with ability

to nourish their infant[151]. In another study test weighing (before and after feeding to demonstrate intake) infants did not affect feelings of competence and confidence in parents of premature infants. Instead they simply increased over time[152].
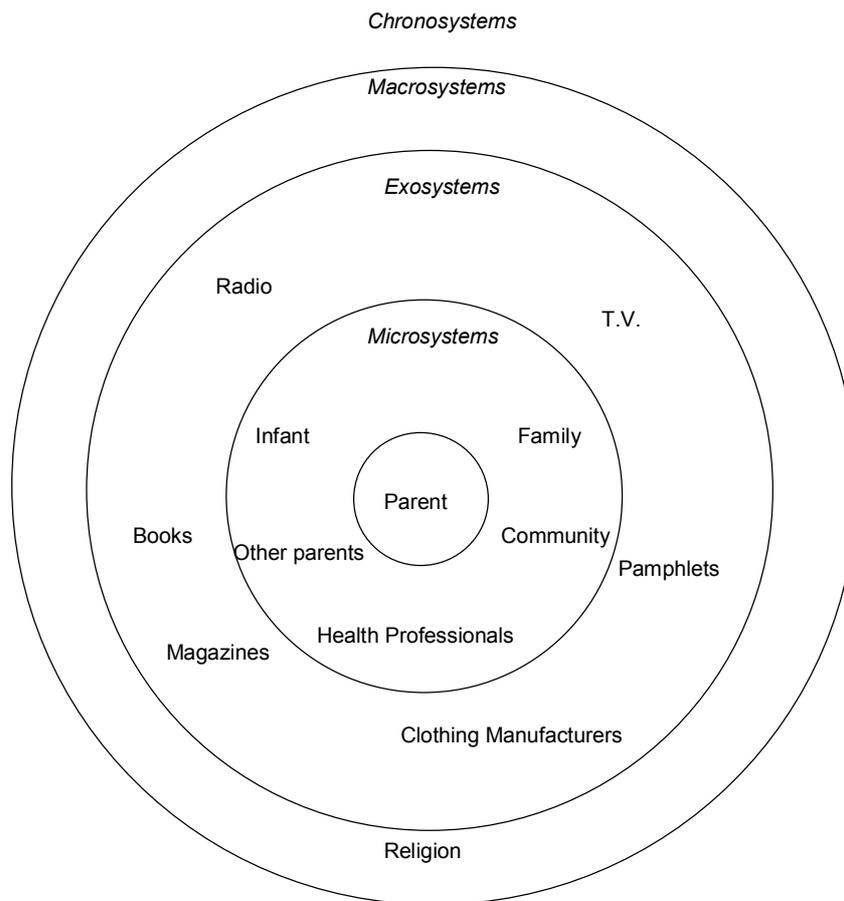
Perceptions of child needs

There was evidence of slightly more concern for underweight and under-eating than overweight and overeating, and mothers were concerned that their babies might be hungry; this was reinforced with the general feeling that parents worried more about under eating and under-weight than over-eating and over-weight[150].

7.4   Influence on views, behaviour, and interpretations of size/growth

In developing themes relating to spheres of influence, reviewers adopted a systems approach after Bronfenbrenner[153]. Bronfenbrenner's original model described ecological effects on children's development. His description of the different systems within which development takes place and the mapping of the ways in which these systems relate to each other has been described as a "framework for disentangling levels of influence"[154] rather than a theory of behaviour. Ecological frameworks have been developed and used to describe family and community influences in a range of settings[154-157]. In Bronfenbrenner's original model five systems are described. Microsystems are the settings within which the individual in the centre of the model has face-to-face relationships. Mesosystem is the collection of groups within the microsystem and the relationships between them, for example the relationship between parents and health professionals. Exosystems are routes by which the individuals or organisations with whom the individual has no direct contact influence the child, for example a government may exert influence through health professionals. Macrosystems are ideologies or belief systems which dictate how elements within the exosystem or the individual at the centre behave, for example published growth norms. Finally the chronosystem represents historical or cultural effects, for example a time-specific response might be expected following the publication of data on health concerns relating to childhood obesity. These systems have been adopted for interpreting spheres of influence in this review, but alternative names have been developed so that the categories make sense to those unfamiliar with this model.

Themes arising from the data in the included studies have been interpreted within this model and are shown below. Some of these themes are derived from data already presented, but following this figure additional data are presented.

Figure 4 Influences of views and behaviour (after Bronfenbrenner)

Chronosystems

Macrosystems

Exosystems

Radio

T.V.

Microsystems

Infant

Family

Parent

Books

Other parents

Community

Pamphlets

Health Professionals

Magazines

Clothing Manufacturers

Religion

Theme: Individual level relationship (microsystems)

Own family

Mothers who had family near-by listened to advice and information from their families. Family members mentioned by mothers included partners and parents (particularly their own mothers), but friends were also important. In one study mothers reported their own mothers as their main source of information[144].

Other parents

Mothers reported influences from wider social networks[152]. They received solicited and unsolicited opinion from relatives and other mothers[141], which was not always welcome[146]: "*it was so difficult to go out and hear negative things about him...you hear these comments over and over and they just feed your own worries and your own doubts...*"(mother of child with faltering growth)[140].

Health professionals

Health professionals were an important source of information and influence. First time mothers used health professionals more often than those with a child[158]. Generally mothers were open to guidance from health professionals[147] and where

there were good relationships, their help was invaluable. Two women in a study where additional midwife support was offered said; *"When I felt worried I contacted my midwife and was reassured*." (office clerk) and "*I don't think I could have got through the pregnancy without the help of the research midwife. . . Her support was invaluable...This kind of support should be available to all women at such a vulnerable time"* (barmaid)[146].

But health professionals could have a negative impact on families: *"The nurses came in and said, 'your baby's still losing weight, you can't feed him', . . . I just dissolved. I often think if I'd been out* [of hospital] *on day 3 it would have been different.."*[151] Authors in one study suggest *"frequent weight checks and visits from health visitors [were] constant reminders of maternal inadequacy in producing a baby that was different from 'normal'."*[140]

Not getting the hoped for information from health professionals was frustrating: *". . . there is nothing physical so it must be your parenting, and then he said, I have to go so we'll talk more about this later.' He left; . . . I was so angry"* (mother of a child with faltering growth)[140], *"I just see them writing down* [weight and height results] *but I don't know how much he has grown. They don't tell me how much he has grown"* (WIC mum)[134]. One study suggested that this support was particularly likely to be lacking for mothers who were not fluent in English

One study concluded that mothers did not ask doctors about their most common concerns, although it is not clear whether the mothers in this sample simply kept these concerns to themselves or sought advice elsewhere[159].

Theme: Organisational level (exosystem)

Pamphlets[158] and books[159] were often used by parents for advice after the birth of their baby as were magazines, T.V. and radio[142]. One study suggests that books were more often used as sources of information than health professionals[159].

Theme: Global influences (macrosystems)

One study mentioned the reliance on God and faith communities for some parents in interpreting their child's medical and growth problems[140].

Theme: Historical or contextual influences (chronosystems)

The studies in this review were all reported in the period 1978-2003. None discussed the historical or cultural context for their findings.

7.5   Summary of findings of thematic synthesis
As described in the findings relating to concepts regarding growth, the predominant concern of participants was normality. This was seen through the creation of norms of growth and models to explain difference. This was conducted across physical, observable characteristics, but included physical unobservable (such as underlying health status) and non physical (such as emotional care) dimensions. Where growth differed from the norm and a plausible explanation could not be found, growth became an important concern for parents. The clearest example of this was

seen in the phenomenological study of parents with children diagnosed with non-organic failure to thrive[140]. For these parents the growth of their children did not conform to any norms they could create, but nor were there explanatory factors to account for this such as family growth patterns or medical cause. *"The constant fear with which families lived was all encompassing. Every aspect of their lives was affected by this emotion".*[140]

In presenting data in this way, we would emphasise that these themes are not an exhaustive account of concepts regarding growth. They are the themes which emerged from the included studies.

8   Findings of narrative synthesis of studies

The narrative synthesis of findings was undertaken by grouping studies according to subgroups. Judgements about the quality of the included studies are shown in Table 2, Appendix 13, and are summarised in this section with regard to the quality of the study *for the purposes of this review* and the methodological quality.

8.1   Findings by participant characteristics

1. Relationship to infants

1a. The views of mothers on early growth

Most of the studies in this review (16/19) explore the views of mothers. This section therefore represents the widest scope of any of the sub-group syntheses. It serves to make explicit the difficulty of summarising across qualitative studies in general, and this group of studies in particular. Whilst all 16 of these studies are concerned with mothers, these mothers vary in terms of the age of their infants, the present health status of their children, the past health status of their infants, their country of residence, their country of origin, income level, socio-economic status (SES), and number of children. In order to consider the breadth of the sample included, a summary of the group characteristics was conducted and is summarised in the Table 30. Formats for reporting of educational level and SES were highly variable, for example use of SES 'scores' with no definition, and in cases where data was included but definitions were not clear it is only stated in the table that this data was reported.

Table 30: Reported characteristics of mothers

Total sample size	1948
Age of mothers	Age range 14-52 years Mean age (from n=768 where age reported)=27.8 years Age unknown n=892
Ethnicity	White = 133 Mexican American=45 African American = 21 Latino = 2 Asian=13 Chinese=8

	Other=14
	Unknown = 1449
Number of children	1 child = 142
	More than 1 child = 615
	Unknown = 928
Income/SES	Low income = 396
	Middle income or higher = 190
	Middle class (social class 1 & 2) =208
	Working class (social class 3,4,5 or other) =482
	Other reporting of SES =112
	Level of education reported=306
Country of residence	USA=762
	Canada = 202
	UK=721
Obstetric complications	Caesarean = 8
	Unknown= 1677
Living with Partner	Yes = 375
	Unknown =1310
Gender of infants	Male = 125
	Female = 116
	Unknown = 1444

It is clear when this data is brought together that very little is known about the background of a large proportion of participants. Sampling was often of poor quality, for example three studies explicitly set out to sample low income groups[134;144;150].   All 3 of these papers used clinics for the US "Special Supplemental Nutrition Program for Women, Infants and Children" (known as WIC) to achieve this. The difficulty of this sampling strategy is that families needed not only to have a low income, but also to register for the WIC programme and attend clinics. In one of these studies[150,] where a comparison group of mothers who did not have low income was sought, 14% of mothers in the non-WIC clinic had incomes low enough to be eligible for WIC programme, thus clinic attendance was not a reliable predictor of income level. Even where we know the proportions of participants from different backgrounds, for example educational background, the studies do not allow comparison between these groups because findings are reported for the entire sample only.

As well as sampling and reporting, difficulties were also experienced due to heterogeneity in study designs and aims. In quantitative studies, methods for assessing and adjusting for heterogeneity are well developed, although there are debates on refining these[160].   In synthesising qualitative research, the field is even less developed. One approach is to include only studies using qualitative methods, population groups or research aims[129;161]. Restricting studies in this way may hamper attempts to answer broader questions.

If we retain the breadth of studies included in this review, practical problems are experienced when presenting the findings of studies. Study descriptions become very long, since quality cannot be easily compared and the context for each study must be described in detail. In our case it is also true that most studies appear in more than one sub-group and assessments must be repeated. This makes the task

of reading the narrative of the review challenging, since the "story" can be concealed by the detail. A strategy was adopted for this review that each study should be described in detail no more than once, and thereafter only issues salient to that subgroup reported. In the case of the views of mothers, where nearly all studies appear, a summary approach to study description was taken. The study summaries below provide a brief description of included studies, a judgement of overall quality, and summary findings. The detailed description of studies is provided in the extended narratives provided within the following sub-groups. In doing this we hoped to maintain the narrative of this important section. As such, it provides an example of way we used a narrative synthesis to comment on common findings.

1. Baughcum et al. 1998[144]. The study sample consisted fo 14 mothers attending WIC clinics in USA with infant aged 12-36 months with a focus on the lack of concern among mothers for overweight in their babies. The study design used was adequate, but the implications of sampling from WIC clinics were not considered. Authors conclude that mothers are more concerned about over- than underweight. Two quotations (not attributable to individuals) are given which state that weight gain is always good, because it means children are eating.

2. Baughcum et al. 2001[150]. Four hundred and fifty-three mothers from health clinics in the USA were asked retrospectively about the growth of their infants in their first year. The focus of this study was overweight in infants and their mothers. The use of closed questionnaires meant that opportunities to explore subjective experiences were limited, and limited reporting (individual questionnaire items were aggregated) means that findings could not be used here.

3. Hewatt & Ellis 1986[151] interviewed 40 mothers in Vancouver with infants aged 9-13 months reflecting on their experience of feeding their babies. The study design was adequate, although retrospective study design may introduce hindsight bias and no details of sampling strategy are reported. Authors state that those mothers who breast fed for longer were less likely to associate infant growth with their ability to nourish their infants

4. May 1997[141] retrospectively interviewed 14 mothers of infants (aged 4-12 months) who had been admitted to NICU. The focus of this study was to consider how these mothers sought out help. The overall study quality was good; the main limitation was the use of NICU staff for initial recruitment. The authors' main conclusion was that mothers of these infants were concerned with the their infants becoming "normal" in size as well as in other aspects of their growth and development.

5. McCann et al 1994[147] used both qualitative and quantitative methods to examine eating habits among mothers of children (aged 9 months to 9.6 years) with faltering growth. Limited reporting made judging the quality of this study impossible. They concluded that there was no evidence to support the view that mothers of children with poor growth had disturbed ideas of infants body size.

6. Pridham 1984 presented two studies within the same paper. They are described here as study a and study b in sections 6 and 7 respectively. Pridham 1984a [159] studied 62 mothers (in the USA) over the first three months of their infants' lives to

determine the types of stress and support they received using parent completed diaries. The study was limited by conversion of respondents' data into statistical information (i.e. no qualitative data were presented), the 10,000 issues logged by parents were converted into 7 categories. Concerns about growth were highest in the 2nd month, although overall growth was not high among concerns for these mothers.

7. Pridham 1984 b [159] used telephone interviews with 22 mothers (in the USA) over the first three months of their infants lives to examine their responses to simulated infant care problems. Authors state that the external source of advice cited most often in relation to feeding and crying was the infant's grandparents. The study was limited by conversion of respondents' data into statistical information (i.e. no qualitative data were presented).

8. Sturm et al 1997[148] studied 132 mothers (in the USA) to determine what causes they attributed to faltering growth. This study had a poor design, with the use of questionnaires with no opportunity to express subjective experience, and no sensitivity to the context in which the study took place. Authors state that participants were more likely to attribute poor growth to medical causes, but responses may have been influenced by the health care setting.

9. Hall et al 2002[152] used a controled trial to assess the effects of weighing infants born too early (30-35 weeks gestation) before and after breastfeeding on feelings of confidence and competence. There were two main criticisms of this study, firstly that the standardised measures used may not have allowed participants to express their views and secondly that the study was introduced into a busy neonatal care unit, and sampling was reliant on the availability of the overstretched staff. The authors main finding was that test weighing did not increase feelings of confidence or competence, but it is conceivable that this is because mothers could be considered to be selectively recruited. They also found that all mothers became more confident over time.

10. Rajan & Oakley 1990[146] sent questionnaires to 467 mothers (in the UK) who had had a low birth weight baby 6 weeks after the birth of a subsequent infant. Some qualitative data was included in this study, although largely quantitative data is reported. This study followed an intervention study, and the implications of this are not discussed despite the fact the quotations reported refer to the importance of the study nurse for these mothers. Overall authors concluded that for most mothers the previous birth of a baby born too small did not result in high rates of concern for low weight in their new infants. Most did not see low birth weight as a problem, and felt that there was too much focus on their babies' small size.

11. Reifsinder, Allan & Percy 2000[134] interviewed 22 mothers using WIC clinics about their explanations for poor growth. Although interviews provide opportunities for mothers to talk about their experiences freely, the use of nurses as interviewers was not taken into account, particularly given that some of the respondents were mothers whose infants had experienced poor growth. Mothers used growth charts to monitor growth, as well as comparisons to other children and clothing. Authors state that eating habits were used as explanations for poor growth as were

genetics, and participants would search for family members with a similar body type to their infant.

12. Sherratt, Johnson & Holmes 1991[142] questionned 228 mothers of 6 month old infants in the UK to compare the concerns and service use of those who were "high risk" at birth and those who were not. The study methodology was poorly reported, and likely to be of poor quality as assessed here. It was not clear whether open-ended questions were used. Findings displayed a contradiction in parental behaviour, where growth did not worry parents very much, but 85% said that weighing their infant was the main reason to attend health clinics.

13. Thomlinson 2002[140] interviewed 11 mothers, plus other family members (in Canada) about their experience of having a child with faltering growth. This was a high quality qualitative study, although reporting was at best adequate; for example the age of the children was not reported. Among these mothers poor growth was a large concern and they were very sensitive to the criticisms by others.

14. Smith 1989[158] sent questionnaires 4 weeks postpartum to 41 parents in Canada (although it is not clear if the entire sample was from Canada) to compare the concerns of first time mothers with mothers with more than one child. The study methodology was poorly reported, but probably of poor quality as assessed here and relied on good literacy among mothers questioned. Growth was low among concerns for mothers in this sample.

15. Kramer et al 1983[162] used questionnaires to study 50 mothers of newborns (in Canada) about factors affecting obesity in infants under 2 years old. The study methodology allowed some opportunity to explore subjective experiences, but there are likely to be difficulties eliciting responses from women who have just given birth. Authors concluded that older mothers preferred leaner infants, but no conclusions about views on causes of obesity are given.

16. Brown 1981[149] used interviews to ask about the concerns of 93 mothers (in the USA) during the first 9 months of their infants lives, comparing those whose babies were born early to others. The interview data appears to have been converted into limited categories by researchers, but little detail of study methodology is given. Growth was an infrequent concern among these mothers.

From this sample of studies it is challenging to make summary conclusions. We would note that growth/size are a concern for parents, particularly achieving average or normal growth. Mothers use a variety of sources to define norms, including growth charts, clothing and familial patterns of growth/size. There is evidence of concern for underweight, but evidence about overweight is unclear.

1b. The views of other family members on early growth

One paper included the views of family members other than mothers on infant growth[140]. A second refers to the views of parents which may include fathers though since this was not made explicit the findings of this paper are not included here [142]. Thomlinson conducted a qualitative study of 12 families of children with failure to thrive and included interviews with 11

mothers, 1 stepmother, 6 fathers and 3 grandmothers in Canada [140]. The study was of acceptable quality, providing 'thick' data and with adequate consideration to sampling and the limitations of generalisability. The author stated that the families were ethnically and socially diverse, but did not provide data to support this. Thomlinson considered the family to be the unit of concern so findings were not presented according to participant type, nor were quotations attributable to individuals. For the purpose of this review it is therefore impossible to determine whether participants other than mothers had different views

1c. The views of health professionals on early growth

Two papers report the views of health professionals in relation to the growth of babies and infants [143;144]. Baughcum, Burklow et al spoke to 16 dieticians attached to 7 WIC clinics (which provided supplemental nutrition for low income women, infants and children for families) in Kentucky, USA about their understanding of the opinion and behaviours of women attending WIC clinics[144]. Vehvilainen-Julkunen sent questionnaires to 263 public health nurses from across Finland concerning home visits to families with new babies[143].

Baughcum, Burklow et al sampled 16 of 17 dieticians attached to 7 nutrition clinics for women on low incomes in the USA. This high level of uptake was achieved through regular team meetings in which the dieticians were interviewed as groups in their work teams. No data was reported on the characteristics of the dieticians except their gender, (all were female), and there was no data on how the 7 clinics were selected or whether they represent all clinics in the study area. Dieticians were asked about "child feeding practices and attitudes to obesity among mothers", for example "Is it a concern for parents if their child is overweight?". The study design allowed for 'thick' data to be collected, which is of adequate quality. The study conclusions were limited because the authors showed no evidence of reflexivity in practice. In the case of the data from dieticians, the views of the dieticians are presented as representative of views of mothers. There is no discussion of potential differences between the views of mothers, and the perceptions of mothers by dieticians. Dieticians were not asked to comment on what they themselves viewed as appropriate or healthy growth, or how they felt about overweight in children. Dieticians in this study did not feel that overweight was an important issue for parents. They offered two explanations for this; firstly that parents felt bigger size was more appropriate either because of family factors or because bigger babies were preferable *"they [mothers] would rather see a overweight than a skinny child"* (dietician). Secondly, because growth was low on the list of priorities for parents:

> *"I think unless a family member. . had a medical problem associated with weight, they really don't see that there is a need for any type of intervention or to do anything. The families have so many things that they are dealing with on a day-to-day basis that to prioritise [weight] for a child who is obviously healthy, happy, can do anything they want*

to do…. The overweight child is not a problem compared to a lot of other issues they are dealing with" (dietician).

Dieticians felt that their advice was not followed, particularly when other family members gave conflicting advice to parents:

"sometimes we're battling ….other family members giving advice, and so we see a 2-week-old child receiving cereal from the bottle…When we see patients prenatally, we kind of address the fact that infants don't need anything in their diets [but milk or formula] *until at least 4 months if not longer. We're even battling that* [early feeding or cereal] *sometimes before the baby's even here because it's something you see quite often"*(dietician).

These views may have been present among dieticians in this study, but may not reflect the views of all dieticians. Even within this study the context may have made it difficult for individuals to dissent from the view presented as common. The views may not have been an accurate reflection of those belonging to the mothers in these clinics, since they were reported at one remove, moreover they reflected a degree of antagonism between dieticians and mothers.

Vehvilainen-Julkunen sent questionnaires to public health nurses across Finland asking about the functions and meanings of home visits provided to families post-natally[143]. Questionnaires were sent to random samples of 100 prenatal nurses, 100 child welfare nurses and 100 non-specialist nurses attached to child welfare clinics. The process of random selection was not described. A good response rate was achieved (87.67%) although it was not reported whether response rates varied across nurse types. Very little detail was reported in this study regarding the data collected, no sample questions or topic guides were given and it is not clear whether only closed questions were used or whether some open-ended questions were included. Thus although the sampling strategy appears thorough and well-described, and the response rate was good, the overall quality of the study cannot be evaluated. Of interest to this review is data reported on the perceived importance to nurses of weighing infants during home visits, and the extent to which nurses saw themselves as an information resource for parents. Twenty-two percent of nurses thought that weighing the child was very important, 40% rated it as important, 31% as of little importance, 2% of very little importance and 3% didn't know. This suggested that, although most nurses (62%) thought weighing the child was important, a third saw little or no value in weighing. Nurses were asked to state which elements of the home visit were very important (it is unclear whether categories were coded from open-ended questions or whether they were choices presented to respondents, but more than one answer was possible). Fifty-five percent felt that the general examination of the newborn infant was very important, 49% the general examination of the mother, 40% general information giving, 36% examining the child's skin and 22% weighing the infant. The author writes

> *"Public health nurses felt that the most important function of home visits has to do with supporting and encouraging parents in child care and with the sense of security they can instil in families."*

The data presented here suggest that they viewed their examination of the infant and mother as a more important element of the visits. A good proportion of the nurses did not place a high value on monitoring the growth of the infant. The 'thinness' of the data makes it difficult to draw conclusions about whether nurses did not consider size important or whether they did not consider weighing the infant necessary to estimate rate of growth.

The data available from these studies do not allow us to make conclusions about the views of health professionals with regard to infant growth.

1d. The views of unrelated others on early growth

Three papers describing 2 studies considered the views of general community samples where recruitment was unrelated to relationship to infants or early growth[138;139;145]. The study by Birgeneau included 135 non-parents (59% of total sample) in order to ascertain whether parents and non-parents viewed infant size differently[145]. The study by Rand & Wright is likely to have included parents, but this data was not recorded. Both of these studies used pictures of infants of different sizes which participants were asked to rate on a number of characteristics and both took place in the USA.

Rand & Wright sampled a broad age range; 303 elementary school children, 427 adolescents, 261 university students, 326 middle-aged adults. The total sample size was 1317. The sampling strategies were unclear. Young adults were recruited from the university where the authors worked, but no data on participant recruitment were given for any other group. The authors acknowledge that the sample did not include older adults (the oldest participant was 55) and was predominantly white (proportions across age groups 74-82%) and female (proportions across age groups 56-66%). There were no data on participation rates, or local population demographics so we cannot assume that this sample is representative. Participants were presented with an array of 9 line-drawings across a spectrum of very underweight to very overweight individuals across different ages (one male and one female of each age), including infants. In each case participants were given statements and asked to select drawing or drawings on the basis of prompts for example:

> *"when you see people at the school, at the mall, or on the beach, what body sizes do you think look OK? Circle the number under all body sizes you think are acceptable in addition to the size you like best."*

This was used to identify socially acceptable body sizes. The second paper (Rand & Wright, 2001) presents findings from the same study analysed according to the proportion of participants selecting thinner female than male body sizes[139]. Findings were presented as mean and standard deviations of body size chosen, where 1 was very underweight, 9 was very overweight

and 5 was the mid-sized body. Very few differences were found between age groups in their assessments of infant size. The range of ideal baby body size was very similar across groups, mean body size for school children was 4 (+/- 0.8), for adolescents 4.3 (+/- 0.9), for young adults 4.4 (+/- 0.8), and for middle aged adults 4.2 (+/- 0.8). The authors stated that:

> *"the body sizes of infants received a greater latitude of endorsement than any other targeted group. Baby body size 3,4,5,and 6 were approved by a majority of each subject group".*

They also stated that "*In regard to babies 32% of young adults preferred thinner females, compared to 12% of children, 125 of middle age and 30% of adolescents*". However, because the reverse data (the number of participants preferring thinner males) is not reported, it is not possible to judge whether thinner females were preferred more often. The data presented in these papers is 'thin'. Participants were given little opportunity to express their views. We conclude that participants in this study found a greater range of body sizes acceptable in infants than in other ages.

The study by Birgeneau was submitted as part of a PhD and used photographs of infants taken from commercial magazines to examine adults' social judgments of babies of different sizes[145]. The recruitment strategy was not described. However, the author states that efforts were made to sample equally from each gender (129 females and 100 males were included), parents and non parents (41% parents and 59% non parents were included) and white and black Americans (41% African American, 58% Caucasian, 1% missing data) from among university students and staff and their friends and colleagues. Photographs of 10 gender-neutral infants were presented to participants, but data on responses to only 2 of the pictures are described. These two were pictures that had been most commonly identified as "average" and "fat" in a pilot of the study material. No details are given on how data collection took place. Participants were asked to report their height and weight, and BMI (Body Mass Index) was estimated using this data. Participants were asked to rate babies on a "series of personality and physical appearance characteristics" on a 7 point Likert scale. Adjectives used were categorised by the author as health-related characteristics (e.g. strong, active, healthy, sickly), weight-related characteristics (e.g. fat, overweight), personality and social characteristics (attractive, playful, kind, likeable, maladjusted, clumsy, greedy) and care of infant "(has a good mother, has a good father, well-cared for, neglected). A full list of terms offered by the author is given in the appendix including terms such as chunky, cuddly, pudgy, skinny, well-fed, malnourished. The infants were labelled with either male or female names and the order of presentation varied so that equal numbers of participants saw each infant with a male and female name. The overall quality of the study was poor. There is a lack of data concerning the appropriateness of the terms selected by the author. No evidence is presented regarding whether participants would agree with the authors' categorisation of terms, for example would participants agree that attractiveness and personality should appear in a single category? A complete list of terms provided and their categorisation is not provided, nor is it made clear why the data from only two infants were presented, or whether they were typical of views of other infants. The author makes conclusions about perceptions of the fat infant, but

no details are reported concerning thin infants, so we cannot know if the important characteristic is 'fatness' or 'other than average'. The author reported that:

> *"Every subject group: women, men; Caucasians, African-Americans, parents, non-parents, and every combination of these groups rated FAT infants* [sic] *significantly less positively (worse socially) then the AVERAGE infants."*

The author reported that gender of the infants did not change perceptions, nor did the estimated BMI of participants change perceptions of infant sociability or heaviness. We would not recommend drawing conclusions regarding perceptions of infant size from this study.

It would seem that unrelated members of the public (including children) tend to prefer infants of mid-range body sizes, but the evidence to support this is thin.

2. Weight status of participants

None of the included studies recruited participants according to participant over- or underweight. One study reported findings according to the weight status of mothers included in the sample[150]. Baughcum, Powers et al used questionnaires to 258 mothers attending health clinics, either WIC or private paediatric clinics. The findings of the questionnaire were compared to weight status of mothers. The authors of this study state that concern about under-eating and being underweight was higher among obese mothers. However it was not found that maternal overweight predicted child overweight, so there was no evidence that this anxiety among over-weight mothers resulted in different behaviour. It was also found that mothers of overweight children were more concerned about over-eating and overweight, suggesting that they were aware of the weight of their children. The structured nature of the questionnaire, and the restrictions in reporting mean that this is a 'thin' study with limited potential for exploring subjective experience. The questionnaires were completed in health clinics, which may have affected mothers' responses.

3. Ethnicity of participants.

Across the group of 19 studies there is a predominance of white, English speaking samples and findings should be considered in this light.

In all but one of the studies findings were not compared between participants of different ethnicities nor did studies purposefully sample by ethnicity. One study compared data between Mexican-American and Anglo-American mothers, but sampling was not carried out on the basis of ethnicity[134]. The sample recruited was predominantly Mexican-American, and participation rates were low (56%). Mothers who moved out of the study area were excluded. For these reasons we consider it inappropriate to consider this sample as representative of Mexican-American mothers.

8.2   Findings by infant characteristics

1. Weight/growth status during infancy (excluding infants born too small or too early)

Four studies recruited participants according to the weight status of their children. Reifsnider et al studied 22 mothers in central Texas, 13 of whom had taken part in a previous study for children with poor growth[134]. Thomlinson interviewed 12 families in Canada (including 21 participants) of children who had been diagnosed with failure to thrive at least three months previously[140]. Sturm used attitudinal ratings scales to determine causal explanations of poor growth among 50 mothers of infants with poor growth, and 82 mothers of healthy babies in the USA[148]. McCann et al interviewed 26 mothers of children with poor weight gain or small size sufficient to be referred to a consultant in Oxfordshire, UK[147].

Two more studies sought out views according to the weight status of infants. Rand & Wright collected data from participants across a wide range of ages, considering pictures of infants of different weights[138;139]. Similarly Birgenaeu explored the views of 229 adults of infants of different weights[145].

Thus we can ask two questions of these data; how do families of infants with poor weight gain view growth, and does the size of infants affected the views of unrelated others? The latter question is dealt with in section PCa.4 above, on comparing views of different sized babies. The experiences of families with poor growth will be considered here.

The study by McCann et al aimed to consider whether disturbed views of eating and body shape were more common among mothers of children with non-organic failure to thrive[147]. The quality of the data collected was mixed, and the qualitative data is described in scant detail. The recruitment and sampling strategy is not clear, and it is the view of the reviewers that it is likely that the mothers interviewed in this study were in fact a sub-sample of a group of mothers from a study exploring the prevalence of eating disorders among parents of failure to thrive children. It should also be noted that the majority of children in this sample were probably outside the range of this review, the children were aged between 9 months and 9.6 years with a mean age of 3.8 years. In this study although all the mothers interviewed had children who were underweight, 58% rated their children as normal or only slightly underweight. However, mothers were concerned about their child's poor growth. Eight said they were relieved when they learnt their child was to be referred and 22 (84%) were reassured when they found that no underlying organic cause of the poor growth had been identified. Mothers found other explanations of poor growth, 16 (61%) said they blamed themselves, but 16 also said that their child was a "finicky" eater. The authors highlight the fact that half of the mothers reported restricting intake of sweet foods, and 30% restricting foods they felt were unhealthy. They suggest that this implies that healthy eating was more important than weight gain for these mothers. They also state that this may reflect food restricting habits in the mother herself, but given that these mothers did not score higher on the assessment of eating disorders there is no evidence to support this statement.

Discussion of findings from the study by Sturm are limited by lack of reporting detail of the study findings. Scores on the attitudinal scales used are only reported in subgroups, rather than responses to individual questions[148]. The subgroups were determined using cluster analysis of responses, so we can only state that responses within these groups were similar to each other. Participants were asked to respond to vignettes reporting the causes of growth deficiency. Authors report that participants were more likely to attribute growth deficiency to medical causes and least likely to attribute them to parenting causes. The data were collected in the waiting room of health clinics, which may have affected data quality. Lack of confidentiality may have affected parents' willingness to discuss parenting issues, and the setting may have been more likely to invite responses about medical causes. Responses of parents of children with poor growth were not differentiated from those of parents with unproblematic growth so we do not know if they differed.

The study by Reifsnider et al interviewed mothers on low-incomes who had previously taken part in a nurse-led intervention for parents of children with poor growth[134]. The age of the children was between 22 and 51 months at the time of the interview, although parents were asked to consider growth in the previous 2 years. There were some problems with the sampling of this study. The sample was atypical in ways not addressed by the study team, for example they only interviewed those who were still in the study area two years after the intervention despite the fact that this excluded nearly half the original sample. Nor do they address the fact that 13 participants had taken part in a nurse-led intervention, and the interviews were undertaken by nurses. This may have affected responses. Details of the interview process were not reported. Quotations were not attributable to individuals, so we do not know which responses come from the group with previous poor growth in their children. In this sample mothers were concerned about their children's' growth. They monitored growth using clothing sizes, growth charts at health clinics, measuring growth at home and comparing to other children. Authors were surprised how many mothers discussed close monitoring of growth using growth charts, e.g. a "struggle to rise above the 5th percentile". They felt that children should be growing and filling out. Their models for explaining growth rate were explored and fell into the following categories: good diets had a positive impact on growth; good care taking had a positive impact on growth; size was genetically determined so they expected their children to be a similar size to them but when they were not, mothers often found another member of the extended family who was a similar body shape. Particularly among those who closely monitored growth charts there was knowledge of growth "similar to the biomedical model"; and finally the growth of the child was a problem for the parent and not the child.

The final study in this group by Thomlinson was conducted in Canada with whole families of children experiencing failure to thrive[140]. This study included 11 mothers, 1 stepmother, 6 fathers and 3 grandmothers. No data was given on the ethnicity of those interviewed. This was a qualitative study with a sound design and adequate reporting except that the age of children involved was not stated. This is a problem for this review of studies relating

to growth between birth and 2 years. In addition, quotations were not attributable to individuals. The main findings were that families of children with poor growth were very anxious about this, and the lack of diagnosis for the cause was particularly worrying. There was an overriding concern that the children should be "normal" which is expressed in the following:

". . you feel like you are asking for the world whenever you say you do not want him to be a genius, you don't want him to be seven feet tall, you just want him to be normal, like everyone else."

Families made constant comparisons between their child and others, and were very sensitive to negative comments about their child. The acknowledgement of their problems by health professionals, and the sense that they were part of a team made families feel supported. Families also gained strength through faith in God and their faith community. Families were concerned with their children's' reaction to eating, and felt that there were few professionals who paid attention to their anxieties.

These 4 studies make it clear that parents and families of children with poor growth are acutely aware of growth as a problem. They monitor growth and discuss growth with others. What they desire is to see "normal" growth in their child, although they will sometimes look for ways that they can interpret growth as normal (for example finding members of the extended family that are of similar body shape). The most common method of assessing size through the studies is by comparison with others, although the use of growth charts and physical measurement is also important.

2. Infants born too small or too early, or admitted to NICU

Four studies considered the experience of mothers of infants born too small or too early. One by May was a qualitative study, with data from interviews with 14 mothers whose babies had been admitted to neonatal intensive care at birth[141]. The second by Hall was a randomised controlled trial[152]. In this study, 60 mothers of infants born at 30-35 weeks gestation, who intended to breastfeed, were assessed on measures of confidence and competence. The intervention was test weighing before and after breast-feeding, and the only outcomes reported were standardised assessments of maternal perceptions of confidence and competence. Diaries were kept by mothers, but no data from these were reported in this paper. Both studies took place in North America. The third study by Sherratt took place in Buckinghamshire, UK and used a questionnaire to compare the views of parents with 'low' risk (derived by taking the infant next recorded as 'high' risk infants in health visitor records) and 'high' risk babies (born too small or admission to special care nursery) on postnatal services provided in the first 6 months[142]. Brown (1981) considered the concerns of mothers of high risk (pre-term or <2.5kgs at birth) and full term infants in the first 9 months in Los Angeles, USA.

Drawing conclusions from Hall's study is problematic. Recruitment and test weighing were undertaken by NICU nurses, adding to their workload. Women were only recruited when the unit was relatively quiet and there was potential for nurses to selectively recruit. It also appears (although this is not stated) that the sample

was drawn from a single NICU. The sample was predominantly white (although again the exact ethnic mix was not reported) and included mothers of 15 sets of twins. Loss to follow up was inadequately dealt with; data from mothers who moved out of the study area within the 4 week follow up period and those who stopped breastfeeding were excluded from analysis. It was not reported whether the 4 women excluded because they stopped breast-feeding came from the test-weighed or non-test weighed group. The study found that test weighing had no significant effect on responses to the Maternal Confidence Questionnaire, or the Parental Sense of Competence Measure. Considering the thinness of the data collected, and the methodological problems with the study, this evidence that test weighing does not have an impact on mothers' perceptions of confidence and competence among pre-term infants is not robust.

The study by May considering the experience of mothers of high-risk infants was a good quality qualitative study[141]. It was considered good on all quality criteria except one, and the study design provided 'thick' data. The methodological concern related to recruitment. The participation rate was not reported, nor is it made clear what proportion of mothers whose infants were admitted to NICU were approached to take part. We know that the sample included 10 white mothers and 4 Latino mothers, that infants were born between 24th and 34th week of pregnancy, and that half of the sample were first time mothers, but quotations from respondents are not attributable to individuals. The study concluded that the main concern of these mothers was to look for evidence that their babies were becoming "normal" reporting one mother saying *"like if you had a normal baby"* [interview 22]. The most important way mothers did this was by comparing babies both to non low-birth-weight babies and to other low-birth-weight babies. Infant weight and size were markers of this, but so were heart monitors, changes in health, technology dependence, signs of development (including crying, feeding, and interaction with others) and changes in "caregiving needs". Receiving good news from physicians was important for mothers and progess in any of these areas made mothers feel more relaxed, comfortable and confident. May reports that "mothers were open to guidance from health professionals" and that they "received solicited and unsolicited opinion from relatives, other mothers, and health professionals." The conclusions from this study are not assumed to be generalisable beyond populations of mothers of preterm low birth-weight infants.

The third study considered views of health visiting services in the UK, although this is not explicitly stated as a study aim[142]. The data were collected using a questionnaire, which was not included in the study report, and it was not made clear whether questions were open-ended or closed. Recruitment introduced potential bias since sampling used health visitor records and links to health visitors was made clear, although part of the study involved participants commenting on their experience of health visiting. Since data collection is not described in detail it is difficult to judge the extent to which this may have affected participants' views. Findings are reported for "parents", but at some points in the report, findings only refer to the experience of mothers. The extent to which fathers' views are included is unclear. The views of 128 parents (58% first time parents) of high-risk infants are reported. Other participant characteristics were not differentiated by high and low risk parents. The study group was 98% English speaking, predominantly of social class 3 (77%) and mostly had a partner living at home (96%). When asked what

worried them most 7% of high risk parents said growth. This was less than for 'medical complaint', 'nothing', specific concerns (e.g. SIDS), poor feeding, general negative comments or sleep problems. Across the whole sample, 85% of high-risk parents said their main reason for attending the child health clinic was to weigh their babies. This was lower than for low risk parents (92%) although the difference was not tested for statistical significance. Authors conclude that growth was not important for parents, and that there were no differences in responses between low-risk and high-risk parents. Without more information on data collected and in the light of difficulties with recruitment, the findings of this study cannot be considered robust.

Brown considered the concerns of mothers of high risk (pre-term or <2.5kgs at birth) and full term infants in the first 9 months in Los Angeles, USA[149]. The aim of this study was to examine "concerns about physical and behavioural development over the first 9 months of life" using semi-structured interviews with mothers. Interviews were coded and only frequency of concerns raised were reported, thus although categories may have been participant driven they have been interpreted by the researcher and no quotations are presented. Limitations in the reporting of sampling and data collection mean that it is difficult to assess the overall quality of this study. The author concluded that concerns about physical growth increased over time for mothers of pre-term infants. For parents of pre-term infants, concerns about physical health (e.g. upper respiratory tract infections, skin disorders) were more common than behavioural concerns over the entire period. Although physical concerns were raised for small babies, size and growth were not high among these concerns.

It is difficult to draw overall conclusions about the views of parents of infants born at some risk (too small, too early or admitted to NICU) from these studies. It seems that monitoring growth was important for parents, but growth and size were low among parental concerns. Comparisons between babies were particularly important and parents were keen to find signs that babies were becoming 'normal'.

3. Infant feeding (breast feeding mothers, bottle feeding mothers, mothers of weaned infants)

Two studies set out to sample mothers who breastfed their babies. The first, by Hall, is a study of the effects of test weighing on maintenance of breastfeeding among preterm infants in Canada[152]. The second study by Hewat and Ellis, also set in Canada, set out to explore women's experience of breast-feeding in a qualitative study[151].

As described above, drawing conclusions from Hall's study is problematic. Considering the thinness of the data collected, and the methodological problems with the study, the evidence that test weighing does not have an impact on breastfeeding or mothers' perceptions of confidence and competence is not robust.

There are also problems with the study by Hewat & Ellis[151]. Authors set out to discuss breastfeeding retrospectively among mothers of 9-13 month old infants whom they had originally intended to breastfeed. The sampling strategy and procedure is not adequately described, except that they were recruited according to

intention to breast-feed shortly after birth and followed up 9-13 months later. All mothers had partners, although it is not clear whether this was an inclusion criterion. Mothers were categorised according to duration of breast-feeding (short/long) although there is no discussion of how cut-offs for duration were determined. Matched pairs of short/long breast feeders were then created, although the strategy for achieving this is unclear. Years of residency in Canada was included, but not ethnicity. The authors concluded that mothers who breastfed for longer were less likely to associate the rate of infant growth with their ability to nourish their babies. These mothers were able to disregard "disparaging remarks made to them" (although evidence of such remarks is not reported). In contrast one woman is quoted who felt that the pressure from nurses in hospital to see weight gain resulted in her decision to stop breast-feeding. They stated that other characteristics of those who breastfed for shorter durations included "rigidity in routine", "inability to relax", "less father support", and "problematic interpretation of sibling behaviour". Thus authors imply that growth was more important for those who did not continue to breast feed. It is also plausible that those who stopped breastfeeding had more difficult lives, or were more inclined to report problems.

None of the studies report data separately for bottle-fed babies or weaned babies. We cannot make any clear conclusions about the views of mothers who intend to breastfeed concerning interpretations of infant growth.

4. *Age of infants*

Findings are recorded by age of infant (0-3 months, 3-6 months, 6-12 months, 12-24 months)

0-3 months

Six studies considered views on infants between birth and 3 months. One of these concerned pre term infants in the first weeks after discharge from hospital, and another looked at pre-term infants in the first 9 months after birth[149;152]. The latter asked parents to complete questionnaires at 0-3 months, 4-6 months and 7-9 months and results therefore appear in each of the age-determined sub-groups. Brief summaries of each of the studies follows.

1. Pridham used diaries of infant care taking tasks and concerns in first 90 days following birth[159]. The data from these diaries were coded and the frequency of concerns and tasks was used as data in this study. The authors noted that mothers were most concerned with growth in the 2nd month. Mothers used a variety of sources of information including partners, books, family and friends. Health professionals were used, but not often for issues regarding growth and development. The most common concerns that mothers had were not raised with health professionals.
2. Hall et al conducted a trial of test-weighing before and after breastfeeding in young infants to examine whether it would increase maternal feelings of confidence and competence[152]. Test weighing did not have an effect

but authors noted that on average all mothers' feelings of confidence and competence increased over time.

3. Rajan & Oakley studied the views of mothers who had previously had low birthweight babies on birth weight in their new infants (aged 6 weeks)[146]. Two thirds of the women did not see low birth weight as a problem in itself, particularly those whose babies had not had medical complications. For those who had felt anxious with their previous baby, or whose babies had been unwell, additional support from health professionals (a research midwife) or their families was very important. Weight gain was low among concerns for babies who were born small; the biggest concern was keeping the baby warm followed by fears that the baby would get an infection or become ill. Many women believed the main problem was prematurity rather than under-weight. The mothers felt that they received undue negative attention; some felt monitored by health visitors and that "too much" was made of their baby's weight. Comparisons with other babies, unsolicited advice, and the fear by those close to them of their small babies were disliked.

4. Smith used questionnaires to examine the main concerns of mothers one month post delivery[158]. Growth was low among concern for both primaparous and multiparous mothers (15th out of 21 concerns listed). However, about one third of mothers (7/19 primaparae and 6/22 multiparae) included growth among their concerns.

5. Kramer et al used attitudinal assessment tools to assess ideas of ideal infant body size and attitudes toward feeding among mothers of babies 1-3 days old[162]. Authors concluded that older mothers preferred leaner infants, although no relationship was demonstrated between preference and attitudes to feeding at this early age.

6. Brown interviewed mothers about their concerns during the first 9 months after birth in premature and full term infants[149]. In the first three months growth was low among concerns for mothers of pre-term and full infants. Only 3 out of 73 mothers of pre-term infants were concerned about growth at all in the first three months, and 5 out of 20 mothers of full term infants.

3-6 months

Two studies considered views towards growth or size in infants aged 3-6 months. Both of these studies included infants who had a low birth weight.

1. As reported above, Brown followed parents through the first 9 months after birth[149]. At 4-6 months, growth remained low among concerns for parents of both pre-term and full term infants.

2. Sherratt et al used questionnaires to study the concerns of mothers of 6 month old infants and their views of service provision among healthy full-term, and high risk at birth infants[142]. Growth was very low among worries about babies, 7% of high risk and 4% of low risk mothers were worried about weight gain. Authors comment that most parents felt that their baby was growing well by 6 months. However 85% of high risk and 92% of low risk mothers said that weighing their baby was the main

reason for attending child health clinics. Authors suggest that weighing babies legitimise visits to the health clinic for other, unstated, reasons.

6-12 months

Five studies considered views towards infants aged 6-12 months. Of these, in two cases no age was specified, but pictures of sitting babies were presented. Reviewers placed these images within the 6-12 month age range.

1. May interviewed mothers of infants between 4 and 12 months (mean age 9 months) who had been admitted to the neonatal intensive care unit at birth[141]. The study aimed to examine how these mothers went about seeking help. The author concluded that mothers in this sample were driven by a desire to see "normality" in their children. They made comparison to other children, and monitored other sources (physicians, cardiac monitors) for signs that their infant was progressing towards "normal" growth or behaviour. Solicited and unsolicited advice came from relatives, other mothers and health professionals. The author suggested that "mothers were open to guidance from health professionals".
2. Brown's findings from mothers when infants were 7-9 months was largely similar[149]. Growth remained low among concerns. However growth was more of a concern for mothers of high risk infants (up to 11 out of 73).
3. Hewat & Ellis interviewed mothers of 9-13 month old infants about their experiences of breastfeeding[151]. The authors concluded that among mothers who breast-fed longer, the rate of weight gain was not associated with nourishment. For mothers who had not breast fed for as long, comments were more likely to reflect concern over poor weight gain and a more rigid and anxious personality. Authors state that comments from health professionals about slow weight gain affected mothers differently.
4. Rand & Wright presented pictures of sitting babies of different body sizes to 1317 individuals (children and adults)[138;139]. All groups found a range of infant body sizes acceptable. Compared to body sizes of children and adults, more sizes of infants were considered acceptable. Middle aged adults were most likely to accept the widest range of body sizes.
5. Birgineau presented 229 adults with photographs of infants (one large and one average weight) to determine whether they associated different characteristics with the fat and average baby[145]. The author concluded that on average, the fatter infant received lower social ratings than did the average infant. There were considerable methodological problems with this study.

12-24 months

Four studies considered growth in infants over the age of 12 months. In three of these the sample included some children older than the upper age limit considered in this review (24 months), but it was not possible to separate data by age.

1. Baughcum, Burklow et al interviewed WIC mothers of infants aged between 12 and 36 months about their beliefs on child feeding[144]. The author concluded that among these mothers growth was a sign of health, and the bigger their babies the better. Heavy children were a sign of competence as a parent, and parents tended to believe that children of heavy parents would themselves be heavy.
2. Baughcum, Powers et al used questionnaires with mothers of 11-24 month old children to obtain views of eating, growth and health[150]. The authors concluded that mothers of overweight children were concerned about overweight and overeating, but across the group as a whole there was greater concern about under-eating and under-weight.
3. McCann et al interviewed mothers of children aged 9 months to 9.6 years who had been born small or diagnosed with poor growth rates about attitudes toward their child's weight and shape[147]. The author concluded that mothers were concerned about poor weight gain, and were glad that they had been referred for help. Most believed that the most beneficial thing about referral was that they were reassured there was no organic cause of poor growth. Mothers blamed themselves for their child's poor growth. Authors argue that for these parents, weight gain was less important than healthy diet, and thus half of mothers restricted the child's intake of sweets and a third "unhealthy" foods. Nearly half of the mothers described their children as picky eaters.
4. Reifsnider et al interviewed mothers of children aged 22.5-51 months (mean age 35 months) about their growth in the previous two years[134]. These children had previously taken part in an intervention for children with poor growth. Authors report the words of participants extensively and many are positive about weight gain. The parents in this sample use clothing sizes, growth measurements and comparisons with other children to gauge growth in their children. There was a perception that there were some things that affected growth – good diets, receiving love and attention but inheritance of size was a common theme. Mothers felt that children inherited their size, or that size was out of their hands – so long as you were taking good care of your child they would simply grow at the rate appropriate for them. Mothers whose first language was not English felt that this was a barrier to getting the information they would have liked from health professionals. The authors noted how closely mothers monitored growth charts and how much emphasis some put on a bio-medical models of causes of poor growth.

We would conclude from these studies that growth appears to be low among priorities for mothers of infants aged 0-3 months. Concern about growth may be particularly low very soon after birth, or for parents of infants with other health concerns (e.g. prematurity). Growth does not seem to be high among concerns of mothers of infants aged 3-6 months, even when infants were born small or early. Study heterogeneity means that we cannot make a summary conclusion about views about growth/size in infants aged 6-12 months. Among older children (more than 12 months) with poor growth there is concern among their parents. Parents want to see good growth in their children, but they also think that love, attention good health and healthy diets are important.

8. 3   Summary of findings of narrative synthesis

The story that emerges from the included studies is of a tension between concern expressed that current size or growth should be 'normal', and the low priority of size and growth relative to other concerns. Although the terms size and growth have different meanings, it was not possible to determine whether participants considered the different implications of assessing current size or rate of growth. None of the studies considered differences of opinion between the meanings of size and growth and it would seem that participants were (unsurprisingly) unclear about the relationship between rate of growth and attained size at any one time. We can say that studies that considered mothers whose children were diagnosed with faltering growth (sometimes called failure to thrive) reported that they were anxious about the progress of their children. As a result of their worries some were pleased when their children received medical attention and acted on advice given.

Monitoring the growth or relative size of children was important for many mothers, not just those whose infants were small. Weighing babies and comparing their size to others appears to be a common practice and considered an important part of health monitoring. In contrast, the reported concerns associated with size and growth *in themselves* are minimal. Mothers of small infants (those with faltering growth and those born small) are more concerned with possible medical complications, such as infections. The focus on observing size or growth is not restricted to mothers. They reported interest and comments of relatives, strangers, health professionals and welfare agencies on the size or growth of their children. Mothers reportedly felt on occasion that this focus on size and growth was inappropriate when children are otherwise well. No data were available to judge health professionals' views of the importance of size or growth relative to other concerns.

Overall there was more concern expressed by mothers in these studies for under than over-weight. As well as receiving advice and help from individuals mothers also looked for advice from books and magazines. None of the studies sought opinions from mothers about the acceptability or advisability of intervening in relation to rate of growth.

In summarising these findings we also note that although there were individual studies of good quality, the conclusions of this review are restricted by the fact that most of the studies were of poor quality *for the purposes of this review* and tended to use designs that are sub-optimal for describing subjective experience. The number of studies, participants and sampling strategies employed means that we would also want to stress that these findings are indicative of participants included, but may not be generalisable.

9.   Discussion of systematic review of lay perspectives on infant size/ growth

The review considered here aimed to address a series of questions. In this review, there were particular difficulties presented by heterogeneity in study aims. Thus two studies that superficially looked substantially similar may often had very different aims. For example the studies by Baughcum, Burklow et al and Hewat are similar in that they are both qualitative studies of feeding decisions of North American mothers[144;151] . Both studies used interviews, in Bauhgcum's case, focus groups. However the hypotheses (and therefore foci) of the studies were different.

The first was that overweight was more common in the offspring of low income mothers, and the second that some mothers did not breastfeed for as long as they wished. The questions asked of mothers and the data reported were therefore very different. By using the new set of questions to interrogate the findings we hoped to find a way to extract what was relevant to our review from the studies. Two methods of synthesis were trailed, which took different approaches to dealing with heterogeneity between studies.

The narrative review was competed by one researcher (PL), using data extracted as described in the methods section. This same researcher conducted the thematic review in collaboration with a second researcher (LA, who also reviewed some of the papers). The two thematic reviewers worked independently on first interpretations of the data, and then in a process of iteration both contributing to the final version. The two reviews ran in tandem, as the thematic review needed time for response and comparison between reviewers. A third researcher discussed the work as it progressed (HR). One of the aims of this part of the review of infant growth was to develop methodologies for answering qualitative questions using systematic reviewing techniques. Having attempted two methodologies it is interesting to compare the strengths and weaknesses of each.

The narrative synthesis grouped the studies into more homogenous groups, each concerning the views of a specific population and could be considered to be a series of related reviews. This technique is familiar from traditional narrative reviews of studies where heterogeneity presents a problem. Using a narrative synthesis enabled us to present a detailed examination of study characteristics including the context for the study. In this way quality and heterogeneity across studies can be considered by the reviewer and readers. This approach allows a large or small number of studies with diverse designs and contexts to be reported. The main drawbacks of this approach are lack of transparency and limitations in the ability to make conclusions. The individual researcher or the research team will determine the sub groups used, the quality assessments and use their judgement to make comparisons between studies. Whilst we might make every endeavour to seek outside views, and to seek agreement between individuals, these methods are not open to external scrutiny. The second limitation is the ability to make conclusions, or create a narrative which will not make the review less accessible to readers. The narrative component of qualitative research is crucial, and is considered by some a quality criterion but becomes more difficult to construct as the data presented becomes more complex and in this case more heterogenous.

The thematic synthesis pooled the findings from all the studies and explored emergent themes. The strengths of the thematic synthesis was in the ability to combine studies with considerable heterogeneity of study designs and aims to draw conclusions from their commonalities. This methodology can be presented in a form that is easier for the reader to assimilate than the narrative synthesis, in a form similar to presentation of primary research. Conclusions from this thematic synthesis fulfilled an important research aim of qualitative research in generating hypotheses. However, pooling findings in this way risks masking the short comings of the studies that make up the review. We have discussed at length the poor quality of many of the studies, and these problems are not represented in the thematic synthesis. This creates a problem in interpreting the findings of the

review. On the one hand the hypotheses that emerge from this synthesis do draw on a broader body of views than any one of the studies (as in a quantitative meta-analysis) and may therefore be more representative. On the other hand we risk making conclusions from a series of studies, none of which individually we may be willing to draw conclusions from. This method may be particularly poor at examining contradictions, as well as commonalities in the data. As with the narrative review, the processes of synthesis may also be open to criticism for lack of transparency.

As predicted by these evaluations of the strengths of each approach, the main difference in the conclusions reached is that only the narrative synthesis is able to address the quality of the studies included. The narrative review contributes significantly to understanding issues that were not addressed by the individual included studies, but which reviewers a priori considered important. These included the views on infant size or growth of those from different ethnic or cultural backgrounds, the effects on views of feeding decisions and the views of those with different weight status. The thematic synthesis was better able to comment on the commonalities between studies, for example the importance for all groups of the importance of 'normal' size. It also enables us to generate hypotheses, for example the differences in views between mothers and health professionals could be interpreted as the discomfort for mothers in applying uniform measures of size to all children, when mothers rely also on social and familial comparisons. The thematic synthesis also suggests an explanation of the finding from the narrative synthesis that parents felt growth monitoring to be an important aspect of health monitoring, but did not view size and growth as defining health. The thematic synthesis suggests that size and growth make visible to parents key aspects of unobservable features of infant care (e.g. love) and health (e.g. undiagnosed illness).

Our purpose in comparing these methods is not to recommend one above the other, but to contribute to a discussion of the potential of different techniques to answer different questions. For the purpose of generating future new research hypotheses the thematic synthesis seems to hold most potential. In contrast, in reviewing the body of literature, and examining the extent of and gaps in what has been studied in the past the narrative synthesis would seems to be more appropriate. Both methods are presented in full in the report to enable readers to reflect on their comparison, and the strengths and weaknesses they may have for other reviews.

9.1 Integrating the systematic review findings into focus groups

In addition to completing a systematic review of lay perspectives on infant size and growth, we carried out some exploratory primary work in East London the purpose of which was to interpret and build on the findings from the review. How salient the research findings of largely North Ameican studies might be in a different context, and what might the implications be for intervention or future work? In our original application, we had intended to test out potential interventions identified by the review of outcomes with people who might be the focus of such interventions. The revised application, with a shorter timescale, led us to a more modest aim for direct work with people for whom services are provided. Discussion with them would, we felt, be one way of refining, addressing and responding to issues arising from the review. As stated by Kitzinger *"group discussion is particularly appropriate when the interviewer has a series of open ended questions and wishes to encourage research participants to explore the issues of importance to them, in their own vocabulary, generating their own questions and pursuing their own priorities"* [163]. This loose discussion structure was well suited to our aim of allowing participants to describe their experiences of infant growth and size, outside of the context of a strong research question. In addition group work encouraged participants to talk about their own interpretations of shared experiences and to compare their beliefs.

At the outset of the project the participants and recruitment methods for the focus groups were considered by the research team and advisory group (see section 10). The recruitment and running of the groups was delayed in the hope that initial findings from the review could feed into these discussions. However, timescales demanded that the groups had to take place before the review was complete. At the point when the groups took place, findings from the outcomes studies reviewed to date were not suggestive of clear evidence of benefits associated with particular patterns of growth. The review of lay perspectives was also well underway, but with no clear conclusions. The analysis of the focus groups was deliberately delayed until after the findings from the review of lay perspectives were complete, and these were used as the basis for analysing the data.

10.   Focus group methods

We conducted 34 brief screening interviews, and 5 focus groups described in more detail below. These explored participants' views of infant size and growth and the possibility of interventions to change growth rate. The work was carried out in March and April 2004.

The transcribed interviews were analysed in terms of the thematic framework created for the synthesis of review findings. Agreements, disagreements, and additional issues were noted. Two researchers (LA and PL) independently analysed all the data and the final interpretations were reached by consensus. The categories under which themes were examined are:

- Constructs regarding infant growth and infant size, including assessment of growth
- Concerns about growth, particularly relative importance among health concerns
- Influences on views and behaviours

In addition participants' responses to the possibility of intervention were analysed.

10.1 Sampling and recruitment

Our recruitment strategy was underpinned by the literature on social networks, and in particular, the role of key individuals in informing decisions made by mothers,[150;164] the dynamic relationship between mothers and those in their support network, the background to advice given, and factors which influence on decision making processes. In the light of this, it was decided that speaking to primary carers together with those in their support network could be informative. Work similar to this has recently been carried out to examine cultural change and eating habits in immigrant communities in Australia[165]. In the Australian study, researchers spoke to family groups, including grandparents, parents and children in a setting where participants could exchange views on what constituted healthy food for children. The use of only one individual's support network meant that the group size was determined by the choices made by the mothers participating. In the work we did, no mother asked for more than two people to be present, so our group size was two or three. Kitzinger (1996:41) suggests that the 'ideal' size for focus groups is 4-8, however we considered that methodologically the process retained the characteristics of focus groups. We therefore we have retained the phrase here, despite the fact that some groups contained only two members

A two-stage recruitment process was developed; an initial screening interview of convenience samples, followed by focus groups selected on the basis of data collected at the screening interviews. Ethical approval for this study was obtained from the City and East London Local Research Ethics Committee. All participants were asked for consent to take part in the study, for our conversations to be recorded, and for these recordings to be used in research. Participants who took part in the long interviews were given vouchers worth £10 to thank them for their time.   Invitations to participate and consent forms are attached as appendix 15.
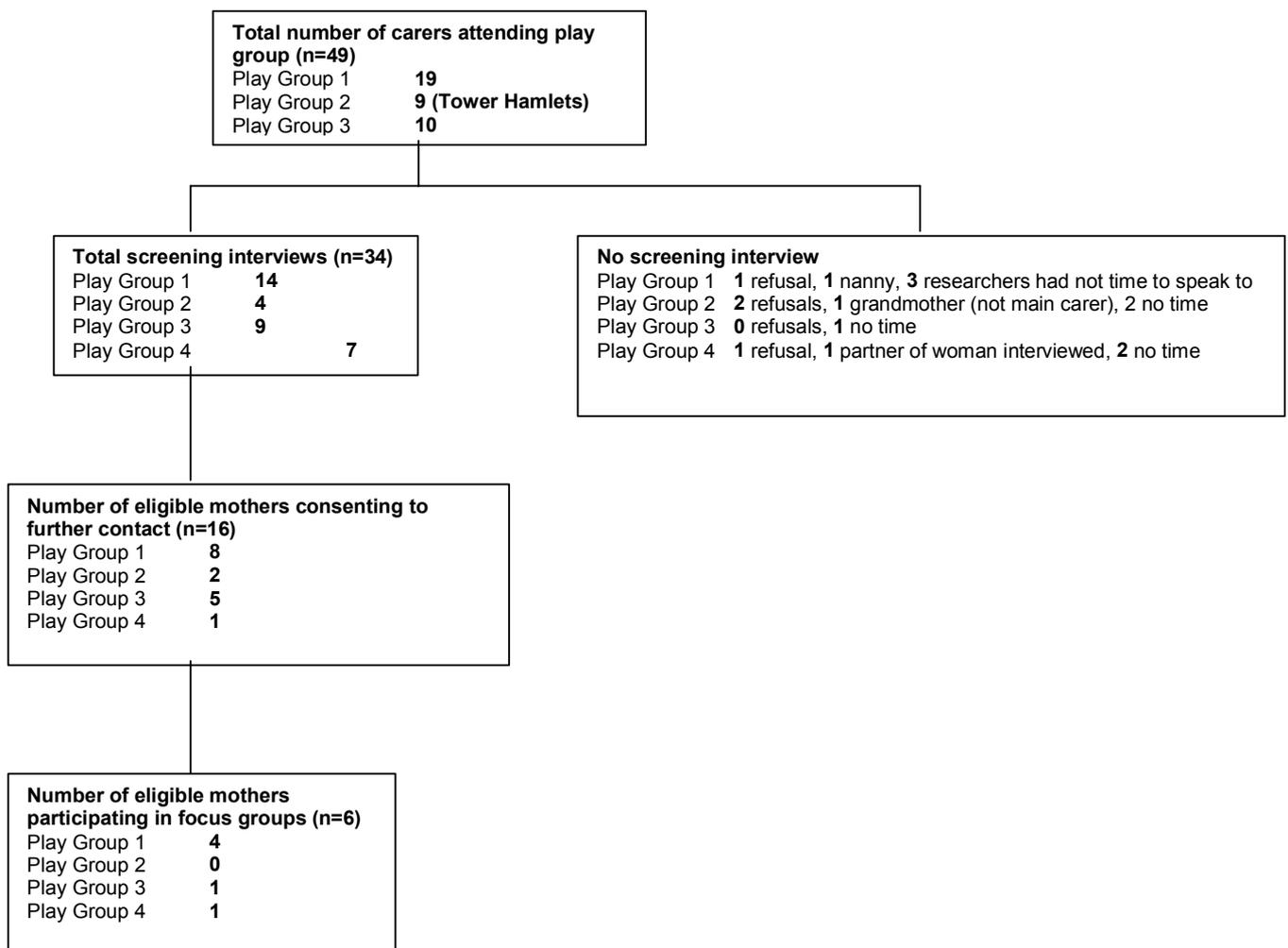
Two local authorities in East London were contacted and asked for comprehensive lists of organisations which provided drop in services for parents with young children. These were screened to identify all those groups where parents stayed for the session, and where under two year olds were included. Services run in health centres were excluded, since we wanted this work to be seen as independent of health services. Organisations were contacted to ask if they would be willing to allow researchers to visit these groups. We prioritised the larger ones. The local knowledge of researchers was important in identifying which these might be, but the number of sessions per week was also considered a good indicator of size.

In total 4 different groups were visited over a two week period in March 2004, 3 were in Hackney and the fourth in Tower Hamlets. Two other groups in Tower Hamlets were approached. One agreed to take part but cancelled when no mothers attended on the day specified. The final group contacted required a female only research team, which was not possible to arrange in the time available.

Two researchers attended and attempted to speak to all carers present at one play session in each location. In three cases the research team comprised one male and one female, and in one case (Group 3) two female researchers. A poster was

put up at least 3 days before visiting the groups (appendix 14). On the day of the visit all adults were approached, told about the study, and given an information sheet (appendix 15). Carers were then given some time to read the information sheet, and researchers then returned to ask whether they would be willing to participate. Those agreeing completed consent forms and then took part in a screening interview during which they were asked a series of brief questions about their experience and opinions of the growth of their youngest child (see appendix 16). These questions were designed to identify which carers were eligible for this study (primary carers of at least one child under 2 years). The questions also provided some data on views.   The numbers of carers spoken to are described in the figure below, as is the number of eligible mothers who consented to contact about further involvement and those who took part in the focus groups. All interviews were conducted with women. Only one man attended and his partner had already been interviewed.

Figure 5: Participation in study; numbers of participants in screening interviews

Total number of carers attending play group (n=49)	
Play Group 1	**19**
Play Group 2	**9 (Tower Hamlets)**
Play Group 3	**10**

Total screening interviews (n=34)	
Play Group 1	**14**
Play Group 2	**4**
Play Group 3	**9**
Play Group 4	**7**

No screening interview
Play Group 1   **1** refusal, **1** nanny, **3** researchers had not time to speak to
Play Group 2   **2** refusals, **1** grandmother (not main carer), 2 no time
Play Group 3   **0** refusals, **1** no time
Play Group 4   **1** refusal, **1** partner of woman interviewed, **2** no time

Number of eligible mothers consenting to further contact (n=16)	
Play Group 1	**8**
Play Group 2	**2**
Play Group 3	**5**
Play Group 4	**1**

Number of eligible mothers participating in focus groups (n=6)	
Play Group 1	**4**
Play Group 2	**0**
Play Group 3	**1**
Play Group 4	**1**

Ethnicity of participants

All those who agreed to take part in the screening interview were asked to self-identify their own ethnicity from a list provided. The numbers of participants identifying themselves as each ethnicity are shown in the table below.

Table 31: Ethnicity of participants

Asian	Indian 1	Pakistani 0	Bangladeshi 0				
Black	Caribbean 1	African 1	Somali 0	British 1			
White	UK 18	Irish 1	Greek 0	Turkish 2	Orthodox Jewish 1	Kurdish 0	Other* 6
Mixed	White & black Caribbean 2	White & black African 0	White & Asian 0				

The majority of those who completed the screening interview were white UK. The ethnic mix is not representative of the two boroughs, and the sampling framework meant that families where both parents work full time and socially isolated carers were unlikely to be reached.

From the 16 screening interviews where participants consented to further contact, families were selected to take part in focus groups.

10.2   Which families were followed up

All mothers of children under the age of two were asked whether they would consider taking part in a further interview. The focus groups was briefly described, and mothers were assured that they were not committing themselves to take part, but simply giving permission for a researcher to contact them by phone to discuss it. Sixteen of the thirty-four interviewed agreed to this. These were then ranked according to criteria we developed for their inclusion in focus groups. Mothers of babies who had had notable growth experiences (e.g. large for age, small at birth) and those with a strong local support network were considered most suitable for the approach we were taking. While those with weak social support networks would be interesting for other reasons, our interest at this stage in exploring review findings in a social context prioritised the former group. Researchers contacted these women in order and asked if they would take part in group discussions. Once four groups had been completed, researchers contacted the remaining mothers to thank them and let them know they would not be asked for further contact. At the contact stage, two mothers decided they would rather not take part, and a further 4 could not be contacted. One of the groups interviewed consisted of 3 of the mothers, who chose each other as their support network. Four mothers were not contacted. Consent for participation was obtained again at the beginning of these interviews.

---

* Scottish, Spanish, Non-orthodox Jewish, Slovakian, Australian, Iraqi

Those taking part are not representative of the areas in which they live in terms of ethnicity or social class, probably for the reasons described above in relation to the screening interviews. While this would be a fatal flaw if we had been focussing only on an intervention, we do not feel that it undermines the more modest and exploratory exercise of testing out the review findings with parents.

Four mothers from group 1 took part in the focus groups, 3 were interviewed together at their regularly weekly get together. The other was interviewed with her mother, with whom she lived. One mother from group 4 was interviewed with her sister-in-law. These mothers all lived in Hackney. One mother from Tower Hamlets was interviewed, along with her partner and her mother.

10.3 Focus group content

Focus group content was derived from the initial project aims, supplemented by questions arising from the outcomes review and the systematic review of lay views of early growth. Discussion content was piloted in an interview with the parents of 2-year-old twins, who lived in the study area and were known to one of the researchers. The mother of the twins identified her partner as her main source of advice and support, along with her mother, and the interview was set up to include these three. The twins' grandmother was not able to join the interview as planned, but the interview went ahead with the parents. The parents in the pilot interview were asked to reflect on the content and running of the interview. They were specifically asked whether they felt any questions were unclear or inappropriate, and whether they felt they had had an opportunity to talk about what they felt was important. Both said that they felt all the questions were appropriate, and that they had enjoyed spending some time thinking about how the last two years had been for them. The main concern of the mother was that they had not spent enough time answering our questions.

At the choice of the parents, all interviews took place in their homes. Vouchers worth £10 were given to each adult participant to thank them for their contribution to the project.

The discussion guide is included in full in appendix 16. This was used to stimulate discussion and interviewees were encouraged to discuss issues they felt to be relevant to the topic. The topics included were opinions about growth, their experiences with their children, sources of advice, and the possibility of intervention. Each interview concluded with the lead interviewer summarising what she thought she had heard, which sometimes initiated further discussion, and a final invitation to interviewees to let us know whether there were other issues important to them that they would have liked to cover.

11 Findings from focus groups

11.1 The participants

The names of all participants and their children have been changed to protect anonymity. Table 32 indicates who took part in each interview so that quotations

used are identifiable. The research team agreed that the pilot interview be included in the analysis. We decided that given the exploratory nature of this work, we would carry out a pilot, and were aware that this might result in significant changes to the topic guide. As we describe below, the topic guide did, in fact, appear fit for purpose in the pilot, particularly given the iterative nature of the discussion format, where each group built on what had come before. The pilot group was recruited using a sampling strategy similar to that of the other interviews and we judged it appropriate to include the results from the pilot. All participants described their ethnic background as white, UK, and Martha was Jewish.

Table 32: Composition of focus groups

Pilot Interview, Hackney	
Sarah	Mother of twins aged 22 months. Working part-time in newspaper publishing.
David	Father of twins, partner of Sarah, working as a psychiatric nurse.
PL (Lead interviewer)	Lead interviewer, and lead researcher on this project. Mother of two children aged 8 and 10 years.
GK (Interviewer)	Male research intern in research unit, no children.
Interview 1, Hackney	
Mary	Mother of three boys aged 2, 5 & 7. Formerly working in book publishing, not working presently.
Ruth	Mother of three boys aged 14 months, 5 & 7. Formerly working as a social worker, currently training as psychotherapist.
Laura	Mother of one boy and two girls aged 28 months, 4 & 7. Working part-time as a picture editor in newspaper publishing.
PL (Lead interviewer)	Lead interviewer as before.
Interview 2, Hackney	
Gemma	Mother of one boy and one girl aged 15 months & 3. Not currently in paid employment.
Louise	Aunt of children, married to Gemma's brother. Lives in Wales; visiting the family. No children. Not asked about her employment.
PL (Lead interviewer)	Lead interviewer as before.
HR (Interviewer)	Head of research unit, Professor of Child health. Mother of 4 adult children, and one granddaughter aged 16 months .
Interview 3, Tower Hamlets	
Mike	Father of two boys and one girl aged 21 months, 5 & 9. Not asked for details of job, but is working.
Anne	Mother of Zoe, grandmother of three. Works part-time in the school the children attend as a meal time supervisor.
Zoe	Mother of two boys and one girl aged 21 months, 5 & 9. Not currently in paid employment.

JM (Lead interviewer)	Female researcher in unit. From new Zealand. No children.
PL (Interviewer)	As before
	Interview 4, Hackney
Sue	Grandmother of one boy and one girl. Not currently in paid employment.
Molly	Mother of one boy and one girl aged 9 months and 27 months. Currently on maternity leave from post as teacher in FE college.
PL (Lead interviewer)	As before
MS (Interviewer)	Researcher in unit. Mother of two boys aged 10 and 4 years, and a girl aged 8 months.

11.2   Understandings of infant size/growth

"Sue: I mean obviously there are normal variations.
PL: Yes, babies, toddlers, people of all ages are naturally
Sue: are different
PL: across a big range?
Sue: Yes, exactly. Different sized bones, different sized everything.
Molly: Everyone is an individual and needs to be healthy within themselves and be the best that they can I suppose."

Findings from the focus groups largely fitted within the framework constructed using the systematic review of existing studies. All participants referred to the importance of "normality" in infant size. Variations in size were acceptable, and where size was seen as unacceptable, the causes of this "unhealthy" size were usually discussed. Table 33 summarises findings. Themes identified from the review of literature are shown in bold and additional themes drawn from interview alone are in italics.

Table 33: Understandings of infant size/growth

Observable Norm Creation
Familial comparative
"Zoe: . . . You'll probably see him when he comes in and think there's nothing tiny about him! But to us he is, compared to the other two and because of his eating ways"
Social comparative
"Laura: well you draw a lot from their peer group don't you? It's kind of . .
Ruth: Yeah, school, I mean nursery whatever.
Mary: what fits into the middle of the peer group was what you'd kind of see as being the average and that's how you measure how extraordinarily large or little you've got."
Medical comparative
"David: You get your red book and they fill in the dots and weigh them"
Developmental comparative
"Gemma: he shouldn't be in a buggy at that age but I'm like well he's only two."
Material comparative

"Molly: . . both me and my mum felt she was getting too fat. She kept having to buy new clothes you see" Visual judgement/looking "normal" *"Gemma: . . But then I can see she is growing. Putting on weight"* *Own experience* *"Laura; . . . .It's funny because with your third you don't tend to go along quite so often"* *Comparing across sources of information* *"Ruth: yeah. I mean he actually probably on percentiles he was probably born on the, I mean he's average but I mean actually he's just grown, . . .but I just know by him and in comparison he looks, he's very small, he's fine boned and he's wiry"*
Observable Explanation of difference
Relation between eating and growth **"Louise: There's normally big, then there is big-big isn't when they're like ob[ese] , from eating unhealthy food, you can tell the difference from when a child is big in build as in big as when it is from eating junk food. . . it's not because his dad's big or whatever, it's because he has been eating unhealthy stuff it's not been healthy living."** **Ruth: ". . . sometimes I think that's where bottle and that's an unnatural thing and you just feed the baby and you give them the bottle and . . . that's why they end up being fatter."** *Physical make-up* *"Sue: Different sized bones, different sized everything."*
Non-Observable Norm Creation
Inherited/familial characteristics *"Ruth:. . . uhuh, I think it's genetic you see. . . the men in my side of the family are unfortunately rather short [laughs] and um, the other two I think have inherited my husband's who are all rather tall"*
Non-Observable Explanation of difference
Nature of child "Zoë: . . I think a lot of it goes together. He ate well, he slept well and he was a good contented baby whereas Davey was a moaner. He didn't eat, didn't sleep well." **Quality of care** **"Mike: You think are they getting enough, rather than are we giving them too much you question are we giving them enough food? You look at yourself and how you think you are bringing them up.** **Anne: It is a reflection on you."** **Medical cause of growth rates** **"Zoë: I think it must be an imbalance because she really is big."** **Fatalistic** *"Ruth: I think that as long as things are alright generally at home as well, then they'll just follow their natural, instinctively"* *Individualistic* *Gemma: "Every baby is different, every child grows differently, at a different rate. So. As long as there is nothing wrong they'll be fine."* **Parental preference** **Sarah: . . . "I do find nice roley poley is quite attractive."** **"Laura: it was quite nice having a tiny baby because Derrick looked like he was about two months old when he was born and I have to be honest I quite enjoyed having this tiny creature come to live with us"**

11.3 Concerns regarding infant size/growth

Sue: "I suppose you have to have a standard so there is not a dangerous level of people getting too fat or too thin but everybody is different, aren't they? I think you are different right from the moment you are born, everybody is individual so you can't make everybody exactly the same. Obviously it helps to have certain sort of rules or, you know . .
Molly: Because if they are very underweight and not feeding properly then you have got to address it but if they are under average but they're eating well…
Sue: And they are really lively.
Molly: And they are growing up well then you have to let them get on with it in a way."

As for participants in the systematic review, growth in itself was not a cause for concern for our participants. Growth was something that parents and other family members worried about because it might mean there was an underlying problem. Table 34 summarises findings. Themes identified from the review of literature are shown in bold and additional themes drawn from interview alone are in italics.

Table 34: Concerns regarding size/growth

Observable Relating to infant
Infant behaviour
Molly: ". . she is constantly being told that her children are underweight but . . the children are very lively and they're full of energy there can't be that much wrong with them."
Social meaning
Laura: "I don't know, I still think at the moment people think it's sweet to be a small girl and whereas it's somehow kind of a weakling to be a small boy perhaps.
Ruth: yeah, I mean that smacks to people of being feminine"
Developmental context (incl prematurity)
Laura: "I think that if you're baby gets really fat it affects their mobility. . . heavier weight children find it difficult to kind of get themselves off the ground and often you find the earliest walkers are the tiniest children"
"Molly: Babies change so much physically I suppose and when they get to five or six they can develop habits and they can continue into adulthood. . To me, I see a baby as a different thing to an adult"
Medical context
"Zoë: ". . you don't know what's worse, having one that's big or having one that's small. When they're big they worry about the diabetes and other things and when they're small they worry about getting infections."
Consistency of growth
Ruth: "I mean we use that a lot in child protection if there was a sudden kind of , if a child was growing on a particular centile all be it in the 9th and suddenly I mean that
Mary: it means something
Ruth: that means something . . .
Laura: major in that child's life"
Observable Relating to parent
None

Non-observable Relating to infant
Quality of care **Nutrition** **PL: "What does a toddler that's too thin look like?** **Ruth: undernourished"**

Non-observable Relating to parent
Feelings of confidence & competence *Mary: "When they're little, you're so vulnerable with your first child so kind of confused as to what, you've got this thing with no instruction booklet that when they do come round you can be quite kind of nervous and the house to look kind of tidy and* *Laura: look like you're coping* *Mary: look like you're coping and you're dressed, which is rare. And I remember being told off I think for having a hat on the babies head in the house. And I was winding one of them and he sort of projectile vomited all over the place and she said, maybe you shouldn't be patting him quite so hard. And it was true in my nerves I was going like this* *[sound of loud patting and laughter]* *Ruth: I felt, I felt quite like that, quite looked at."* I used to take him to get him weighed and they said he was a bit underweight but he was well, I knew he was well so I didn't worry about it. ." **Perceptions of child needs** *Zoë: ". . when you are breast feeding, you don't know what they're getting."*

11.4 Influences on views and behaviour

Parents were asked about where they went for information or advice and who commented on the size or growth of their babies'. Participants spoke most about the influences that were closest to them, their friends, families and those they were in regular contact with. Participants in these groups expressed more negative views about the role of health professionals and the place of the 'baby clinic' than was reported in the views literature. This could be explained by the predominance of clinic attendees in the study samples in the review, but may also reflect local concerns in a part of London where health resources are particularly stretched. Table 35 summarises findings. Themes identified from the review of literature are shown in bold and additional themes drawn from interview alone are in italics.

Table 35: Influences on views and behaviour

Influences from individuals and those within day-to-day life (microsystems)
Families **PL: "Of all the advice you got, which has been the most useful?** **Gemma: I think my mum's or my sisters. Yeah, cos they've probably been through it. And I can trust what they say."** **Other parents** **Mike: "We have three children and I work with blokes younger than me who have children and we work in a big office with women there but we still talk about our children, the men, and we talk about football and all that but we do get on to talk about our children more than the women do sometimes. But Zoë's mate might ring her and ask for advice but in**

work we more or less tell each other things rather than giving advice. . . I'll tell them what I did, but I wouldn't give advice."

Sarah: ". . . The most useful tips and information that I have ever received, are through just parents that you meet on the street, you don't know, you just talk to them for 10 minutes, you come away with a great "hey that's how I can do it then, that's worked for them". And I think that's the most powerful…"

<u>Health Professionals; Health visitors</u>

David: "That's what I mean, I think, what I sense from, I'm sure there's great ideals behind, but it's always come across as pressure rather than support, you know. "You've got to come up and get them weighed", "you've got to do this", "we want you to do this", "we want you to do that". It's never been, "how can we help you do that" or, you know. . . . So monitoring for signs of neglect is not exactly what I want as a parent."

PL: "Are there other people that you would ask advice from?

Gemma: Health visitor, um, friends. There's a health visitor that comes to Ivy Street she's got a little girl so that's handy. I always pounce on her!"

Health Professionals; GPs

Ruth: "the GP was actually really helpful . . . But I had quite a well established relationship with him."

<u>Strangers</u>

Zoë: "Going back to when I had Sam everyone was the same, and it upset me a bit didn't it Mum, 'Isn't he big'?"

Trusted individuals

Zoë: "Not a lot of them are like the old school, do you know what I mean, older people? You have some of the older types of midwives and you feel safe with them, do you know what I mean? I mean a student delivered Davey and it was her first baby that she delivered and she was brilliant, but sometimes you feel more secure with the older ones.

JM: You want experience and knowledge.

Zoë: Not just someone who has just read it in a book.

Mike: When that student delivered Davey she was meant to finish but she stayed on, which was nice.

Zoë: Yeah, I trusted her but sometimes you want more experience."

Laura: ". ., it wasn't the GP it was a locum a really young guy who gave me really pathetic patronising advice . . . But the advice was next to useless and a part of that might have been because that guy didn't have children of his own and had never been through this

Ruth: yeah. I do think experience has a lot to add"

Influences at organisation and community level (macrosystems)

Extended Family

PL: "Who is it who talks about the size of your baby?

Ruth: relatives

Laura: Yeah if you say relatives it's kind of looking at maybe previous generations it's the kind of thing that grandparents might say, or aunts and uncles, great aunts and uncles kind of things. So yeah must be something that comes from previous generations."

Government

Sarah: "Well there's your bog standard NHS leaflets and stuff, which we've read."
Commercial publications
Mary: "and it was the fact, and it was something about cross referencing books and reading different books was better than health visitors or my mum whose ideas were 20 years out of date."
Only one mother had looked at books/magazines and not found anything of use in them, Ruth found them "irritating".
Historical and Cultural Context (chronosystems)
<u>*Cultures*</u>:
it also reflects my being in California. I mean you always say that Mary, also there's culture that goes into being, you know, culture, class. ."
Media attention:
Zoë: ". . Plus on the news on the telly a lot it says don't give them this, the things you read, if I am looking through the paper and it says these are no good, I think I won't buy them no more."
PL: " when we sat down Francis was talking about, you know, well you can hardly avoid the concern there is for obesity in childhood at the moment
Laura: especially recently
Ruth: every day
Laura: you can't avoid it"
<u>*Historical changes*</u>
Ruth: "Also, I'd never dream of, you know my grandmother used to leave my dad in a cot with bottles tied to the, tied to it when she went out and she's a doctor . . .
Mary: it's only now that I realise that all these things are subject to fashion
Laura: absolutely
Ruth: like Dr Spock"
Sue: " I think people used to think being chubby was more healthy than being not so chubby but nowadays you don't think that, everybody is so conscious of not getting fat and obesity and all that sort of stuff so you would be more wary of your baby being overweight rather than."

11.5 Opinions about possible interventions

On the whole parents were willing to acknowledge that intervention might be appropriate sometimes, but thought it inadvisable for most cases. The reasoning that participants went through when considering this question brought together many of the themes that have been discussed in terms of constructs, concerns and influences on views regarding growth. Individuals within each interview often disagreed with each other. For example, in the following discussion, Molly and her mother were considering a friend of Molly's who had a baby they both thought was too fat. This text has been annotated in capitals in terms of the themes researchers felt they were referring to:

Molly: "Well I think the health visitor should have been the one to notice that with her, that she was fatter, it would be her place to say that[HP].
PL: but do you think that's acceptable advice

[HP] Influence of Health Professionals

Molly: I think it is acceptable. I think you have got to be able to tell people the other way as well. I think people are so interested in feeding children up all the time but someone should take notice if a child is eating too much as it is equally damaging to have too much than it is to not have enough^{NG}

MS: but people don't think like that

Molly: they don't think that that's a worry so much, as long as they want it they should eat it.   But if you give them the wrong, even if they eat loads and loads of vegetables it might give them an upset stomach, they have got a mouth thing and they like to put things in their mouth and they don't know they are actually eating. I think you should be told. With breastfeeding and everything it is all on demand, the baby knows exactly how much they need. Well maybe they don't always. Maybe they are just greedy, or they don't. I mean Ewan liked to keep sucking on the breast forever and ever, he wouldn't let go if he had the choice^{NG;IB ; ED}

Sue: But you don't know that that was her way of developing, that she needed to build up and then let it go later.^{IN}

Molly: Well that's what she did do in the end.^{IN}

Sue: Yes, so you don't know that that might have been right for her because she is okay now.

Molly: Yes, she is okay but the thing that worried me with her was the type of food and her kidneys.^{NG; MC}

Sue: It's rather the type than rather the quantity really.

PL: I get the feeling that Chloe feels sure about cutting back but you are not so sure, you wouldn't feel comfortable restricting food like that?

Sue: Well I suppose if it made the child unhappy, if it seemed as though it really needed and wanted more food.^{IB}

Molly: Well then I could see that if food was the only thing it wanted and would make it happy but then you would have to try and give it

Sue: but then you would make sure that the things that you gave it were more

Molly: it wasn't apple crumble and custard. Yes I think it really does boil down to give them another piece of fruit or a biscuit^{NG}

Sue: I would say it's not quantity it's what they having really rather than . . .

PL: So advising to cut back quantity even if that made the child uncomfortable or unhappy or cry a bit more?

Sue: No, I wouldn't. Not unless they were absolutely obese I suppose. No, no I suppose if you had a really obese child that was really you know was obviously^{CN}

Molly: But I think that's more into children I think its very difficult to say what an obese baby is, I don't think you can make that judgement because they are so different but it is very easy to judge what an obese adult or an obese child is but this is where the problem lies as babies are so different and some of them that are enormous and fat and chubby turn into the skinniest beanpole adult. They are very different babies"^{CN; F}

---

^{NG} Nutrition and growth
^B Infant Behaviour
^{ED} Explanation of difference
^{IN} Individualism
^{MC} Medical context
^{NG} Nutrition and growth
^{CN} Creation of norms
^F Fatalism

As reported above, most participants felt that most of the time intervention was inappropriate. They said things like:

Ruth: "I think that I would find it very odd."

Mary: "I would be totally sceptical
Ruth: likewise
Mary: I'd think what a load of old rubbish. Because just surely there's more to it than just the rate you put the ounces on."

Gemma: ". . My nephew, they're told my nephew to go on a diet. He's only three. It is ridiculous."

Suggesting possible longer-term harm was likely to decrease parents' feelings of competence, and cause them anxiety rather than encourage them to accept intervention:

"Ruth: "no, you're saying hypothetically
PL: I'm saying hypothetically
Laura: Oh!! I thought you were telling me that was the case!
Mary: Getting worried! Oh no! Whoever would have thought such a thing!"

Ruth: "I just find that hard to, I mean maybe I feel slightly defensive about it."

David: ". . Because the last thing that you want is an anxious parent. That's got, surely that's going to have more of a detrimental effect on children than catching up their weight."

Sue: "Well, if it was absolutely 100% certain that it was going to affect them in that way in later life then yes, I think you would try and restrict things a bit.
MS: Would you believe it was 100% certain just because somebody told you?
Sue: Just because somebody told me, no not necessarily, no, no!"

There were times when parents did feel intervening to change growth was appropriate, but these tended to be for older children. For example:

Zoë: "In Davey's class is a little girl who is really big, she is five and I reckon she is seven stone easy. You notice when you go on school trips, we went to Greenwich and walking up the hill she was bright red, out of breath, moaning because her legs were rubbing.
Anne:: Oh bless her.
"Zoë: And I wanted to cry because I would think if that was my child I would have to cut out everything, I would, I really would . . "

Gemma: "She sorted him out. She's contacted the school, he's not allowed to have tuck shop. Because he's, she'll send him with fruit. He's not allowed tuck shop. . .. . he's grown tall and he's got lots of friends now, he's very popular now and he's very happy, so, I think she did the right thing."

Almost all participants felt that the most appropriate intervention would be to change diets for over weight children, reflecting a belief that most over weight is caused by inappropriate eating. This also emphasises their feeling that only weight outside the bounds of what they could interpret as normal should be attended to. What is important is creating and maintaining a healthy diet and healthy growth patterns, not just size:

JM: "Do you think it is okay to restrict food when they are babies and infants?
Anne: If they are really podgy, yes. I wouldn't restrict it so much as what Zoë says, limit it to like one a day and give them more fruit"

PL:" You [could give] higher calorie foods, for example only egg and milk custards that have a lot of calories in rather than vegetables.
Molly: I think that is all right in the short term but you wouldn't want to do that in the longer term as it would develop tastes for sweet rich food which might not be so healthy.
PL: So again it's about the healthiness of the diet?
Molly: Its overall healthiness"

PL: "so that rather than , if the options were to supplement feed her and build her as up as quickly as possible til she gets to average weight or there abouts. Or the alternative that, you know, take it really slowly we'll monitor the growth we prefer slow growth so we won't do additional feeding we won't do, um,
Mary: I think that would be acceptable. It kind of makes sense in a way that they're just on a track
Laura: yeah, just to follow your centile, as long as you are kind of progressing along a path."

Parents were also presented with hypothetical scenarios relating infant growth to outcomes in either childhood (such as development) or adulthood (such as IHD). For example, we asked what participants would think if it were to be found that the way that babies grew between 2 months and 6 months might predict their readiness for school when they were coming up to 5 years of age. In both cases participants agreed that if such a link were proven they might act upon it, but they considered the existence of a simplistic relationship implausible. Their view was that they would need to be convinced that the evidence was overwhelming and incontrovertible. They did not comment on plausibility in relation to outcomes in the shorter or longer term.

12. Discussion of findings from focus groups

The findings from these interviews are drawn from a small, purposive sample. Most of the mothers in this sample did not work outside the home, and those that did worked part time. All fathers were in full time employment. All participants were from a white UK background. None of the children of participants had experienced serious problems with growth, and all the participants had more than one child. In line with principles of qualitative research, the purpose of these interviews was to generate new material. The views presented here make no claim to being representative. On the basis of what is reported, we would *hypothesise* that intervention appears unlikely to be acceptable except in cases of extreme under- or

over-weight. Even this may be difficult to recommend since parents are resistant to interpretations of "abnormal" or "unhealthy" growth in infants. We would also suggest that in order for intervention to be acceptable there would need to be clear and unequivocal evidence of benefit that was accepted as genuine by parents and families as well as health professionals. It would seem that current media attention has made parents concerned about the possibility of over-weight in children, but that at present they are unsure how to act in response to this concern.

These views also support the findings of the systematic review, in that they provide evidence that the thematic synthesis of review findings presents views and opinions that appear salient in a current UK context.

13. Discussion

13.1 Introduction to discussion

This section first sets out the original aims and objectives of the study. It then summarises the findings from the different components of the study. Next it discusses the learning from this study, in terms of the methods and methodological development, and the process of undertaking the study. The findings of the study are then discussed in relation to the original aims and objectives. Finally it draws conclusions on what the study has added and discusses the implications for families, policy, practice and research. Wherever possible, this discussion is integrated so that all components of the study are considered together.

Our overall aims were to

- Systematically assess and collate scientific evidence and lay perspectives on infant size or growth with a view to determining whether optimal patterns of growth can be defined

- To identify gaps in the evidence, assess the implications of those gaps, and make recommendations for addressing the gaps.

We also identified a number of themes which would cut across all layers of the review. These were: the possibility of beneficial and adverse outcomes in relation to early growth; the balance between early and late effects of infant growth on health; consideration of inequalities in people's lives; and capacity to inform public policy.

We refer back to these aims, objectives and themes during the following discussion.

13.2 Summary of main findings in the different parts of the study

13.2.1 Review of life course outcomes

The association of infant growth status with health and wellbeing across the life course was studied in depth. Over 160,000 thousand abstracts were screened and 279 papers assessed in full to select the 119 papers (relating to 108 studies) which contributed to the final review. Despite this large effort, relatively little or no information was found for many outcomes which, a priori, we had considered important. We identified many gaps in the research literature linking infant growth status to health and well-being across the life course. These included lack of information on many diseases which are major causes of child and adult mortality and morbidity. Similarly in the case of health-related behaviour, no studies relating to smoking, physical inactivity, drug or alcohol misuse or unsafe sex were identified. The exception here was obesity, which can be considered to be a proxy for patterns of eating and physical activity, where a number of papers on childhood and adult obesity were found. Thus, because of these gaps in the literature, we have no capacity to answer questions on the associations of infant growth status with many beneficial outcomes (for example, employment or healthy lifestyles). Only one unpublished study was identified relating infant growth status to non-health related quality of life. We have no reason to suppose that infant growth status would be causally linked to these outcomes, but it may be an indicator, so the lack of

research evidence in this field is disappointing. In addition, published research rarely allowed estimation of the size of effect of the relation of infant growth status to outcome, nor a description of optimal infant growth status in quantitative terms.

We were, however, able to consider the relation of infant growth status to many outcomes describing health in early and later life, identifying 107 studies on this topic. Of these, 52 studies considered infant growth status in relation to various measures of childhood development, which may be construed as a measure of wellbeing.

For adult health outcomes, we found reasonable numbers of studies (5 or more) for ischaemic heart disease and adult obesity. This was also true for insulin dependent diabetes which results in burden of disease in both adolescence and adulthood (although in this review the majority of studies were based on adolescent subjects). Other important adult outcomes, as assessed by burden of disease and disability, such as cancer, mental health, stroke, dementia and non-insulin dependent diabetes had few or no studies associating them with infant growth status. Of the studies available there was more information on infant size than growth. In general, there was evidence that smaller size (mostly determined by weight) in infancy was associated with increased risk of ischaemic heart disease. The direction of association was the same with risk of suicide and stroke, and more equivocally with non-insulin dependent diabetes (all based on three or fewer studies). Both adult obesity and insulin dependent diabetes were associated with larger size (weight and weight-for-height) in infancy and higher rates of infant growth, though there were fewer studies on growth than size. The majority of the studies of adult obesity focused on the relation of infant obesity (usually defined in terms of weight or weight-for-height) to adult obesity, and therefore were often considering only one part of the distribution of infant growth.

The context and setting of these studies is important in drawing conclusions from these results. For the studies of insulin dependent diabetes and of adult obesity, there were a range of research settings and researchers. However, for other adult outcomes, the research literature in this area is dominated by 3 cohort studies. These are the British 1946 cohort study, and the Hertfordshire (UK) and Helsinki (Finland) cohorts. This means that the findings may (at least partly) reflect specific features of the cohort's members, the time and place of the cohorts, or the research groups studying them. We return to these discussion points later.

Few studies examined infant growth status in relation to children's health. We identified a reasonable number of studies of sudden infant death syndrome (SIDS), which suggested that infants who died of SIDS grew less well as infants, though this was not a marked effect. Nine studies examined infant size or growth in relation to childhood obesity. Childhood obesity was more likely in those who had either been obese (variously defined) or who had gained weight rapidly in infancy.

There were 5 eligible studies relating to mortality or morbidity in childhood. Three of these related to respiratory or diarrhoeal infection and demonstrated that better growth in infancy was associated with lower levels of morbidity and mortality. A single study explored the association between head size and a range of childhood

cancers. Similarly only one study explored the relation between head size and autism. Both of these studies were at high risk of bias in their findings.

Few studies were identified in relation to adolescent health and those identified addressed only 2 of the review outcomes. A single study found no relationship between growth faltering in infancy and measures of anxiety and low self-esteem. Three studies of infant growth status in relation to adolescent obesity and seven in relation to insulin dependent diabetes were also identified. These were included with studies of adult subjects to enable synthesis.

Fifty-two studies associated infant growth status with measures of child development (45 on cognitive development, 10 on motor development with 3 looking at both). Both infant growth status and motor and cognitive development were measured in many different ways, limiting our capacity for meta-analysis, though some was possible. In general, larger size in infancy - head circumference, weight or length- was associated with better cognitive development in childhood. In addition, studies of children whose growth had faltered during infancy showed that they had a modest reduction in cognitive function. For motor development, the studies were less conclusive. A larger head size in infancy was associated with better motor development. The evidence for an association between weight or length in infancy and motor development was less compelling, though was in the same direction (i.e. larger size, better motor development).

Over a third (n=19) of the studies relating to development were based on children who had been LBW (defined as birthweight less than 2.5kg). There were no studies relating to other review outcomes that focused solely on LBW infants.   Comparison of review findings for the LBW infants and normal birthweight infants revealed that the trends and associations observed were consistent across both groups: infant size in terms of head circumference, length and weight was positively associated with cognitive development, growth faltering was associated with poorer cognitive outcomes and head size in infancy was positively associated with motor development in both LBW and normal birthweight infants.

Assessed against our criteria, the overall quality of the studies in this review was moderate and many carried at least a medium risk of bias. This may in part be a function of the type of research in this field, and we return to this point later (page 159, section 13.4.2). However, it increases the need to be cautious in our interpretation.

Due to a lack of information in the published studies, we had little capacity to address the effect of inequalities in people's lives on these associations. Differences between sexes were sometimes described in the studies relating to IHD, suicide, lung cancer, OA, child and adult obesity and IDDM. However there were few examples of sex differences in the associations between infant growth status and the outcome, and many studies may not have had adequate statistical power to explore these. The only studies demonstrating sex differences were those relating to IHD and suicide based on the Hertfordshire and Finland cohorts. The inverse association between infant size and IHD was present in the men but not the women in both Finland and Hertfordshire. Similarly the inverse association

between infant size and suicide was demonstrated in the Hertfordshire men but not the women.

Very few studies described differences in relationships by socio-economic or ethnic differences.   Only a single study (by Fall et al) reported findings in relation to IHD according to social class[97] Similarly only one study of cognitive development by Nelson et al reported findings in relation to race. It is difficult to assess how many studies had sufficient information to report their findings according to socio-economic status or ethnicity[45]. Many studies, particularly those relating to childhood development, either adjusted for socio-economic factors such as maternal education or reported socio-economic status at baseline. However few studies reported sufficient baseline information to allow assessment of their capacity to group their findings according to these variables and many would have had insufficient statistical power to do so.   Overall 72% of studies considered socio-economic status in their analyses. However this relatively high proportion masks the fact that for many outcomes consideration of socio-economic status was poor with very few of the studies reviewed assessing its effect. This applied to SIDS and obesity in childhood and to non-insulin dependent diabetes and obesity in adulthood.

Overall, this section of the study showed that larger size in infancy tends to be associated with better cognitive development, reduced risk of SIDS and some adult diseases including IHD and non-insulin dependent diabetes, and with increased risk of insulin dependent diabetes. Infants who are at the heaviest or fattest end of the distribution are more likely to become obese in childhood and adulthood. There were significant gaps in the literature which in itself was of variable scientific quality.

13.2.2 Review of lay perspectives

The review of lay perspectives was organised around the overall study aims i.e. what are lay, particularly parental perspectives on infant size and growth; what issues do parents think important in relation to size and growth; where does infant growth lie in terms of priorities of those who care for infants and to what extent is infant growth a salient issue for parents and children.

Nineteen studies were included, most of which addressed the views of parents, usually mothers, although children and adolescents, and health professionals were also represented. Notions of what constituted healthy size and growth were dominated by the concept of 'normality.' This operated through the creation of norms of growth and models to explain difference using both physical, observable characteristics and physical unobservable constructs (such as underlying health status) and non-physical dimensions (such as emotional care). Where growth differed from the norm and a plausible explanation could not be found, growth became an important concern for parents.

It was not always possible to interpret a difference in the meanings of size and growth by participants. As a result, there are some contradictions in the data, for example the contrast between the importance ascribed to growth monitoring compared to the lack of health value given to size alone. These differences may be due in part to reporting, in part to research design, and in part to the picture which

frequently emerges in qualitative research (and every day life) of emerging, changing and less than entirely consistent views.

The story that emerges from the included studies is of a tension between concerns associated with *size* and the low priority of *growth* as a health concern. Studies that considered mothers whose children were diagnosed with faltering growth (sometimes called failure to thrive) reported that they were anxious about the progress of their children. As a result of their worries they were often pleased when their children received medical attention and acted on advice given. Monitoring the growth of children was not just important for those whose infants were small. Weighing babies and comparing their size to others appears to be a common practice and one that mothers consider important, particular so for a first child. In contrast, the reported concerns associated with size and growth in themselves are minimal. Mothers of small infants (those with faltering growth and those born small) are more concerned with possible medical complications, such as infections, associated with their smallness. The focus on observing growth is not restricted to mothers. They reported interest in and comments from, relatives, strangers, health professionals and welfare agencies on the growth of their children. Sometimes mothers felt that too much is made of growth alone, when children are otherwise well. Overall there was more concern expressed by mothers in these studies for under- than over-weight. As well as receiving advice and help from individuals, mothers also looked for advice from books and magazines. None of the studies sought opinions from mothers about the acceptability or advisability of intervening to affect rate of growth.

13.2.3 Focus groups and interviews

In addition to the systematic review of lay perspectives on infant size and growth, we carried out exploratory primary work in East London to interpret and build on the findings from the review. How salient might the review findings be in a different context, and what might the implications be for intervention or future work? We conducted 34 screening interviews, and 5 focus groups with a total of 12 individuals to explore participants' views of infant size and growth and the possibility of interventions to change growth rate. The broad themes identified in the systematic review were also found in the interviews.

All participants referred to the importance of "normality" in infant size. This was reflected in the ways that normal size was defined and assessed, and the ways in which differences between infants were explained. Variations in size were acceptable, and where size was seen as unacceptable, the causes of this "unhealthy" size were usually discussed.

Growth was seen as indicative of the nature or temperament of their children. Mothers described children as "good eaters" and associated this with good behaviour. Parents were concerned that if they could see no cause for unusual patterns of growth then something was wrong, and this was very likely to be a health problem. Participants often felt that children were simply going to be the size that they were going to be, and that there was little they could do to change this. This was an optimistic feeling, in that provided the child was well cared for, size was not something that parents needed to worry about. More powerful than the fatalistic notion of growth, was the sense that norms were simply inappropriate. The concept

of individual variation was used to explain differences between children, and to account for patterns of growth that others might want to categorise as 'abnormal'. For some the very act of defining a growth path was alien to their system of understanding. Parents talked about their preferences for the way babies looked, and on the whole this was a preference for larger babies. As with participants in the systematic review, growth in itself was not a cause for concern for our participants, but it was seen as an indicator of underlying problems. If children weren't growing appropriately that might be a sign that something was going wrong. Growth was something that parents and other family members worried about because it might mean there was a problem.

We identified many gaps in the research literature exploring the views of those around babies to their size or growth. These included lack of information from family members other than mothers particularly the views of fathers, views from those whom we know to have breast fed, bottle-fed or to have weaned their babies, views of parents of first or subsequent babies and views of different ethnic or cultural groups. In addition there was a lack of studies which combined qualitative and quantitative data. For example the study that test weighed preterm infants before and after feeding would have benefited from further quantitative data (infant outcomes as well as maternal outcomes) and qualitative data (describing the response of the mothers to the intervention). The limited scope of this study means that while we can state this intervention was not effective for mothers, it does not contribute to our understanding of what might be helpful to mothers or babies.

13.3   Discussion of the methods and process

13.3.1 Developing the study question

Like many projects, the precise origin of the question which this study was designed to answer is obscure and we are reliant on our memories (and hence subject to recall bias). Three of us (CL, JB, HR) had been present at discussions in various fora as to whether the evidence relating early growth to adult disease (especially cardiovascular disease) could be used to inform policy or to design an intervention to change infant growth. We believed that these discussions were not sufficiently broad in their scope for three reasons. Firstly, we felt that other adult outcomes apart from cardiovascular disease should be considered. Secondly, we felt early life should not be regarded simply as a preparatory phase for adulthood, but as of value in itself, and therefore that earlier outcomes should be included. Thirdly, we believed that, on principle, parents' views of infant growth were critical, and in practice, interventions to change infant growth would probably depend on parental co-operation, and might well be built on existing expertise among parents, rather than drawing on a deficit model.

With these beliefs in mind we posed a related but different question. What is optimal infant growth? As this question is conditional (optimal for what?), we then defined our methods based on the beliefs stated above, and these methods are discussed below. We recognise that this has influenced our design and therefore the results of this project. For example, had we posed the question, "what is optimal infant growth for adult health", we might have come to different conclusions, because we would have been reviewing only a proportion of the studies described in this report, and

the participants in these studies come from a limited range of populations and periods of history, and have been studied by a small pool of researchers. We might also have included other studies of risk factors (for example, blood pressure or the normal distribution of body mass index). Whilst our review of lay perspectives might have reached similar conclusions, the discussions at the focus groups would have been differently structured and emphasised, which is likely to result in different findings.

13.3.2 Basing the methods on a life course approach

The optimum is the condition, degree, amount or compromise that produces the best possible result (Collins Concise Dictionary). Thus, optimal infant growth can only be defined in terms of this result. For the review of infant growth in relation to life course outcomes, we defined the result as health across the life course, as we believed that the public in general, and parents in particular, would attach importance to early as well as later health, and to 'being' as well as becoming – in the sense that the well-being of one's child in the here and now is important to the parents of infants. We also believed that they would value health both as freedom from disease and as a more positive feeling of wellbeing. Thus our framework of outcomes was developed to capture experience of health across the life course. As well as including outcomes based on disease (morbidity, mortality and longevity) we also included outcomes on health related behaviours, health related quality of life, non-health related quality of life, and, in studies of children, development. This framework allowed us to weigh up the associations of infant growth with early and late outcomes, and the possibilities that changes in infant growth (for example, increased weight gain) would be associated with health gain or harm. However, it meant that the search strategies were complicated, and resulted in a heavy screening load. It also revealed many gaps in the research evidence.

The review of lay perspectives was designed to assess the views of lay people, on infant growth. As such, we sought to understand their views of growth and size and to describe their experience of infant growth. The focus groups took a similar perspective. In addition, focus group facilitators considered with participants the relationship of infant growth to earlier and later health, as well as the issue of when intervention might be appropriate and in what form.

13.3.3 Developing the outcomes

The framework of outcomes used for the review of infant growth in relation to life course outcomes was developed for this study as no existing suitable framework could be identified (see page 15, section 2.2.1). It was based on the premise that both the public and policymakers would be most concerned about health states which are associated with a high burden of disease or disability. Thus, we based our adult health outcomes on the Global Burden of Disease Study[10] which attempts to quantify the burden of different diseases and risk factors in a common unit of disability adjusted life years (DALYs). However, this results in most burdens being ascribed to diseases of adult life. As there were no equivalent studies for the period before adulthood, we constructed the framework for infant, childhood and adolescent outcomes on a similar "burden of disease" approach (that is including preferentially those health states that are common and/or serious to the individual) based on information from the published literature and, where this was lacking, a theoretical approach. However, our basic premise could be challenged. Whilst

burden of disease is an important consideration in health policy making, policymakers must also take account of rarer but serious (and therefore newsworthy) diseases, particularly if public policy may be implicated in their aetiology. In the UK, recent examples include autism and Bovine Songiform Encephalopathy (BSE). Nonetheless, we think that selection of outcomes based on burden is appropriate for a review aimed at informing policy, as health impact and cost effectiveness are important considerations in policymaking.

If we accept the rationale for a burden of disease approach, our choice of the global burden of disease study as a point of reference could be challenged, as the study has been criticised for its cross-sectional nature and the fact that much of the vital registration data on which it is based may be incomplete or unreliable, although this is less likely to be a problem for developed country estimates. One of the main drawbacks of GBDS is that DALYs estimates are based on a set of assumptions and as such have wide confidence intervals around them. Nonetheless, it was useful in allowing us to identify the diseases carrying most burdens, and to construct our framework specifically for informing policy in developed countries (the GBDS ranks causes of DALYS, for the world, and by economic status of countries). Due to lack of literature at the time we were developing our protocol, our framework for pre-adult outcomes had to be constructed in an ad hoc fashion, and with some judgement as to what were the most important components of each health domain. However, projects published recently on or using indicators of childhood health (and after our protocol had been agreed) show significant overlap in selection of indicators with our study outcomes particularly in relation to selection of non-health related quality of life and important chronic disease of childhood which influence health-related quality of life, including asthma and diabetes[166;167].

A further area of debate is the way in which our framework influences the research evidence that we reviewed. The way in which we constructed the framework meant that all health domains were considered explicitly at 4 periods of life – infancy, childhood, adolescence and adulthood. This is likely to weight the selection of studies to those looking at earlier (pre-adult) rather than later (adult) outcomes. It may also mean that the high burden of disease experienced in late adult life from causes of DALYs that are not in the top10 diseases in the GBDS is under-represented in the review. We believe this approach may be consonant with parents' views on the meaning of growth. Parents, as one might expect, did talk more about earlier than later outcomes, but when presented with potential effects in early and later life did not comment on the relative importance of these.

The methods were developed in order to inform an intervention (either as a policy or further research) directed at infant growth. These might differ from a review aimed at informing whether infant growth caused particular health states. For the review of infant growth in relation to life course outcomes, an aetiological review might focus only on biologically plausible relationships, whereas our review included all relationships, whether biologically plausible or not. Our reasoning was that plausibility may have different meanings to scientific and lay audiences (for example sections of scientific community and the lay public take different views on the plausibility of a link between measles, mumps and rubella (MMR) triple vaccine and autism.

13.3.4 The methods of the review

We followed the methods recommended by the Centre for Reviews and Dissemination at the University of York[9]. These are based on experience of trials and other types of intervention studies, and on observational studies. However, experience of the latter is relatively limited and that of integration of quantitative and qualitative evidence even more so[122;129;131;137;164]. Methodological development was time-consuming, and, in retrospect, we should have requested more time for this aspect of the study, although this would have increased costs. In particular, construction of the framework of outcomes required several iterations, and on the qualitative side, finding a way of presenting the findings which did justice to their scope, while attempting to make them sufficiently succinct to be meaningful was a challenge. Similarly design of the databases and quality assessment tools (appendix 5) were iterative, and it was necessary to have some experience of data abstraction to grasp the range of methodological and quality issues associated with quantitative and qualitative studies of infant growth. Furthermore, both the review components (of lay perspectives and of infant growth in relation to life course outcomes) included different study designs. This made it necessary for the databases for abstracted information and the quality assessment tools to be sensitive to such differences. For example, separate tools were developed for case control and cohort studies, sharing common criteria where appropriate, but also recognising the fundamental differences in design.

The approach adopted in the review of qualitative literature had to account for the differences in study designs within a far smaller pool of studies (19). Two methods were adopted to allow abstraction across these. Firstly, study quality was assessed in relation to the utility for this review, as opposed to theoretical considerations of study design. Secondly, findings were abstracted by interrogation with research questions posed by the review. The source of data (conclusions of original author or data abstracted by reviewers) was made explicit.

13.3.5 Searching the literature

Most of the studies in the review of life course outcomes were observational. Searching for observational evidence presents particular problems in that, unlike randomised controlled trials, the terms used to index observational studies vary widely and it is not possible to use a study design filter to identify relevant papers in bibliographic databases. Therefore it was important that searches were of sufficient breadth to identify all study designs. A further issue in this review was that we expected there to be relatively few studies in relation to some of the outcomes. It was important that our searches were of high sensitivity in order to maximise our chances of finding studies. To this end our search strategies were comprehensive with search terms used to define outcomes often numbering in excess of 50 and 13 terms being used to describe infant growth status. Where outcomes were grouped, for example in relation to mental illness, outcomes terms were often in excess of 100.

Due to the need for high sensitivity our searches lacked specificity and so many irrelevant abstracts were identified. Overall more than 160,000 abstracts were identified through electronic searching but only 108 relevant studies met the inclusion criteria. This is unusual and low even for a review of observational

evidence. Comparison with other observational systematic reviews suggest that the specificity of searches is directly related to the questions being asked in the review. For example, in the review of water fluoridation by CRD which addressed a much narrower set of questions, a relatively high proportion of references screened met inclusion criteria: 3,200 studies were identified of which 254 (8%) met inclusion criteria[14]. The scale of our review, which was addressing a whole series of questions, together with the lack of biological plausibility of some of the associations we were exploring are likely to be the main factors contributing to the scale of the searching task. We had intended to search the foreign language literature for relevant studies. However, due to the time constraints of the review and the large scale of the searching task, we were not able to do this for all outcomes. We focussed our attention on childhood development, given that this was the area where we had identified most studies. A search of LILACS did not lead to identification of any development studies that we had not already identified in our other searches. We cannot say for certain that the same result would have been observed had we performed foreign language searching for other outcomes. However, this experience suggests that it is likely.

Searching for qualitative work raises particular problems, which are described in the body of the report, and the more sophisticated search mechanisms in the medical/health databases are likely to have introduced bias.

Search techniques for qualitative and views studies are not well developed, and there was a potential risk that by using search terms for attitudes and views, studies would be missed. To broaden the scope of the searches, all searches conducted for the review of outcomes were also reviewed for inclusion in the review of views of growth. A similar approach has been taken in other systematic reviews combining qualitative and quantitative findings[122-124]. This approach is thorough but time consuming and in this case proved to be inefficient, since no new studies were identified.

13.3.6 Synthesising the results
We faced a number of challenges in synthesising our results across all components of the project. Firstly, the methodological challenges to synthesis of different study designs and of quantitative and qualitative research are daunting and there are few examples in the published literature on which to build[124;129;136;164]. Second, the component parts of our project were all aimed at addressing an overarching question, "should we intervene to change infant growth?" However, each component asked a slightly different, subsidiary question. The review of infant growth in relation to life course outcomes aimed to describe the association between infant growth status and each of a range of outcomes. The review of lay perspectives and the focus groups aimed to collate lay views on the meaning, interpretation and importance of infant growth. Other models of synthesis have used different study designs to answer the same question. For example, a recent study synthesised studies of children's views of eating fruit and vegetables with interventions to promote their fruit and vegetable consumption. Both types of studies were aimed at an overall question, "what is known about the barriers to and facilitators of fruit and vegetable intake among children"? Synthesis of the qualitative studies allowed this author[164] to develop post hoc criteria for intervention strategies which could then be tested against the effectiveness of the intervention

studies reviewed. Our synthesis of lay perspectives does not allow us to re-interpret the associations of infant growth with health outcomes across the life course. However, it may allow us to discuss whether the associations observed are likely to be of concern to parents.

At a more technical level, we also faced challenges in synthesis within each component of the review. In both reviews there were relatively few studies, particularly for subgroup analysis. In the review of infant growth and life course outcomes, small numbers of studies for each outcome prevented synthesis in all but a few areas. Where there were sufficient studies, meta-analysis was often impossible, because of differences in exposure and outcome measures, and heterogeneity of effect. As well as heterogeneity in study design, the review of lay perspectives struggled with heterogeneity in study aims and populations.

To avoid bias once the literature had been accessed, narrative synthesis was constructed around health outcomes for the review of infant growth in relation to life course outcomes (appendix 3) and around pre-specified questions for the review of lay perspectives (section 6.1).

Within each outcome review we grouped studies according to the aspect of infant growth status being considered and in relation to the scientific question being addressed in the study. This allowed us to synthesise the results from studies considering the size of infants separately from those that considered growth. However, this strategy presented some challenges for those studies where specific patterns of growth were being examined. For example we grouped studies that set out to explore the relationship of failure to thrive (or growth faltering) with cognitive development separately from studies which examined the whole range of weight but did not include a definition of failure to thrive. We assessed the potential for this strategy of grouping studies to bias our overall findings and found that it did not. For example, both lower infant weight in the normal range and growth faltering were consistently associated with poorer cognitive development.

In general, we confined sub-group analyses to pre-specified categories, with some categories common to both reviews. However, the major limitation on sub-group analysis was lack of evidence or a small number of studies. In particular, the sub-group analyses discussed at the beginning of the project to examine dimensions of inequalities in infant growth and health, and infant feeding, were not possible across the whole review because of lack of primary studies.

The development of the methods for the review of lay perspectives was incorporated into this project. Since only a small number of studies were identified (n=19), thematic and narrative syntheses were undertaken and both are reported here. Reporting both allows readers to consider the strength of each method and the fitness of each for the data type and for the purposes of future reviews.

In the review of infant growth and life course outcomes, basing the narrative synthesis primarily on life course outcomes was to some extent a pragmatic decision, as it matched our searching strategies. These reflected health research paradigms (a physiological systems based approach based, to a considerable

extent, on disease), which in turn shaped the research literature available for assessment.

13.3.7 Discussion of the process

This study was designed to use multi-disciplinary methods in an integrated way to address a question which we believed to be of policy and public health importance. The study was planned by four members of the study team, who also applied together (successfully) for funding for the study (the others were employed once the study had been funded). The team met at regular intervals during the study, and in smaller groups, communicated by email and teleconference, and circulated notes and other information. We had three meeting with our expert advisory group, with some of whom we also had helpful contact between meetings. At the end of the study, we feel that we have succeeded in working together, but we would like to reflect on the challenges to this and the limitations which we found.

Our major challenge was the combination of geographical distances between our workplaces, the usual time pressures of academic life, and uncertainties in the scope of the study, given that this type of review is relatively unusual. We were spread across three academic units at the beginning of the study and five at the end. Whilst some of us knew each other well, others had to make new working relationships at a distance. Travelling added to the time commitment we each had to make for meetings, and we were unable to have the opportunistic conversations about the study that might occur had we been located in the same place. We were conscious that some members of the team felt better able to communicate in teleconferences and by email than others. Planning for meetings was hampered by uncertainties on the scale of the study, which meant that we often had to readjust our timescales.

An advantage of distance and "virtual" working for this and other studies, is that it allows combinations of skills, expertises and perspectives to be brought together independent of institution and location. It is unlikely that any of the institutions in which the team were located at the start of the study could have mounted such a study using only their own institutional resources. Based on our learning from these processes we would in future consider meeting on a regular basis (for example, every six weeks), rather than planning to meet at particular milestones of the study. This might seem luxurious, but would allow regular focus on the common ground of the study, provide an opportunity for ad hoc conversations, and would mean that uncertainties in the timescales of the study would become less influential in determining meeting times. We also think that studies involving several institutions may need additional time to carry out work, particularly if it is the first time investigators are working together, and that this additional time should be considered in study planning and funding. Finally, we would recommend that arrangements for management, meeting and communication are set out at the start of the study (and in applications for funding), taking into account individual team members preferences and experience of different methods of communication. With increasing use of virtual working, some researchers may find it helpful to have training in specific communication skills (e.g. teleconferencing).

13.4 Discussion of results

An extended summary of the results is given at the beginning of this discussion (page 137, section 13.1). In addition, there is discussion on specific topics within particular sections of the results. Thus this section of the report aims to draw out generic issues across the whole study, integrating the three strands of work.

13.4.1 Did we fulfil our study aims?

Our study aimed to synthesise the scientific evidence and lay perspectives to define optimal infant growth. Whilst we were able to meet this aim by carrying out a synthesis of the available research evidence, we found that this evidence was insufficient to define optimal infant growth. We lacked information on the size of effect of the relation of infant growth to outcomes across the life course, and the range of outcomes for which associations with infant growth were described was limited. The literature review on lay perspectives found that much literature was not suited to addressing the question of defining lay views of optimal growth. What information existed also tended to emerge in focus group discussions, which suggests that what was salient in these (mainly North American) studies was also salient in the current UK context. Furthermore, many of the studies included in our review were subject to some degree of bias. These issues are covered in more detail elsewhere in the report.

Our second study aim was to identify gaps, to assess the implications of those gaps and to make recommendations for addressing them. For the review of lay perspectives there were large gaps in literature which does not derive from a health research paradigm. Perhaps because health-related databases are "search-friendly", the studies which we identified and reviewed tended to be dominated by those with a health focus. The research literature on lay perspectives of other meanings of infant growth – in terms of fashion, "attractiveness" and other non-health related factors – were not accessed using our methodology, despite our efforts to cast a broad net. The review of lay perspectives also suffered from gaps within included studies, particularly a lack of description of participants.

In the review of infant growth with outcomes across the life course, we were unable to locate studies examining the association of infant growth with many significant causes of morbidity and mortality throughout life, with health related behaviours, and with many aspects of wellbeing, particularly in adult life. Particularly for adult outcomes, we found that a limited number of studies and researchers contributed to the included studies. There were few longitudinal studies and so, despite taking a life course approach in our choice of outcomes, most of the evidence we considered looked at cross-sectional points in the life course rather than giving a true life course perspective. This represents an important gap in the literature. Furthermore, though inequalities in infant growth and in outcomes across the life course are well documented, few of the included studies allowed us to see the extent to which they influenced the observed relationships. Similarly, though infant feeding is an important influence on infant growth (and the target of potential interventions), few studies allowed us to assess its influence.

We have reflected on why these gaps exist. Across the wide range of subject areas reviewed, we do not pretend that there are simple answers which apply equally to

all areas. However, there are some common features. First, research design may be an issue. For example, the research between a putative exposure (infant growth) and a disease with a long latent period (cardiovascular disease or cancer) is challenging logistically. Prospective cohort studies are rare and none yet exist which can answer this question although some studies will soon have that capacity[168;169]. The main limitation of many of the studies within the review of lay perspectives was that they were context specific in their orientation. All but two were explicitly undertaken within the context of health care, making it difficult to untangle the reported importance of a medical model of growth from the health care context in which the studies took place.

Second, research paradigms in biomedical science in particular, and to some extent in public health or social science favour studies of single systems, diseases or social constructs. Thus some researchers may concentrate on, for example, cardiovascular disease, and others on respiratory disease. These disease-based paradigms may be reinforced by the success of individual researchers, which is likely to reflect their personal qualities and funding streams, as well as the importance of scientific discovery. In social science too, funding streams affect what gets done, and some kinds of research are more susceptible to funding than others. Qualitative work is probably more likely to attract funding now than it was a decade ago, and mixed qualitative and quantitative work has become relatively attractive to funders even more recently. Some important gaps looking for instance at general or specific aspects of well-being in children may have been relatively neglected in large-scale studies because of the problem oriented approach of much research. The same is true of an emphasis until relatively recently on risk rather than resilience. The growing tendency to seek advice from 'users' on the formulation of research questions and in research commissioning may change this over time.

Third, the research process often depends on our ability to measure parameters of health or function. In some areas, risk factors for cardiovascular disease, for example, or measures of intelligence, we have a range of measures to use. In others, for example, biological risk factors for many cancers or indices of motor development or function, we have fewer measures, and this lack of knowledge perpetuates research gaps. In our review, we decided not to include risk factors for adult disease. We are aware of many studies of cardiovascular risk factors in relation to early growth, and few for risk factors for other disease (for example respiratory function). We felt that including risk factors would have implicitly biased the weight of evidence to cardiovascular disease.

13.4.2 Quality of studies
We constructed quality assessment tools for the review of life course outcomes which varied according to the study design (appendix 7). For the review of lay perspectives we used the system of Popay et al (1998)[132]. Though these were based on standard tools it is possible that our amendments or application to this subject resulted in methods which were overly harsh in the quality standards expected. We did not use quality of study as an inclusion or exclusion criterion, nor did we apply a "fatal flaw" rule (that is if a study has a high risk of bias for a single scoring item, then it has a high risk of bias overall). However, we found that many of the studies included in the review were, at best, only moderately high quality and so carried a medium risk of bias. In the review of infant growth and life course

outcomes just under 20% of studies were judged to have a low risk of bias. There is a particular bias in the review of lay perspectives because all the work we located was based in the health field and conducted almost exclusively by health professionals. This is likely to have created a bias towards 'medicalised' explanations in the nature of the questions asked and data collected. Our conclusions overall are, in any case, cautious because of the overall lack of research evidence. If we take into account the risk of bias in that research evidence, our conclusions become even more cautious.

One possible reason for the medium to high risk of bias in many of the studies may be due to the nature of observational research. For example, many of the studies of infant growth in relation to adult morbidity and mortality relied on historical cohort studies. Unless there are comprehensive systems for record linkage and/or tracing, a high proportion of loss of participants is extremely likely in such designs. A second issue is that we interrogated studies to answer specific questions about associations with, or perceptions of, infant growth. However, the studies were often set up to answer a different question for example, the relation of growth from birth to adolescence to an adult outcome or parents' experiences of child health services. Areas where studies were often considered to have high risk of bias were in consideration of confounding factors of importance in the relationship between infant growth status and outcomes and in high rates of attrition. Our findings concur with those of recent research which suggested that many published epidemiological studies are not of high quality particularly in relation to consideration of confounding[170]. However, it is possible that had our reviewed studies been designed or written to answer "our" questions, they may well have been judged to be of higher quality. For example in the review of infant growth in relation to life course outcomes, studies could have taken more account of potential confounders.

In the review of lay perspectives, in order to think about why we considered studies poor for our purposes, we considered what the ideal study might have been. For the purpose of this review, the ideal study would have asked about lay (particularly parental) views of the importance and meaning of early growth, with a view to understanding whether growth adaptation is likely to be acceptable or even welcome to parents. Such a study would have been likely to have both qualitative and quantitative elements, and include a well designed survey, both preceded and followed by qualitative methods exploring subjective experience, with a robust design and sound execution. Our search strategy did not identify such a study. Instead the topic of growth/size emerged from data from studies with diverse study aims. The single largest quality problem in the qualitative studies related to poor reporting of qualitative data, and in some cases, poor reporting of salient participant characteristics.

In the review of life course outcomes, the ideal study for addressing most of the outcomes considered would be a prospective cohort study set up to explore the influence of infant growth status on early and later health. Such a study would use systems such as record linkage to minimise attrition and would collect information on demographic and socio-economic variables likely to confound the associations of interest. The exception would be studies of rare outcomes such as SIDS where a case-control design is appropriate.

There were too few studies to be able to perform a sensitivity analysis i.e. considering whether our results would have been different if only studies of high quality were included. Instead, we considered quality in our narrative syntheses, particularly in considering inter-study differences

13.4.3 Infant growth or size

Much debate has focused on whether it is attained size or the pattern of growth that is associated with a variety of outcomes. We aimed to describe associations with and lay perspectives on either infant size or growth, and furthermore with any measure of size or growth (for example, weight, body mass, head circumference). However most of the studies we reviewed considered size rather than growth. We also had difficulties integrating some information, as the terms growth and size were sometimes used imprecisely in the scientific literature and the definitions used for particular patterns of growth status (e.g. growth faltering, infant obesity) varied widely between studies. Furthermore, we are unsure how lay people in some of the lay perspectives literature would have defined these terms, which are often used inter-changeably in common conversation.

Both the literature and the interviews suggested that parents struggled to identify what was and what was not appropriate growth. Some participants in our interviews said that consistency of growth was what was crucial regardless of where on the chart you were. There was also discussion and a literature on catch-up growth, where reaching a particular size was key. What was looked for was not following a line on a chart, but being in the correct clothes size. This was particularly so in relation to faltering growth and premature babies.

In the review of life course outcomes, the use of different measures of either size or growth across studies limited the numbers of studies with similar measures which we could group for synthesis. However, it was possible to group a number of studies considering head size, weight or growth faltering in infancy in the review of cognitive development. The predominant description was of infant size, particularly weight, and the paucity of studies considering infant growth represents an important gap. This probably reflects the relative ease of measuring size. For the outcomes of adult and childhood obesity, researchers often looked at the relationship of these outcomes with being at one end of the growth distribution in infancy, usually infant obesity (variously defined). The variability of descriptions of size and growth in infancy is a serious limitation to synthesis of the evidence. A number of studies carried out more than one set of measurements during infancy but did not then go on to explore the association of infant growth with the outcome. For example, of the 25 studies of head size, weight or height and cognitive development 18 studies took measurements at more than one point during infancy but only 12 of the studies reported findings that made use of these measurements. This suggests that accessing raw data in order to carry out secondary analyses might address some of the current gaps relating to infant growth in the existing research.

13.4.4 Are the associations described causal?

We set out to describe the associations of infant growth with outcomes across the life course. It could be argued that, unless it is known whether infant growth causes these associations (or is on the causal pathway), this information is of only theoretical importance. However, at present, we believe that our ability to infer

causation is limited. We could, for example, apply the Bradford Hill criteria[171] to the evidence we have collated, but we have not sought to answer some of the questions that the criteria pose (for example, whether the association is biologically plausible). One of the classical methods for inferring causality is by observing the effect of experimental manipulation of the exposure variable (in this case, infant growth) on an outcome. Animal models might be used, but these will always be limited in their application to humans. Experimental intervention in infants would have to be justified by a systematic collation of relevant evidence, including explicit consideration of the possibility for harm[8]. We hope that our review can contribute to that systematic collation of evidence. Thus, we do not assume causality, but argue that our work may be used in this debate.

In the work on lay perspectives, size and growth were seen as part of a causal relationship with outcomes in childhood and adulthood, but not as the antecedent cause. The causal relationship most often identified was with medical outcomes. The broad themes identified within the review of the literature were replicated in the interviews, as illustrated by the views of one mother:

"Zoë": you don't know what's worse, having one that's big or having one that's small. When they're big they worry about the diabetes and other things and when they're small they worry about getting infections, all different things."

Some participants felt that advice to restrict or promote growth in babies and young children would be seen as untrustworthy. Parents would consider intervention, but it needed to "make sense", to conform to their understanding of the causes and consequences of unhealthy growth. When presented with possible causal links, participants did not consider consequences in childhood health or well-being as more plausible that adult outcomes.

13.4.5 Context of studies
Observational studies need to be considered in their context, both in time and place. Of necessity, those studies which examined infant growth in relation to adult outcomes were conducted on people whose infancy was experienced decades ago. Their lives would have been different in many ways to today's infants. With relevance to growth, they are likely to have had a different experience of breastfeeding and weaning, if bottle fed, the composition of their milk would have been different, they would have had a higher burden of infection, their family social structure would have been different and the educational attainment and relative income of their mothers is likely to have been lower. The historical context is important because there have been considerable changes in childhood growth in recent years, with secular trends in increasing childhood weight and body mass, even at young ages. In areas where we had sufficient studies, we were able to assess whether the association of infant growth with some outcomes were the same in studies carried out many years ago and those carried out more recently. In the review of childhood obesity 6 of the 9 included studies were published before 1979. The findings of these studies were less conclusive than the studies published more recently with only 2 of the older studies demonstrating the statistically significant positive association between infant growth status and adult obesity that was observed in the more recent studies. However, the quality of these older studies was also lower than the more recent studies (with 2 studies having high risk

of bias and the remainder being of medium risk) and they were also of smaller sample size. Given these methodological weaknesses, which may in part be due to less rigorous requirements for journal publication in the past, it is difficult to place too much weight on the differences in findings of earlier and later studies.

Across the rest of the review there were relatively few studies published prior to 1979: 4 of the 40 studies related to cognitive development, 3 of the 8 studies relating to SIDS, 2 of the 9 studies relating to adult obesity and 1 of the 7 studies relating to insulin-dependent diabetes. In all cases the earlier studies were of lower quality than the more recent ones and were more likely to have inconclusive findings or findings that were inconsistent with more recent studies. However given there were relatively few studies, their influence on the overall findings of the review was limited.

We believed that the literature on lay perspectives on infant growth would be particularly sensitive to time, and wanted it to be relevant to today's infants. We therefore set a date limit on our inclusion criteria, to include only those studies published in the past 25 years. This period of time was chosen to represent a generation.

We did not exclude reviews on the basis of where they were conducted, but our review was primarily designed to inform the development of public policy in developed countries, and particularly in the UK. We therefore included studies which could address our questions in relation to a UK context. Thus, for example, a study in a developing country relating infant growth to childhood asthma would be included, but one which related infant growth in the presence of malaria to childhood asthma would not. In the lay perspective work, we also tried to include studies that considered cultural influence by including studies from countries that contribute significantly to current UK population. Thus a paper considering the views of mothers living in townships in South Africa to overweight was excluded since the context for these mothers was food shortages and an absence of medical care, making this group substantially unlike most mothers in the UK. In practice, exclusion of studies on these criteria excluded very little research.

For the review of lay perspectives, only 3 UK studies and one other European study met our inclusion criteria. The remainder were North American. This presented us with some contextual difficulties in interpreting issues such as ethnicity where the views of Mexican American or Latino mothers may be less useful in a UK context. Low income and the ways in which health and medical services are provided also introduced contextual problems. Our exploratory direct work with a UK sample was intended to address some of these contextual issues.

13.5 Conclusions

13.5.1 What this study has added, its limitations and its strengths
The study has synthesised a broad range of literature on the associations of infant growth with health across the life course and integrated that with a review and qualitative assessment of lay perspectives on infant growth. Taken together, our view as researchers is that the evidence is not sufficient to propose the development of an intervention to alter infant growth with a view to improving adult

health. Though many beneficial health outcomes and developmental milestones in childhood are associated with larger size during infancy, both insulin dependent diabetes and obesity in childhood and adulthood are also associated with larger infant size. The evidence reviewed had, in general, at least a medium risk of bias, and large gaps in the consideration of important outcomes (for example, childhood morbidity). Analysis of the literature on lay perspectives concluded that parents viewed growth as an indication of underlying health status, but not as a dimension of health in itself. Participants in the focus groups were not in favour of intervening in an infant's growth. Thus there is no evidence to suggest that interventions to alter infant growth are high on the agenda of parents, but we cannot conclude this with certainty without further study.

In conducting our review we tried to break out of established research paradigms to consider information relevant to policy. Thus we have synthesized literature that is not usually considered together (for example on causes of infant and adult mortality), reflecting policy needs to consider the whole age span of the population and individuals within it. In doing this, we hope we have contributed to methodology of evidence synthesis. We have identified and developed our work to meet (to some extent) the challenges of synthesising different types of evidence, incorporating user perspectives into review and considering the entire life course. The framework of outcomes which we developed for this review is now undergoing further development and should in future be useful for other study types (for example public health evaluations) as well as systematic review.

We have also tried to reflect on the process and experience of our work, and to document it in this report, to inform those who commission and carry out research.

We acknowledge our limitations as researchers. Across such a broad field, we were inevitably considering research literature in which we were not expert. We sought to attain high methodological standards but some of our aspirations had to be sacrificed due to time constraints. For example, we could not carry out foreign language searching for all the review outcomes. The resource required to carry out the study was difficult to estimate at the outset, because of the broad nature of the subject and innovative approach, and had we had unlimited time and resource we could have improved our technical processes, though it is uncertain whether this would have altered our conclusions.

As in all systematic review, we were limited to the available literature, mainly that which was published. Whilst we tried to ascertain where unpublished analysis existed, our capacity (and that of any team attempting this) was limited to contact with those known to be working in this field. We did not attempt to access data to carry out analyses ourselves, as this would have been unfeasible in the time available across this broad range.

A particular limitation in such a broad ranging review is that the results are multifaceted and therefore the answers to the study questions are complex. They depend, in part, on the weight of importance attached to different types of information and these weights are likely to be subjective and vary by user. For example, a woman whose older child is obese and being bullied may wish to avoid childhood obesity for her infant, and so place a greater weight on the avoidance of

high infant weight i.e. she places a high value on childhood outcomes. Conversely, a woman who had recently lost a close relative to ischaemic heart disease, might value the avoidance of small size in infancy (regardless of whether this is a causal relationship to IHD), and place a higher value on adult outcomes. A policymaker or health professional may be particularly anxious to avoid causing harm – for them, the gaps in the evidence (with the spectre of unknown harm) may be of great significance when advising on patterns of infant growth. Whilst synthesis of the evidence can inform these values, it does not decide them.

We believe that the study demonstrated some strengths. It had a multi-disciplinary team who brought their own and the strengths of several institutions to the study. We used transparent methodology, basing it wherever possible on standard methods and adapting it for this approach, though some innovation was necessary. We tried to adopt a self-critical approach, and to report our findings openly. It was undoubtedly a strength of this study that it benefited from the expertise of an advisory group, skilfully chaired, and was supported financially by the Department of Health for England. Full details are given in the acknowledgements section and in appendix 1.

13.5.2 Implications

For parents/carers: we hope that parents and carers may find the knowledge of what information is and is not available to link infant growth to health across the life course helpful. We do not believe that our findings indicate any change in principle in the way parents currently care for their infants. However, we note the value placed by parents in the review of lay perspectives and the focus groups on "being normal" - that is, being like everyone else. If trends in infant growth continue towards greater fatness, the definition of being normal may well include infants who are fatter than in the past. Conversely current widespread concern about levels of overweight and obesity made lead to misplaced anxiety among those who are unlikely to feel the ill effects of overweight. One possible mechanism for changes in perceived norms suggested by both the systematic review and the focus groups, is the use of clothing sizes to judge appropriate size. Changes in manufacturers' sizes driven by fashion or by changing body shape might directly influence judgements about appropriate size made by parents.

For policy and practice: as stated above, we do not believe that the balance of evidence which relates infant size or growth to life course is sufficient at present to justify a policy or practice intervention to alter infant growth with the aim of promoting adult health, nor do we think such an aim is currently a priority for parents. Firstly, there are large gaps in the evidence base. Secondly, a single optimal pattern of infant growth cannot be defined quantitatively, nor can desirable cut off points be defined. Thirdly, we have no evidence to suggest parents would support this and limited evidence to suggest that they would not. However, there is good evidence that infants in the highest part of the size distribution for weight or body mass index are more at risk of adult and childhood obesity. A policy priority might be to define the characteristics of this group of infants and to assess whether specific interventions in this part of the infant size distribution should be developed and tested. In the absence of further evidence, a precautionary principle approach might be needed.

For research: many of the large gaps in the evidence base may be plugged relatively easily. For example, ongoing cohort studies will be able to examine the relationships of infant growth to childhood morbidity, and some will have the capacity to test the influence of ethnicity, socio-economic status and infant feeding on those relationships. Examination of routinely collected datasets (for example child health surveillance) may also prove helpful in the future, particularly once record linkage methodology is developed.

We also need to understand how, why and in what way infants grow differently. Some of this will require basic science research, but insights may also be gained by looking at the predictors of patterns of growth within cohort studies. Examination of patterns of growth during infancy should not only focus on the postnatal period but should take account of the trajectory of growth beginning in utero. Whilst many interventions can be tested and implemented without full knowledge of the causal relationships between exposure and outcome (as has occurred for many years with smoking and cancer), such knowledge may aid the design of interventions and optimise the effectiveness.

Filling the gaps in knowledge of the association of infant growth with adult outcomes is more challenging. The cohort studies featured in this review may be able to test some of these relationships. However, these are limited in their geographical and ethnic diversity. Supplemental information may be required using pre-morbid states in other cohorts, or using different study designs. Further research will also need to consider the applicability of these associations to today's children. Given that the pathways between early and adult outcomes remain poorly understood, future research will need to take account of a range of social and biological factors that are related to both growth and later health, including infant feeding and socio-economic status.

As discussed in the comparison of narrative and thematic syntheses of the review of lay perspectives we can perhaps draw different implications for research from each method. If we focus on the narrative synthesis we can see where gaps in the existing body of literature lie. For example, we need to be able to compare the views of parents with different experiences to growth and size. Do parents of average, big and little babies have the same concerns? We also need to understand how the information we give is used by parents. The thematic analysis suggests some of the routes by which parents are influenced and we can see that the context in which an infant is growing is likely to affect the way parents behave. The role of health professionals in this relationship is complex, and there isn't always a good fit between the expectations of each party. Research aimed at understanding influences suggested by this review in relation to health behaviour in general, as well as growth in particular, would help front line professionals to support families.

In this review we set out to assess whether there was sufficient evidence of optimal patterns of infant growth to justify the development of interventions in early life to optimise infant growth as a way of preventing adult disease, and to assess whether the public, particularly parents, would support this. From the evidence amassed,

we were unable to define optimal size or growth. Therefore we do not recommend either the development of a testable intervention to prevent adult disease by optimising infant growth, nor a policy change to achieve this. Instead, we recommend further research to fill the gaps exposed by this review. This research should include assessment of population trends in infant growth and the effects on those trends of the many policy changes aimed at the early years and at preventing childhood obesity.

14. Acknowledgements

15. Source of funding

Reference list

1. Bundred P, Kitchiner D, Buchan I. Prevalence of overweight and obese children between 1989 and 1998: population bases series of cross sectional studies. *BMJ* 2001;**322**:1-4.

2. Stamatakis E. Anthropometric measurements, overweight, and obesity. In Sproston K, Primatesta P, eds. *Health Survey for England 2002: The Health of Children and Young People*, London: The Stationery Office, 2002.

3. House of Commons Health Committee. Obesity - Third Report of Session 2003-04, volume 1. 2004. London, The Stationery Office Limited.

4. Barker DJ, Winter PD, Osmond C, Margetts B, Simmonds SJ. Weight in infancy and death from ischaemic heart disease. *Lancet* 1989;**2**:577-80.

5. Eriksson JG, Forsen T, Tuomilehto J, Osmond C, Barker DJ. Early growth and coronary heart disease in later life: longitudinal study. *BMJ* 2001;**322**:949-53.

6. Barker DJP. Mothers, Babies and Health in Later Life. Edinburgh: Churchill Livingstone, 1998.

7. Singhal A,.Lucas A. Early origins of cardiovascular disease: is there a unifying hypothesis? *Lancet* 2004;**363**:45.

8. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D *et al*. Framework for design and evaluation of complex interventions to improve health. *BMJ* 2000;**321**:413-19.

9. NHS Centre for Reviews and Dissemination. Undertaking systematic reviews of research on effectiveness: CRD's guidance for those carrying out or commissioning reviews. 4 (2nd Edition). 2001. CRD.

10. Murray CJL,.Lopez AD. Regional patterns of disability-free life expectancy and disability-adjusted life expectancy: Global burden of disease study. *Lancet* 1997;**349**:1347-52.

11. Office for National Statistics. The health of children and young people. 2004.

12. Social Determinants of Health. Oxford: Oxford University Press, 1999.

13. Ezzati M, Lopez AD, Rodgers A, Vander Hoorn S, Murray CJL, Comparative Risk Assessment Collaborating Group. Selected major risk factors and global and regional burden of disease. *Lancet* 2002;**360**:60.

14. NHS Centre for Reviews and Dissemination. A systematic review of public water fluoridation. 2000. York, NHS Centre for Reviews and Dissemination.

15. Williams SM, Taylor BJ, Ford RP, Nelson EA. Growth velocity before sudden infant death. *Arch Dis Child* 1990;**65**:1315-8.

16. Brooks JG, Gilbert RE, Flemming PJ, Berry PJ, Golding J. Postnatal growth preceding sudden infant death syndrome. *Pediatrics* 1994;**94**:456-61.

17. Williams SM, Scragg R, Mitchell EA, Taylor BJ. Growth and the sudden infant death syndrome. *Acta Paediatr.* 1996;**85**:1284-9.

18. Peterson DR, Benson EA, Fisher LD, Chinn NA, Beckwith JB. Postnatal growth and the sudden infant death syndrome. *Am J Epidemiol.* 1974;**99**:389-94.

19. Jorgensen T, Biering-Sorensen F, Hilden J. Sudden infant death in Copenhagen 1956--1971. IV. Infant development. *Acta Paediatr.Scand.* 1982;**71**:183-9.

20. Blair PS, Nadin P, Cole TJ, Fleming PJ, Smith IJ, Platt MW *et al*. Weight gain and sudden infant death syndrome: changes in weight z scores may identify infants at increased risk. *Arch Dis Child* 2000;**82**:462-9.

21. Naeye RL, Ladis B, Drage JS. Sudden infant death syndrome. A prospective study. *Am J Dis Child* 1976;**130**:1207-10.

22. Froggatt P, Lynas MA, MacKenzie G. Epidemiology of sudden unexpected death in infants ('cot death') in Northern Ireland. *British Journal of Preventive & Social Medicine.25(3):119-34,* 1971.

23. Suminoe A, Matsuzaki A, Kinukawa N, Inamitsu T, Tajiri T, Suita S *et al*. Rapid somatic growth after birth in children with neuroblastoma: A survey of 1718 patients with childhood cancer in Kyushu-Okinawa district. *J Pediatr.* 1999;**134**:178-84.

24. Victora CG, Barros FC, Horta BL, Martorell R. Short-term benefits of catch-up growth for small-for-gestational-age infants.[see comment]. *International Journal of Epidemiology.30(6):1325-30,* 2001.

25. Yoon PW, Black RE, Moulton LH, Becker S. The effect of malnutrition on the risk of diarrheal and respiratory mortality in children < 2 y of age in Cebu, Philippines.[comment]. *Am J Clin Nutr* 1997;**65**:1070-7.

26. Fonseca W, Kirkwood BR, Victora CG, Fuchs SR, Flores JA, Misago C. Risk factors for childhood pneumonia among the urban poor in Fortaleza, Brazil: a case--control study. *Bulletin of the World Health Organization.74(2):199-208,* 1996.

27. Courchesne E, Carper R, Akshoomoff N. Evidence of brain overgrowth in the first year of life in autism.[see comment]. *JAMA.290(3):337-44,* 2003.

28. Stettler N, Zemel BS, Kumanyika S, Stallings VA. Infant weight gain and childhood overweight status in a multicenter, cohort study. *Pediatrics* 2002;**109**:194-9.

29. Poskitt EM,.Cole TJ. Do fat babies stay fat? *Br Med J* 1977;**1**:7-9.

30. Stettler N, Bovet P, Shamlaye H, Zemel BS, Stallings VA, Paccaud F. Prevalence and risk factors for overweight and obesity in children from Seychelles, a country in rapid transition: the importance of early growth. *International Journal of Obesity*

& Related Metabolic Disorders: Journal of the International Association for the Study of Obesity.26(2):214-9, 2002.

31. Mellbin T,.Vuille JC. Physical development at 7 years of age in relation to velocity of weight gain in infancy with special reference to incidence of overweight. *Br J Prev.Soc Med* 1973;**27**:225-35.

32. Eid EE. Follow-up study of physical growth of children who had excessive weight gain in first six months of life. *British Medical Journal.2(701):74-6,* 1970.

33. Asher P. Fat babies and fat children. The prognosis of obesity in the very young. *Archives of Disease in Childhood.41(220):672-3,* 1966.

34. Mei Z, Grummer-Strawn LM, Scanlon KS. Does overweight in infancy persist through the preschool years? An analysis of CDC Pediatric Nutrition Surveillance System data. *Sozial- und Praventivmedizin.48(3):161-7,* 2003.

35. Johnston FE,.Mack RW. Obesity in urban black adolescents of high and low relative weight at 1 year of age. *American Journal of Diseases of Children.132(9):862-4,* 1978.

36. Wilkinson PW, Parkin JM, Pearlson J, Philips PR, Sykes P. Obesity in childhood: A community study in Newcastle upon Tyne. *Lancet.1(8007):350-2,* 1977.

37. Tenovuo A, Kero P, Korvenranta H, Piekkala P, Sillanpaa M, Erkkola R. Developmental outcome of 519 small-for-gestational age children at the age of two years. *Neuropediatrics.* 1919;41-5.

38. Astbury J, Orgill AA, Bajuk B, Yu VY. Sequelae of growth failure in appropriate for gestational age, very low-birthweight infants. *Dev Med Child Neurol* 1986;**28**:472-9.

39. Koller H, Lawson K, Rose SA, Wallace I, McCarton C. Patterns of cognitive development in very low birth weight children during the first six years of life. *Pediatrics* 1997;**99**:383-9.

40. Ford G, Rickards A, Kitchen W. Relationship of growth and psychoneurologic status of 2-year-old children of birthweight 500-999 g. *Early Human Development* 1986;**13:**329-37.

41. Fisch RO, Bilek MK, Horrobin JM, Chang PN. Children with superior intelligence at 7 years of age: a prospective study of the influence of perinatal, medical, and socioeconomic factors. *Am J Dis Child* 1976;**130**:481-7.

42. Fernandez-Carrocera LA, Chavez-Torres R, Casanueva E, Barrera-Reyes RH, Del P, I, Martinez-Cruz C. Intrauterine growth retardation and neurodevelopment at one year of age in Mexican children. *Nutrition Research* 2003;**23**:1-8.

43. Pryor J. Physical and behavioural correlates of 12-month development in small-for-gestational age and appropriately grown infants. *Journal of Reproductive & Infant Psychology* 1996;**14:**233-42.

44. Camp BW, Jamieson-Darr K, Hansen R, Schmidt B. Growth parameters and attention to faces at 4 to 6 months of age. *J Dev Behav.Pediatr.* 1990;**11**:229-33.

45. Nelson KB,.Deutschberger J. Head size at one year as a predictor of four-year IQ. *Dev Med Child Neurol* 1970;**12**:487-95.

46. Lasky RE, Klein RE, Yarbrough C, Engle PL, Lechtig A, Martorell R. The relationship between physical growth and infant behavioral development in rural Guatemala. *Child Dev* 1981;**52**:219-26.

47. Lucas A, Morley R, Lister G, Leeson-Payne C. Effect of very low birth weight on cognitive abilities at school age. *N Engl J Med* 1992;**326**:202-3.

48. Ernhart CB. Size and cognitive development in the early preschool years. *Psychological Reports* 1961;**61:**103-6.

49. Stathis SL, O'Callaghan M, Harvey J, Rogers Y. Head circumference in ELBW babies is associated with learning difficulties and cognition but not ADHD in the school-aged child. *Dev Med Child Neurol* 1999;**41**:375-80.

50. Hack M,.Breslau N. Very low birth weight infants: effects of brain growth during infancy on intelligence quotient at 3 years of age. *Pediatrics* 1986;**77**:196-202.

51. Ong LC, Boo NY, Chandran V, Zamratol SM, Allison L, Teoh SL *et al*. Relationship between head growth and neurodevelopmental outcome of Malaysian very low birthweight infants during the 1st year of life. *Ann.Trop.Paediatr.* 1997;**17**:209-16.

52. Gale CR, O'Callaghan FJ, Godfrey KM, Law CM, Martyn CN. Critical periods of brain growth and cognitive function in children. *Brain.127(Pt 2):321-9,* 2004.

53. Bendersky M, Koons A, Lewis M, Hegyi T. Developmental implications of head growth following intracranial hemorrhage. *Clin Pediatr.(Phila)* 1998;**37**:469-76.

54. Hack M, Breslau N, Weissman B, Aram D, Klein N, Borawski E. Effect of very low birth weight and subnormal head size on cognitive abilities at school age.[see comment]. *New England Journal of Medicine.325(4):231-7,* 1991.

55. Skuse D, Pickles A, Wolke D, Reilly S. Postnatal growth and mental development: evidence for a "sensitive period". *J Child Psychol.Psychiatry* 1994;**35**:521-45.

56. Boddy J, Skuse D, Andrews B. The developmental sequelae of nonorganic failure to thrive. *J Child Psychol.Psychiatry* 2000;**41**:1003-14.

57. Corbett SS, Drewett RF, Wright CM. Does a fall down a centile chart matter? The growth and developmental sequelae of mild failure to thrive. *Acta Paediatrica* 1996;**85**:1278-83.

58. Drewett RF, Corbett SS, Wright CM. Cognitive and educational attainments at school age of children who failed to thrive in infancy: a population-based study. *J Child Psychol.Psychiatry* 1999;**40**:551-61.

59. Drotar D. Early psychological outcomes in failure to thrive: Predictions from an interactional model. *Journal of Clinical Child Psychology* 1985;**14:**105-11.

60. Drotar D,.Sturm L. Prediction of intellectual development in young children with early histories of nonorganic failure-to-thrive. *J Pediatr.Psychol.* 1988;**13**:281-96.

61. Mackner LM, Starr RHJ, Black MM. The cumulative effect of neglect and failure to thrive on cognitive functioning. *Child Abuse Negl.* 1997;**21**:691-700.

62. Mackner LM, Black MM, Starr RH, Jr. Cognitive development of children in poverty with failure to thrive: a prospective study through age 6. *Journal of Child Psychology & Psychiatry & Allied Disciplines.44(5):743-51,* 2003.

63. Kerr MA, Black MM, Krishnakumar A. Failure-to-thrive, maltreatment and the behavior and development of 6-year-old children from low-income, urban families: a cumulative risk model. *Child Abuse Negl.* 2000;**24**:587-98.

64. Kelleher KJ, Casey PH, Bradley RH, Pope SK, Whiteside L, Barrett KW *et al.* Risk factors and outcomes for failure to thrive in low birth weight preterm infants. *Pediatrics* 1993;**91**:941-8.

65. Singer LT. Cognitive development in the failure-to-thrive infant: A three-year longitudinal study. *Journal of Pediatric Psychology* 1984;**9:**363-83.

66. Glaser HH, Heagarty MC, Bullard DMJ, Pivchik EC. Physical and psychological development of children with early failure to thrive. *J Pediatr.* 1968;**73**:690-8.

67. Reif S, Beler B, Villa Y, Spirer Z. Long-term follow-up and outcome of infants with non-organic failure to thrive. *Isr.J Med Sci* 1995;**31**:483-9.

68. Field M. Follow-up developmental status of infants hospitalized for nonorganic failure to thrive. *J Pediatr.Psychol.* 1984;**9**:241-56.

69. Hack M, Merkatz IR, Gordon D, Jones PK, Fanaroff AA. The prognostic significance of postnatal growth in very low--birth weight infants. *American Journal of Obstetrics & Gynecology.143(6):693-9,* 1982.

70. Mitchell WG, Gorrell RW, Greenberg RA. Failure-to-thrive: a study in a primary care setting. Epidemiology and follow-up. *Pediatrics.65(5):971-7,* 1980.

71. Tudehope DI, Burns Y, O'Callaghan M, Mohay H, Silcock A. The relationship between intrauterine and postnatal growth on the subsequent psychomotor development of very low birthweight (VLBW) infants. *Aust Paediatr.J* 1983;**19**:3-8.

72. Abramson L. Facial expressivity in failure to thrive and normal infants: Implications for their capacity to engage in the world. *Merrill-Palmer Quarterly* 1937;**37:**159-almerQuarterlyyy37.

73. Dykman RA, Casey PH, Ackerman PT, McPherson WB. Behavioral and cognitive status in school-aged children with a history of failure to thrive during early childhood. *Clinical Pediatrics.40(2):63-70,* 2001.

74. Berkman DS, Lescano AG, Gilman RH, Lopez SL, Black MM. Effects of stunting, diarrhoeal disease, and parasitic infection during infancy on cognition in late childhood: a follow-up study. *Lancet* 2002;**359**:564-71.

75. Monckeberg F, Tisler S, Toro S, Gattas V, Vega L. Malnutrition and mental development. *Am J Clin Nutr* 1972;**25**:766-72.

76. Paine PA. Postnatal growth and psychomotor development in small for gestational age Brazilian infants. *Dev Psychol.* 1920;**20**:363-6.

77. Gherpelli JL, Ferreira H, Costa HP. Neurological follow-up of small-for-gestational age newborn infants. A study of risk factors related to prognosis at one year of age. *Arq Neuropsiquiatr.* 1993;**51**:50-8.

78. Ross G, Krauss AN, Auld PA. Growth achievement in low-birth-weight premature infants: relationship to neurobehavioral outcome at one year. *Journal of Pediatrics.103(1):105-8,* 1983.

79. Rose SA. Relation between physical growth and information processing in infants born in India. *Child Development* 1965;**65:**889-902.

80. Corbett, S. S., Durham, M., Wright, C. M., Tymms, P., and Drewett, R. F. Weight gain in infancy and educational attainment at ten. 2004.

81. Whaley SE, Sigman M, Espinosa MP, Neumann CG. Infant predictors of cognitive development in an undernourished Kenyan population. *J Dev Behav.Pediatr.* 1998;**19**:169-77.

82. Latal-Hajnal B, von Siebenthal K, Kovari H, Bucher HU, Largo RH. Postnatal growth in VLBW infants: significant association with neurodevelopmental outcome.[see comment]. *Journal of Pediatrics.143(2):163-70,* 2003.

83. Upadhyay SK, Saran A, Agarwal DK, Singh MP, Agarwal KN. Growth and behavior development in rural infants in relation to malnutrition and environment. *Indian Pediatrics.29(5):595-606,* 1992.

84. Connors JM, O'Callaghan MJ, Burns YR, Gray PH, Tudehope DI, Mohay H *et al*. The influence of growth on development outcome in extremely low birthweight infants at 2 years of age. *J Paediatr.Child Health* 1999;**35**:37-41.

85. Jaffe M,.Kosakov C. The motor development of fat babies. *Clin Pediatr.(Phila)* 1982;**21**:619-21.

86. Kohlhauser C, Fuiko R, Panagl A, Zadra C, Haschke N, Brandstetter S *et al*. Outcome of very-low-birth-weight infants at 1 and 2 years of age. The importance of early identification of neurodevelopmental deficits. *Clin Pediatr.(Phila)* 2000;**39**:441-9.

87. Simon NP, Brady NR, Stafford RL. Catch-up head growth and motor performance in very-low-birthweight infants. *Clin Pediatr.(Phila)* 1993;**32**:405-11.

88. Cheung YB, Yip PS, Karlberg JP. Fetal growth, early postnatal growth and motor development in Pakistani infants.[see comment]. *International Journal of Epidemiology.30(1):66-72,* 2001.

89. Johansson C, Samuelsson U, Ludvigsson J. A high weight gain early in life is associated with an increased risk of type 1 (insulin-dependent) diabetes mellitus. *Diabetologia* 1994;**37**:91-4.

90. Hypponen E, Kenward MG, Virtanen SM, Piitulainen A, Virta-Autio P, Tuomilehto J *et al*. Infant feeding, early weight gain, and risk of type I diabetes. *Diabetes Care* 1999;**22**:1961-5.

91. Hypponen E, Virtanen SM, Kenward MG, Knip M, Akerblom HK, Childhood Diabetes in Finland Study Group. Obesity, increased linear growth, and risk of type 1 diabetes in children. *Diabetes Care.* 2000;**23**:1755-60.

92. Baum JD, Ounsted M, Smith MA. Letter: Weight gain in infancy and subsequent development of diabetes mellitus in childhood. *Lancet.* 1975;**2**:866.

93. DiLiberti JH, Carver K, Parton E, Totka J, Mick G, McCormick K. Stature at time of diagnosis of type 1 diabetes mellitus. *Pediatrics* 2002;**109**:479-83.

94. Ramachandran A, Snehalatha C, Joseph TA, Vijay V, Viswanathan M. Height at onset of insulin-dependent diabetes in children in southern India. *Diabetes Research & Clinical Practice.* 1994;**23**:55-7.

95. Bruining GJ. Association between infant growth before onset of juvenile type-1 diabetes and autoantibodies to IA-2. Netherlands Kolibrie study group of childhood diabetes.[see comment]. *Lancet.356(9230):655-6,* 2000.

96. Drewett, R. F., Corbett, S. S., and Wright, C. M. Physical and emotional development, appetite and body image in adolescents who failed to thrive as infants. 2004.

97. Fall CH, Vijayakumar M, Barker DJ, Osmond C, Duggleby S. Weight in infancy and prevalence of coronary heart disease in adult life. *BMJ* 1995;**310**:17-9.

98. Osmond C, Barker DJ, Winter PD, Fall CH, Simmonds SJ. Early growth and death from cardiovascular disease in women. *BMJ.307(6918):1519-24,* 1993.

99. Forsen T, Osmond C, Eriksson JG, Barker DJ. Growth of girls who later develop coronary heart disease. *Heart (British Cardiac Society).90(1):20-4,* 2004.

100. Martyn CN, Barker DJ, Osmond C. Mothers' pelvic size, fetal growth, and death from stroke and coronary heart disease in men in the UK. *Lancet* 1996;**348**:1264-8.

101. Hales CN, Barker DJP, Clark PMS, Cox LJ, Fall C, Osmond C *et al*. Fetal and infant growth and impaired glucose tolerance at age 64. *BMJ.British Medical Journal* 1991;**303**:1019-22.

102. Eriksson JG, Forsen T, Tuomilehto J, Osmond C, Barker DJ. Early adiposity rebound in childhood and risk of Type 2 diabetes in adult life. *Diabetologia.46(2):190-4,* 2003.

103. Bhargava SK, Sachdev HS, Fall CH, Osmond C, Lakshmy R, Barker DJ *et al*. Relation of serial changes in childhood body-mass index to impaired glucose tolerance in young adulthood.[see comment]. *New England Journal of Medicine.350(9):865-75,* 2004.

104. De Stavola BL, dos SS, I, McCormack V, Hardy RJ, Kuh DJ, Wadsworth ME. Childhood growth and breast cancer. *American Journal of Epidemiology.159(7):671-82,* 2004.

105. Sayer AA, Poole J, Cox V, Kuh D, Hardy R, Wadsworth M *et al*. Weight from birth to 53 years: a longitudinal study of the influence on clinical hand osteoarthritis. *Arthritis & Rheumatism.* 2003;**48**:1030-3.

106. Barker DJ, Osmond C, Rodin I, Fall CH, Winter PD. Low weight gain in infancy and suicide in adult life. *BMJ* 1995;**311**:1203.

107. Thompson C, Syddall H, Rodin I, Osmond C, Barker DJ. Birth weight and the risk of depressive disorder in late life. *Br J Psychiatry* 2001;**179**:450-5.

108. He Q,.Karlberg J. Prediction of adult overweight during the pediatric years. *Pediatr.Res.* 1999;**46**:697-703.

109. Whitaker RC, Wright JA, Pepe MS, Seidel KD, Dietz WH. Predicting obesity in young adulthood from childhood and parental obesity.[comment]. *N Engl J Med* 1997;**337**:869-73.

110. Rolland-Cachera MF, Deheeger M, Guilloud-Bataille M, Avons P, Patois E, Sempe M. Tracking the development of adiposity from one month of age to adulthood. *Ann.Hum.Biol* 1987;**14**:219-29.

111. Guo SS, Roche AF, Chumlea WC, Gardner JD, Siervogel RM. The predictive value of childhood body mass index values for overweight at age 35 y. *Am J Clin Nutr* 1994;**59**:810-9.

112. Eriksson J, Forsen T, Osmond C, Barker D. Obesity from cradle to grave. *International Journal of Obesity & Related Metabolic Disorders: Journal of the International Association for the Study of Obesity.27(6):722-7,* 2003.

113. Mossberg HO. 40-year follow-up of overweight children. *Lancet* 1989;**2**:491-3.

114. Charney E, Goodman HC, McBride M, Lyon B, Pratt R. Childhood antecedents of adult obesity. Do chubby infants become obese adults? *New England Journal of Medicine.295(1):6-9,* 1976.

115. Stettler N, Kumanyika SK, Katz SH, Zemel BS, Stallings VA. Rapid weight gain during infancy and obesity in young adulthood in a cohort of African Americans.[see comment]. *American Journal of Clinical Nutrition.77(6):1374-8,* 2003.

116. Garn SM,.LaVelle M. Two-decade follow-up of fatness in early childhood. *American Journal of Diseases of Children.139(2):181-5,* 1985.

117. Tienboon P,.Wahlqvist ML. A prospective study of weight and height going from infancy to adolescence. *Asia Pac.J Clin Nutr* 2002;**11**:42-7.

118. Monteiro PO, Victora CG, Barros FC, Monteiro LM. Birth size, early childhood growth, and adolescent obesity in a Brazilian birth cohort. *International Journal of Obesity & Related Metabolic Disorders: Journal of the International Association for the Study of Obesity.27(10):1274-82,* 2003.

119. Heald FP,.Hollander RJ. The relationship between obesity in adolescence and early growth. *Journal of Pediatrics.* 1965;**67**:35-8.

120. Barker, D. J. P., Eriksson, J. G., Forsen, T., and Osmond, C. Infant growth and income fifty years later. 2004.

121. Petrosino A, Boruch RF, Rounding C, McDonald S, Chalmers I. The Campbell Collaboration Social, Psychological, Educational and Criminological Trials Register (C2-SPECTR) to Facilitate the Preparation and Maintenance of Systematic Reviews of Social and Educational Interventions. *Evaluation and Research in Education* 2000;**14**:206-19.

122. Harden A, Garcia J, Oliver S, Rees R, Shepherd J, Brunton G *et al*. Applying systematic review methods to sutides of people's views: and example from public health research. *Journal of Epidemiology and Community Health* 2004;**58**:794-800.

123. Oliver, Sandy, Harden, Angela, Rees, Rebecca, Shepherd, Jonathon, Brunton, Ginny, Garcia, Jo, and Oakley, Ann. An emerging framework for interating different types of evidence in systematic reviews for public policy. 2004.

124. Shepherd, J, Garcia, J, Oliver, S, Harden, A, Rees, R, Brunton, G, and Oakley, A. Barriers to, and facilitators of the health of young people: A systematic review of evidence on young people's views and on intervention in mental health, physical activity and health eating. 2002. London. EPPI-Centre, Social Science Research Unit.

125. Office for National Statistics. The UK population: by ethnic group, April 2001. 1-9-2003.

126. The Literacy Trust. Languages spoken by pupils in London. 1-9-2003.

127. Spencer, Liz, Ritchie, Jane, Lewis, Jane, and Dillon, Lucy. Quality in qualitative evaluation: a framework for assessing research evidence. 2003. London,UK, The Strategy Unit, Admiralty Arch.

128. Mays N,.Pope C. Assessing quality in qualitative research. *Br Med J* 2000;**320**:50-2.

129. Dixon-Woods, Mary, Agarwal, Shona, Young, Bridget, Jones, David, and Sutton, Alex. Integrative approaches to qualitative and quantitative evidence. 2004. London, Health Developmental Agency.

130. Arai, Lisa, Popay, Jenny, Roberts, Helen, and Roen, Katrina. Preventing accidents in children-how can we improve our understanding of what really works?: Exploring methodological and practical issues in the systematic review of factors affecting the implementation of child injury prevention initiatives. 2003. London, Child Health Research & Policy Unit, City University.

131. Popay, Jenny, Roen, Katrina, Sowden, Amanda, Rodgers, Mark, Roberts, Helen, Arai, Lisa, Petticrew, Mark, and Baldwin, Sally. Developing methods for the narrative synthesis of quantitative and qualitative data in systematic reviews of effectiveness. 12-10-2004.

132. Popay J, Rogers A, Williams G. Rationale and standards for the systematic review of qualitative literature in health services research. *Qualitative Health Research* 1998;**8**:351.

133. Shaw RL, Sutton AJ, Jones DR, Booth A, Dixon-Woods M, Miller T *et al*. An evaluation of these methods of appraising qualitative research. *Journal of Epidemiology and Community Health* 2004;**58**:A25.

134. Reifsnider E, Allan J, Percy M. Mothers' explanatory models of lack of child growth. *Public Health Nursing* 2000;**17**:US.

135. Hopkins DP, Briss PA, Ricard CJ, Husten CG, Carande-Kulis VG, Fielding JE *et al*. Reviews of evidence regarding interventions to reduce tobacco use and exposure to environmental tobacco smoke. *American Journal of Preventive Medicine* 2001;**20**:16-66.

136. Garcia J, Bricker L, Henderson J, Martin M-A, Mugford M, Nielson J *et al*. Women's views of pregnancy ultrasound: A systematic review. *Birth* 2002;**29**:225-50.

137. Cox, Kate, Stevenson, Fiona, Britten, Nicky, and Dundar, Yenal. A Systematic review of communication between patients and health care professionals about medicine-taking and prescribing. 2003. King's College London, GKT Concordance Unit, Guys' King's and St Thomas' School of Medicine.

138. Rand CSW,.Wright BA. Continuity and change in the evaluation of ideal and acceptable body sizes across a wide age span. *International Journal of Eating Disorders* 2000;**28**:US.

139. Rand CSW,.Wright BA. Thinner females and heavier males: Who says? A comparison of female to male ideal body sizes across a wide age span. *International Journal of Eating Disorders* 2001;**29**:US.

140. Thomlinson EH. The lived experience of families of children who are failing to thrive. *Journal of Advanced Nursing.39(6):537-45, 2002 Sep.(12 ref)* 2002;537-45.

141. May KM. Searching for normalcy: mothers' caregiving for low birth weight infants. *Pediatric Nursing.23(1):17-20, 26-7, 1997 Jan-Feb.(27 ref)* 1997;17-20.

142. Sherratt F, Johnson A, Holmes S. Responding to parental concerns at the six-month stage. *Health Visitor.* 1991;**64**:84-6.

143. Vehvilainen-Julkunen K. The function of home visits in maternal and child welfare as evaluated by service providers and users. *Journal of Advanced Nursing.* 1994;**20**:672-8.

144. Baughcum AE, Burklow KA, Deeks CM, Powers SW, Whitaker RC. Maternal feeding practices and childhood obesity: a focus group study of low-income mothers. *Archives of Pediatrics & Adolescent Medicine.* 1998;**152**:1010-4.

145. Birgenaeu, C. T. Body Image in Infancy. Adult body weight-related biases applied to infants. 1-127. 2001. Univeristy of Massachesetts Boston, USA.

146. Rajan L,.Oakley A. Low birth weight babies: the mother's point of view. *Midwifery.6(2):73-85, 1990 Jun.(5 ref)* 1990;73-85.

147. McCann JB, Stein A, Fairburn CG, Dunger DB. Eating habits and attitudes of mothers of children with non-organic failure to thrive. *Archives of Disease in Childhood.70(3):234-6, 1994 Mar.(18 ref)* 1994;234-6.

148. Sturm LA, Drotar D, Laing K, Zimet GD. Mothers' beliefs about the causes of infant growth deficiency: is there attributional bias? *Journal of Pediatric Psychology.* 1997;**22**:329-44.

149. Brown, M. M. An exploration of parental concerns about preterms and full term infants during the first nince months of life. 1-122. 1981. University of California, Los Angeles, USA.

150. Baughcum AE, Powers SW, Johnson SB, Chamberlin LA, Deeks CM, Jain A *et al*. Maternal feeding practices and beliefs and their relationships to overweight in early childhood. *Journal of Developmental & Behavioral Pediatrics.* 2001;**22**:391-408.

151. Hewat RJ,.Ellis DJ. Similarities and differences between women who breastfeed for short and long duration. *Midwifery.2(1):37-43, 1986 Mar.(20 ref)* 1986;37-43.

152. Hall WA, Shearer K, Mogan J, Berkowitz J. Weighing preterm infants before & after breastfeeding: does it increase maternal confidence and competence? *MCN, American Journal of Maternal Child Nursing.27(6):318-27, 2002 Nov-Dec.(28 ref)* 2002;318-27.

153. Bronfenbrenner U. Development ecology through space and time: A future perspective. In Moen P, Elder GH, Luscher K, eds. *Examining lives in context: Perspectives on the ecology of human development*, pp 619-47. Washington, DC: American Psychological Association, 1995.

154. Dishion, Thomas J. Deviant peer contagion within interventions and programs: An ecologcial framework for understanding influence mechanisms. 22-10-2004.

155. Franks P, Campbell TL, Shields GG. Social relationships and health: The relative roles of family functioning and social support. *Social Science and Medicine* 1992;**34**:779-88.

156. Cochran M, Larner M, Riley D, Gunnarson L, Henderson CR. Extending families. The social networks of parents and their children. Cambridge: Cambridge University Press, 1993.

157. National Criminal Justice Reference service. Preventing crime: what works, what doesn't, what's promising. Chapter 4: Family Based Crime Prevention. Lawrence L Sherman. 27-10-2004.

158. Smith MP. Postnatal concerns of mothers: an update. *Midwifery.* 1989;**5**:182-8.

159. Pridham KF. Information needs and problem solving behavior of parents of infants. *Birth Defects: Original Article Series.* 1984;**20**:125-65.

160. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *Br Med J* 2003;**327**:557-60.

161. Jones K. Mission drift in qualitative research, or moving toward a systematic review of qualitative sutdies, moving back to a more systematic narrative review. *The Qualitative Report* 2004;**9**:95-112.

162. Kramer MS, Barr RG, Leduc DG, Boisjoly C, Pless IB. Maternal psychological determinants of infant obesity. Development and testing of two new instruments. *Journal of Chronic Diseases.* 1983;**36**:329-35.

163. Kitzinger J. Qualitative research: introducing focus groups. *Br Med J* 1995;**311**:299-302.

164. Thomas J, Harden A, Oakley A, Oliver S, Sutcliffe K, Rees R *et al*. Integrating qualitative research with trials in systematic reviews. *Br Med J* 2004;**328**:1010-2.

165. Green J, Waters E, Haikerwal A, O'Meill C, Ramsan S, Booth K *et al*. Social, cultural and environmental influence on child activity and eating in Australian migrant communities. *Child: Care, Health & Development* 2003;**29**:441-8.

166. Waters, E., Goldfeld, S., and Hopkins, S. Indicators for child health, development and wellbeing. 2002. Melbourne, Centre for Community Child Health, Royal Children's Hospital Melbourne.

167. Rigby MJ, Kohler LI, Blair ME, Metchler R. Child health indicators for europe. *European Journal of Public Health* 2003;**13**:38-46.

168. Wadsworth MEJ. The Imprint of Time. Oxford: Clarendon Press, 1991.

169. Power C, Matthews S, Manor O. Inequalities in self rated health in the 1958 birth cohort: lifetime social circumstances or social mobility. *BMJ* 1996;**313**:449-53.

170. Pocock SJ, Collier TJ, Dandreo KJ, De Stavola BL, Goldman MB, Kalish LA *et al*. Issues in the reporting of epidemiological studies: a survey of recent practice. *BMJ* 2004;doi:10.1136/bmj.38250.571088.55 (published 6 October 2004).

171. Hill AB. The environment and disease: association or causation? *J Royal Soc Med* 1965;**58**:295-300.