Intelligence and reading abilities in eight year old children who failed to thrive in infancy

Sally Suzanne Corbett

Intelligence and Reading Abilities in Eight Year Old Children who Failed to Thrive in Infancy

Sally Suzanne Corbett

Thesis submitted to the University of Durham
Department of Psychology
for the degree of Doctor of Philosophy
September, 1998

Declaration

Except where otherwise stated, the research contained in this thesis was carried out by the author between January 1995 and September 1998 while a post graduate student in the Department of Psychology at the University of Durham. None of the work contained in this thesis has been submitted in candidature for any other degree.

Statement of Copyright

Intelligence and Reading Abilities in Eight Year Old Children who Failed to Thrive in Infancy

Submitted for the degree of Doctor of Philosophy by Sally Corbett

The aim of the work reported in this thesis was to investigate intelligence and reading ability at school age in a population based sample of children who failed to thrive in the first two years of life.

Weights for an annual cohort of term infants, retrieved from clinic records, were compared with weight expected conditional upon early weight. Those with weights in the lowest 5% in two or more age bands (3, 6, 9, 12 and 18 months) were identified as cases (n =136). Cases were stratified by age, sex and deprivation level of their area of residence at eighteen months of age, and the same number of controls selected from each stratum. Two controls were later omitted as they were found to have been born preterm (< 37 weeks). Between ages 7 and 9 years 79% of cases and 82% of controls were traced and studied. Height, head circumference and weight were measured, and an IQ and reading test administered. Information about socio-economic status, family structure and medical history was gathered during a home visit. The mother's height was measured, the father's being reported by the mothers, and the mother's IQ tested. All testing was carried out blind to the child's case or control status. The child's medical records were retrieved where admission to a hospital or outpatient clinic was reported and the conditions diagnosed were coded blind for their probable effects on cognitive outcomes or growth.

At eight years of age mothers in the case group reported more feeding problems in infancy and more organic conditions. Cases were shorter, lighter, thinner and had a smaller head circumference than controls. These anthropometric differences were all statistically significant and remained so after allowing for parental stature. There were no statistically significant differences in IQ and reading ability either before or after adjusting for maternal IQ, organic condition or the few covariates found to differ between the groups.

Dedication

To my mother who was dedicated to the education and welfare of children, especially those in her own family.

Acknowledgements

I would like to thank the Wellcome Trust, who made this project possible with their generous funding.

I am particularly indebted to Robert Drewett for his support and help during the commission of this research, and formative work which provided the foundation for the current project. I also owe an enormous debt of gratitude to Charlotte Wright whose enthusiasm and drive helped to maintain the momentum of the project.

Thanks are also due to Jane Callum and Eve Brown for all the ground work they did on my behalf, and their good humour throughout. Also thanks to Kathryn Parkinson for carrying out reading tests on my behalf, with such efficiency and diligence.

I am grateful to Head Teachers for allowing me to see the children in their schools and for the support they gave me with the parents and children.

Finally, I am most grateful to the parents who allowed me into their homes and above all to their children for their eager participation.

Table of Contents

Chapter Two

Aims and method

Chapter Three

Demographic characteristics, attrition and medical history of the study sample

Chapter Six

Discussion

List of tables

Chapter 5

List of Figures

Chapter 3

Chapter 4

Chapter 5

Appendix

Chapter One

Introduction and literature review

1.1.1 Introduction

"Failure to thrive" is the term used in respect of 'the infant or young child whose growth falls substantially behind that of his or her peers' (Wilcox et al, 1989). It is a common problem in infancy, prevalence being estimated between 1.3% (Dowdney et al 1987) and 20.9% (Edwards et al, 1990). The detection of poor weight gain in infancy requires growth monitoring (Hall, 1996), and in studies in the United States was found to account for between one and five per cent of paediatric admissions to hospital (Sills, 1978, Berwick, 1980). More recently, multidisciplinary home based intervention programmes have been advocated (Black, 1995, Wright, 1996). The detection and treatment of poor weight gain thus requires considerable resources.

In addition to the implications for public health, concerns have been raised about the cognitive development of children with failure to thrive and the increased likelihood that they will experience greater learning difficulties at school (Elmer et al, 1968, Hufton and Oates, 1977, Oates et al, 1985, Dowdney et al, 1987). However, some recent studies have cast doubt on these findings (Mitchell et al, 1980, Corbett, 1994, Boddy, 1997) having found few differences between cases and their controls. The aim of the research presented in this thesis is to ascertain if there is an association between failure to thrive in infancy and enduring psychological and educational deficits.

1.1.2 Problems with the definition of failure to thrive

A major problem with previous studies has been that failure to thrive has proved difficult to define, both in terms of its aetiology, which is not well understood (Boddy and Skuse, 1994), or its presenting feature, poor growth. Hall (1996) has questioned the value of repeated growth monitoring for all children as there is 'no simple way of

assessing whether growth limits are outside the range of normality'. In cases of failure to thrive he recommends that, 'the whole clinical picture rather than the weight chart alone' should be taken into account 'mainly because psychosocial factors are sometimes associated with poor growth'. In some cases slow growth in infancy is attributed to an underlying organic condition and is referred to as 'organic failure to thrive' (OFT). Where no such attribution has been made, it is referred to as non-organic failure to thrive (NOFT). This distinction will be discussed later in this review.

Whilst it may be appropriate to the management of an individual infant's poor growth to evaluate the whole clinical picture, it is not appropriate for research, as it is unclear which criteria other than poor growth are sensitive or specific indicators of failure to thrive in a population. Kotelchuck and Newberger (1983) were highly critical of previous work which has focussed on the inadequacies of the mother as an aetiological factor in failure to thrive. In their study of 42 mothers of infants failing to thrive and a matched control group, they found that although the mothers were more socially isolated and had lower levels of educational attainment, the case children were also more sickly. Their evidence could not support the view that failure to thrive is the result of inadequate mothering.

Skuse (1985) argued that the primary cause of failure of thrive is poor nutritional intake which may be associated with poor parenting, but also states that 'no universal attribute of caretakers has been found'. He proposes a hierarchy of factors which predispose to poor nutritional intake and thus to failure to thrive, such as the quality of parenting, the interaction between the caregiver and infant, marital problems and poverty. However, in the study of Kotelchuck and Newberger (1983) the problems mothers encountered with their child were very specific and confined to feeding and then health. No demographic or social class differences were found between their cases and the comparison group. As far as nutritional intake is concerned, many studies report feeding problems in children who fail to thrive (Pollitt and Eichler, 1976, Mathisen et al, 1989, Wilensky et al, 1996), but in a study of four year olds with chronic failure to thrive whose food intake was observed and calculated, the mean daily energy intake was not significantly different from a control group and in fact was higher in the case group when calculated per unit body weight (Hepinstall et al, 1987).

In order to ascertain how failure to thrive was defined, Wilcox et al (1989) reviewed 22 current paediatric textbooks and 13 papers published in journals. Of the textbooks, four made no mention of failure to thrive, two gave no definition at all and six used subjective definitions which did not include any anthropometric or other criteria for abnormality. Failure to thrive was described in quantitative terms in ten of the 22 textbooks, but there was a lack of consensus in the anthropometric indices used and the criteria for abnormality. Similarly, of the thirteen journal articles reviewed, four contained no definition of failure to thrive and the remaining nine used various anthropometric criteria. However, different definitions of poor growth may identify different populations of children failing to thrive, with resulting inconsistencies when comparing outcomes.

Drotar (1990) attempts to clarify a procedure for defining failure to thrive. Firstly a child's growth pattern should be compared with age-related norms for physical growth. A cut off point should be defined, such as weight more than two standard deviations below the reference population median, or a less stringent criterion such as weight below the 5th centile (-1.64 standard deviations). Sometimes a decline in relative standing against population norms is used (Mitchell et al, 1980, Edwards et al, 1990). Drotar also advocates the use of additional measures, such as length, or height, or head circumference which is associated with poor brain growth (Winick and Rosso, 1969). These additional measures may provide more information about the characteristics of poor growth which are associated with adverse psychological outcomes.

Length in infancy is almost impossible to measure single handedly and difficult to measure accurately, and there is a discontinuity between supine length and standing height (Tanner, 1989). Weight on the other hand can be measured by a single person. Waterlow et al (1977) point out that weight for age is a particularly useful measure in infants under one year old and, if length is not measured accurately, may be the most valid measure available. Weight is the index most often used to monitor growth in infancy at the primary care level (WHO, 1986). For these reasons, the remainder of this discussion will be mainly concerned with assessing attained weight or rate of weight gain.

Three issues need to be considered. Firstly, caution should be exercised in the choice of a reference population. Secondly, the choice of cut off point is conventional and has not been shown to have a clear theoretical basis. Thirdly, using a single measure, for example an attained weight, biases the definition towards the inclusion of low birth weight children, so serial weights are more often used to assess weight gain over time. However, measures of decline in relative position on a cross sectional growth chart are difficult to define and interpret.

1.1.3 Choice of reference population

Age related norms are available for two broad categories of growth measure, sometimes called distance and velocity measures. Distance measures are most often used to monitor growth in the primary care setting, velocity measures are used more to determine response to treatment (Tanner, 1989).

A distance measure can be defined as the height or weight attained by a certain age, analogous to a child having travelled a certain distance towards adult size (Tanner et al, 1966). Assessment of normality is made by comparing an individual's height or weight with cross sectional data from a reference population. Growth charts have centile lines marked on them to indicate the percentage of individuals in the reference population above or below the line. Thus a child with a weight plotted below the third centile line is amongst the lightest three per cent of children of that age. Population norms exist for height, weight, head circumference, body mass index (i.e. weight adjusted for height) and many other less commonly used anthropometric measures.

Normative standards should ideally be available for genetically different or geographically localised populations and these should be updated periodically (Tanner 1989) in order to take account of secular trends. The revision of the old British standards of Tanner et al (1966a, 1966b) was long overdue as they were based on anthropometric measures of infants growing up in the fifties who were mainly bottle fed (Whitehead and Paul, 1984). The standards have also been criticised for using a

small and unrepresentative sample of British infants living in South East England (Freeman et al, 1995). Today weight gain is more rapid in early infancy, and slower in later infancy compared to these older standards (Wright et al, 1993).

The World Health Organisation (WHO, 1981) has advocated the use of the National Center for Health Statistics standards (Hammill et al, 1979). These are based on a more recent sampling (Health Examination survey, 1962-1970 and Health and Nutrition Survey, 1971-74) of the whole child population of the USA, the numbers sampled being proportionate to the population of each geographical area. The large reference sample facilitates the comparison of children occupying the outer centiles.

Criticisms of these standards question their appropriateness for breast fed infants and for a UK population. Large differences between breast fed and formula fed babies were found using the NCHS standards (Dewey et al, 1992), the weight of breast-fed boys dropping below the NCHS median at 8 months and that of girls dropping below the 25th centile by 12 months. Formula fed infants remained above the median. Dewey and her colleagues express the concern that the slower weight gain of breast fed infants after 3 months of age in comparison with the standards may be incorrectly perceived as growth faltering, and advocate the development of new growth charts for breast fed infants. Although concurring with the use of the NCHS standards to monitor undernutrition in sub-populations within countries, Tanner (1989) argues that their use in Britain to monitor individuals is inappropriate because height, weight, and weight for height tend to be greater in children in the USA compared with British children. Wright et al (1993) found discrepancies in weight gain in UK children when compared to NCHS standards similar to those found with the Tanner weight standards in infancy.

Conveniently for this study, new British weight reference centiles for children from birth to 20 years of age were published in 1995 (Freeman et al, 1995). The sample for the new British standards was intended to be more representative of the population as a whole. The new weight reference centiles for infants were constructed using data from three sources, the Human Measurements Anthropometry and Growth (HUMAG) Research Group, which sampled 16 areas in England and Wales, the Cambridge Infant

Growth Study of 252 infants living in the Cambridge area, and birth anthropometry from 756 babies from 32 weeks gestational age born at the Whittington Hospital, London, during a six month period.

The HUMAG study, although providing the most representative sample geographically, actually contained data for only a small number of infants and no information on gestational age was recorded, so it was not possible to correct for this. The data for the Cambridge study, although geographically less representative, have the advantage that they provide longitudinal weight data, collected on infants from four weeks after birth to two years. No weights between birth and four weeks were available in this set, but socioeconomic data was recorded, allowing adjustments to be made for social class. The Whittington study provided data on birth weight and weights at two days old. The last two data sets also recorded gestational age.

The British weight reference curves 1990 (Freeman et al, 1995) were produced using the LMS method (Cole, 1990). The data were transformed to produce a normal distribution using a power term (L) to correct for any skew, allowing centiles at any age to be generated by the median (M) and standard deviation (S). The fitting process ensures that the values of L, M and S change smoothly with age and are represented as smooth curves. The LMS method has a number of advantages. Allowing for the effect of skew in the distribution enables data to be converted to standard deviation scores making comparisons easier across ages (the variance differs with age), and providing better estimates of extreme centiles than a non-parametric approach.

1.1.4 Theoretical basis for a growth threshold

In a review of current chart centiles, Cole (1994) argues that there is no theoretical basis for the threshold which defines abnormality, instead, the convention of using the 3rd centile in Europe is the result of using charts with the preferred centile divisions of 3rd, 10th, 25th, 50th, 75th, 90th and 97th centiles marked on them. These centiles are equally spaced in terms of standard deviation scores, each major centile division corresponding to approximately 0.65 of a standard deviation. In the USA the National

Center for Health Statistics uses centile charts with the 5th to 95th centiles marked on them as the compilers of these standards thought the 3rd centile too extreme (Cole, 1994).

In previous studies of failure to thrive a number of different centiles have been used to define a child as a case. Wilcox et al (1989) in their review of criteria used for failure to thrive, found that in eight out of nine textbooks in which the criteria for abnormal growth was stated, indices below the third centile were used. However, this cut off point is arbitrary and the fifth (Drotar et al, 1985) and tenth centile (Dowdney et al, 1987) have been used as well. Dowdney et al (1987) argue that the use of a lower threshold for failure to thrive increases the likelihood that poor growth will have an organic aetiology. Kristansson et al (1981), however, found that failure to thrive with an organic origin was more likely to be detected by evaluating the rate of weight gain rather than low weight for age.

Cole (1994) recommends two thresholds. Children falling below the much lower 0.4th centile should be immediately referred to hospital or specialist clinic, whilst those below the 2nd centile should be monitored for referral. This, he argues, would reduce an unacceptably high false positive rate. However, such a stringent criterion would increase the number of constitutionally small children diagnosed as failing to thrive, whilst reducing the number of initially large infants diagnosed. In any case, it is unclear at which threshold an effect on cognitive outcomes is likely to be found, or whether it is linearly related to a continuous scale of growth, though, as Drotar (1990) points out, the less stringent the cut off criterion the more likely that psychological deficits will be mild.

1.1.5 Changes in relative weight in infancy

Growth in early infancy is difficult to characterise accurately because it is both rapid and has a rapidly diminishing velocity, the rate of weight gain may show short term changes and diurnal variation, there are problems with measurement error or

inaccurate recording of data, and there maybe substantial individual differences in rate of weight gain.

Immediately after birth an infant's growth velocity (i.e. the increments in weight from one age to the next) is high (Tanner, 1989). But this phase of rapid growth also rapidly declines, so that on average, from birth to four months an infant's weight will double, whereas between four months and a year it increases by only one half (Tanner et al, 1966b).

The rapid rate of growth and the fall in growth velocity in the first year of life are further complicated by fluctuations of growth over short periods. Giani et al (1996) observed weight gain of infants from birth and found that fluctuations in weight gain, which are greater in males, occur as a series of decreasing oscillations to the age of two. As growth charts graph cross sectional centiles at different ages, they do not provide the limits of variation that can be expected for individual fluctuations in weight gain.

A further problem with both the standards of Tanner et al (1966a, 1966b) and the NCHS standards is that data for infant weights was only gathered at three month intervals (Whitehead and Paul, 1984), so neither standard provides precise reference data for this period of rapid and fluctuating growth in the first year of life. Gairdner Pearson growth charts (1985) provide standards for infants aged 0 to 24 months and were the standards used in child health clinics in Newcastle Upon Tyne in the late eighties (Edwards et al, 1990). They are based on the standards of Tanner et al (1966a, 1966b) and so have the same limitations. Tanner acknowledges the limitations of using these standards to interpret 'the finer details of the growth of very young babies' (Whitehead and Paul, 1984). Nevertheless, he argues that a change in rank order of a child's growth over time does not reflect the inconsistencies found in patterns of infant growth, but identifies a consistent change, especially when the effect of illness or environmental adversity is being investigated (Tanner, 1989).

1.1.6 Measurement error

Hall (1996) points out that repeat measurements made on the same child are rarely the same, because in addition to diurnal variation in the child, there may be measuring error due to poor maintenance and installation of equipment and non-standardised measuring technique. The problem of measurement error has also been investigated by Davies and Williams (1983) who found that in health clinics, infants were often weighed with clothes on and the weight of the clothes was not always deducted, poorly calibrated scales were used, weights were recorded using imperial and decimal indices interchangeably and mistakes were often made when writing weights and dates on records. When individual infants are assessed, measurement error may lead to their incorrect identification as a case. This has implications for research where routinely collected data is used, necessitating the checking of outlying data and the building in of precautionary measures in any study design, such as avoiding over reliance on single data points.

However, routinely collected population data for infant weights has been found to compare well with weight data that has been collected with greater accuracy. Cole (1995) compared early and late weight standard deviation scores in routinely collected weight data for a cohort of children born in Newcastle and data collected for the standardisation sample in Cambridge. He showed that both the mean of standard deviation scores and the mean of standard deviation scores for weight gain between any two ages was close to zero in both datasets and the distributions were similar, the Newcastle children gaining weight slightly faster in comparison with the standardisation sample in the first year. Cole argues that the small differences between the two datasets confirm that the reference for weight gain based on the standardisation sample can be used for other data.

1.1.7 Individual variation

Assessment of an individual's growth is rarely made on the basis of a single weight alone, as a low weight for age may not be the effect of failure to thrive, but of slow

intrauterine growth. In order to allow for the effect of slow intrauterine growth, changes in the relative position of weights at different ages are evaluated. Thus, if a child were found to be on the 3rd centile for weight, it would be important to know whether this was because weight gain was slow or simply reflects a low birth weight.

Kristiansson, et al (1981) compared two methods of screening infants using weight gain criteria, with screening using an attained weight. One method used the difference between a late standard deviation score and an earlier score, divided by the length of the time period. The other method involved converting the weight gain in kilograms between the two time points to a standard deviation score. Levels of ascertainment of organic disease were more accurate for both of these methods in comparison with screening using a single measure of attained weight. Only 12 children out of 28 with organic disease had a weight equivalent of more than two standard deviations below the mean, whereas 22 of these children had a low rate of weight gain. Kristiansson et al conclude that the rate of weight gain is more sensitive than attained weight.

Two other ways of assessing weight gain are either to compare weight velocity with a standardisation sample, or to monitor changes in centile position on distance charts. Velocity measures for height and weight are rarely used in primary care clinics as they are difficult to interpret. However, Tanner (1989) supports the use of a velocity measure in clinical work as 'velocity picks out the pathological cases better than does distance because velocity represents what is happening now, whereas distance represents the sum of all that has happened in the past'. Peterson et al (1985) particularly recommend monitoring weight velocity, as well as attained weight, in order to differentiate stunted children with a low rate of weight gain from short normal proportionate children.

The recommended practice (Davies and Williams, 1983) in community health clinics is to plot serial weights against age on a distance chart with major centile lines (i.e. the 97th, 90th, 75th, 50th, 25th, 10th, 3rd) marked on them. These weights are used to plot an infants' weight gain so that an easy comparison can be made with their own earlier centile position. A fall across major centile lines often occurs in infant growth as a result of short term illness, but Edwards et al (1990) found that a fall across two

or more major centile lines, from weight attained at four to eight weeks of age, persisting for a month or more, predicts significant anthropometric differences in the second year of life.

However, the definition used by Edwards et al (1990) and the method of comparing serial weights has been shown to be problematic in identifying infants failing to thrive. Previous estimates of the prevalence of failure to thrive at the community level range from 1.3% (Dowdney et al, 1987) to 9.6% (Mitchell et al, 1980). The definition of Edwards et al (1990) produced a prevalence of failure to thrive of 20.9% in a very deprived area of Newcastle upon Tyne and when applied to all infants born in Newcastle in one year, including those born in more affluent areas of the city, this definition identified 33% of infants as failing to thrive (Wright, 1995). In a subsequent analysis of the weight data of the infants studied by Edwards et al the criterion was found to be highly sensitive but with low specificity, identifying as cases large infants showing a fall in relative weight that was within normal limits (Corbett, Drewett and Wright, 1996).

The problem with using serial weights to assess the rate of weight gain of an infant is that individual infants do not maintain their rank order for attained weight over time and considerable variability in relative weight has been observed. Davies (1980) found that only 12% of the infants he studied maintained rank order for attained weight in the first year of life. Tanner (1989) argues that since growth standards are based on cross sectional analysis of weight for age, no confidence limits are provided for shifts in relative weight, which occur longitudinally.

Berkey et al (1983) considered longitudinal data from 309 children from birth to six years, and found that a shift across two major intercentile spaces on the NCHS standards (i.e >97, 90-97, 75-90, 50-75, 25-50, 10-25, 3-10, <3) toward the 50th centile is more likely than a shift away from the 50th centile. The most labile measure is weight before 1.5 to 2 years of age, after which changes in centile position were unusual. They also found greater variability in the weight of males. Reference data are provided which depend on the two time points being considered, and the correlation of weights at different ages which reflects the likely degree of change in centile position

over a certain time. They estimated that only 33% of infants between the 50th and 75th centiles for weight at three months would retain their centile position by one year and that about 3% of these infants would fall below the 10th centile by one year old. Berkey et al warn that the crossing of weight distance centiles should not be used to derive a measure of normal or abnormal weight gain over time, as to do so assumes that normally individuals track along centile lines. In fact, they show that considerable changes in centile position can occur in the first year of life.

The essence of the problem Berkey et al (1983) identified is that simply plotting a fall in centile position, or subtracting an earlier standard deviation score from a later score, makes no allowance for a child's earlier weight. The relative weight of large or small children tends to move towards the mean. This is known as regression to the mean, a well recognised phenomenon of repeated measurements. Since most values cluster around the mean, after an initial extreme value, a later value is more likely to be closer to the mean and less likely to be further from it. Large babies will tend to fall towards and smaller babies catch up towards the population mean.

This was precisely the effect that was demonstrated in the reanalysis of the weight data of the cases originally identified by Edwards (1988) using the criterion of a fall across centile lines. Six children, whose initial weight was well above average, had been identified as cases and yet, despite a fall across two major centiles, their actual later weight was still above the average weight expected for children with the same early weight (Corbett, Drewett and Wright, 1996). Potentially cases may also have been missed as the weight of initially small infants would be expected to be closer to the population mean. Thus, even a small fall in relative weight in an initially low weight infant would mean that they are well below the weight expected in infants with a similar early weight.

1.1.8 A velocity based definition of failure to thrive

Accurate definition of failure to thrive must involve two elements. It must be a measure of post natal weight gain only and take account of an individual infant's initial

weight. A way of incorporating these elements into case ascertainment for failure to thrive has been provided by the Thrive Index method (Wright et al, 1994a). An expected weight, calculated conditional upon a previous weight, allows for regression to the mean. Expected weight for age is then compared with actual weight. This way of identifying cases was first suggested by Healy (1978) and developed by Wright et al (1994a) and Cole (1995).

After correction for skew, weight standard deviation scores (SDS) have a mean of 0 and a standard deviation of 1, so to calculate an expected later weight the regression takes the form:

$$\text{Expected } SDS_{\text{time 2}} = r * SDS_{\text{time1}}$$

Because the scores are standardised the regression coefficient here, r, is the correlation between the weights of infants in the population at time 2 and time 1. The difference between an infant's actual SDS_{time2} and their expected SDS_{time2} can then be used as an index of failure to thrive. It is this difference that is referred to as the Thrive Index (Wright et al, 1994a), as shown below.

$$\text{Thrive Index} = \text{Actual weight SDS} - \text{Expected weight SDS}$$

The Thrive Index is normally distributed, with zero mean and a standard deviation equal to the residual standard deviation from the above regression, $\sqrt{(1 - r^2)}$. Where values for the Thrive Index are less than zero a child has not reached their expected weight for age. A threshold can then be applied to determine the cut off which constitutes failure to thrive. It can be predicted that 5% of children will have a Thrive Index value of $-1.64 * \sqrt{(1 - r^2)}$, and this has been demonstrated empirically to be so (Wright, 1994a), using a large standardisation population.

The Thrive Index thus provides a longitudinal reference, as relative weights of individual infants are compared with the mean relative weight for infants with the same earlier weight.

A drawback of the method is the reliance on two datapoints, an early and late weight, both of which are subject to measurement error. It is also not clear which baseline weight to use to calculate expected later weights. Edwards et al (1990) recommend four to eight weeks weight as a baseline value, as early weights are influenced by intrauterine factors; small infants, in particular, often showing fast catch up growth in the early post natal period. They argue that after six weeks the effect of the intrauterine environment is reduced and that the genetic contribution is greater from the ages of four to eight weeks. However, the use of a later baseline prevents the screening in of cases with early onset failure to thrive.

1.1.9 Summary of the definition of failure to thrive

The problem with conducting research on failure to thrive is that no universally accepted criteria for diagnosis have been agreed. This has led to the use of many different criteria, sometimes not even addressing the main feature of the condition, which is relatively poor post natal weight gain. The effect of this diversity of diagnostic criteria has meant that populations or samples identified and studied may not be comparable, resulting in inconsistent findings.

Abnormal growth in infancy has proved difficult to characterise. A definition must exclude the likelihood that low attained weight might be attributable to poor prenatal growth. For this reason a longitudinal measure of post natal weight gain alone is necessary. However, it is not appropriate simply to plot serial weights on a growth chart in order to detect a fall in relative weight, because, whilst in general early weights are predictive of later weights, individual differences in rate of weight gain lead to considerable variation in an individual's relative weight over time. These individual differences contribute to regression to the mean; small infants catch up to the population mean and large infants fall towards the mean. The Thrive Index provides a method of comparing successive measurements whilst allowing for early weight and so, provides a measure of the extent to which weight gain has deviated from its expected pattern compared to a standard population. A problem with the Thrive Index method is the reliance on only two datapoints, an early and late weight, when

measurement error is common. The method is also likely to miss cases with a very early fall in relative weight, if, as recommended, a six week weight is used as a baseline measure. However, its strength is that since it allows for early weight, the Thrive Index method provides a measure of post natal weight gain only.

1.2.1 Cognitive outcomes of failure to thrive

Any review of previous studies of cognitive outcomes of failure to thrive is not only complicated by the variety of criteria used to define cases, but also the variety of the types of research studies. The earliest are investigations of institutionalised children (Bowlby, 1953, Goldfarb, 1943), but these are specialised populations which do not have relevance to the cases who are cared for at home within the family and whose rate of growth, far from deteriorating when admitted to hospital, as described by Spitz (1946), is often found to accelerate (Casey and Arnold, 1985). More recently the long term effects of undernutrition on growth and cognitive development have been studied in developing countries (e.g. Stoch and Smythe, 1963, Cravioto et al, 1966. Hertzig et al 1972) and there have been studies of the effects of nutritional supplementation (e.g. Freeman et al, 1977, Lasky et al, 1981, Grantham-McGregor et al, 1991, Pollitt et al, 1993). But as Pollitt (1969) points out these are difficult enough studies to evaluate in their own context, and they may have less relevance to countries where sanitation is better and infant health care is more comprehensive. The child abuse literature also contains follow up studies of growth retarded children (Powell et al, 1967, Money et al, 1983). But whilst there may be some overlap of children who have failure to thrive with those who are abused, estimates of the prevalence of failure to thrive in the primary care setting are far greater than for abuse. For failure to thrive estimates of prevalence range from 1.3% (Dowdney et al, 1987) to 20.9% (Edwards et al, 1990, 1994), but the estimated incidence of non-accidental injury as a result of abuse between one and five years is 0.33% (Butler and Golding, 1986), and it is unlikely that all of the abused children would have been diagnosed as failing to thrive. And differences are found in behavioural and cognitive development between cases referred to hospital for failure to thrive and cases referred for abuse (Oates et al, 1984). Follow up studies of short normal children reveal little effect on cognitive performance of short stature once

social class and familial factors are controlled (Lacey and Parkin, 1974, Skuse et al, 1994b). Downie et al (1997) found an average difference of five IQ points between short children and controls, but it is not possible to infer that these children were failing to thrive in infancy as well, because insufficient infant weight data are reported. This review will therefore focus only on infants diagnosed specifically as failing to thrive in countries where undernutrition is not endemic.

In this review, I have divided a number of key of studies into those based on cases referred to hospital and those based on an unselected population. These two categories effectively also divide the review into earlier studies, which relied heavily on samples referred to hospital or specialist clinics, and later studies which were based on whole populations. The reasons for this methodological change will be discussed as part of this review.

Key features of the studies discussed, such as the origin of the sample, the criteria for selection, the number of cases successfully studied as a proportion of those initially screened in, whether controls were recruited, age at follow up, tests used, whether the study was carried out blind, the outcome and additional observations, have been summarised in Appendix 1.

1.2.2 Studies based on infants referred to hospital or clinics for failure to thrive

Three early studies of children hospitalised for failure to thrive during the first year of life (Glaser et al 1968, Elmer et al, 1969, Hufton and Oates, 1977) all report a constellation of problems in behavioural and cognitive development.

Glaser et al (1968) followed up 40 children after an average of four and a half years from discharge. One child was psychotic, three had a non-psychotic emotional disorder, seven had mild behaviour disorders and six had been removed into care. They report that IQ scores although normally distributed were 5 points below that of a standardisation population (mean IQ approximately 95) and that six children (15%) had borderline or retarded intelligence (IQ 75 or below). However, 65% of the

families were in social classes IV or V and despite the evidence of deprivation in the sample, 29 children (72%) showed no evidence of behaviour problems, and 26 children (65%) had IQ scores within the normal range with a further six (15%) scoring above 120 IQ points.

Elmer et al (1969) found that at an average of five years after hospitalisation ten out of fifteen of their sample showed evidence of mental retardation (using the Oppenheimer rating for mental retardation) and seven had abnormal behaviour. Clearly these results demonstrate levels of mental retardation and behaviour problems which are higher than would be expected in a standard population, but again there was a high level of family dysfunction and economic deprivation. Seven out of 13 sets of parents had separated, a number of fathers had disappeared or were in prison or psychiatric hospital, only one mother had finished more than 11th grade at school and all but two families were receiving public assistance.

The third study (Hufton and Oates, 1977) followed up children six years after discharge. Only 14 of the original 30 children were tested using the WISC. It is reported that one child had a full scale score of more than 110 points and three had scores less than 90 points, with the remaining 10 children scoring between 90 and 110 points. Most of those tested had scores within one standard deviation of the mean of a standardisation population (ie. 85 - 115 IQ points). Hufton and Oates report high levels of parental separation, mental illness, families on welfare or with financial problems, and housing difficulties with frequent moves and overcrowding. In view of the reported social problems in the children's families, it is not surprising that slightly fewer children performed well.

The WISC manual (Wechsler, 1992) indicates that in a standard population, 5% of those tested could be expected to have a discrepancy greater than 24 points between the verbal and performance subscales. Large discrepancies between the verbal and performance subscales were reported by Hufton and Oates (1977). These suggested that the children were particularly prone to verbal deficits. Of ten children scoring within 10 points of a standardisation population mean of 100, two (14%) had verbal scores 30 points lower than their performance scores. Only one discrepant score

where verbal performance quotient exceeded performance quotient is reported (19 points higher), but the raw data are not published to show whether the direction of this discrepancy is isolated. Hufton and Oates also report that of 18 cases given a graded reading vocabulary test, two thirds of the children had a reading age one or two years behind their chronological age.

Although these papers are frequently cited, none provides a comparison group with a similar socio-economic background. The assumption is that deficits are attributable to failure to thrive, despite the majority of children's IQ scores being within the average range of a standardisation population, and despite the presence of a number of socio-economic factors which are associated with poorer test and educational performance. A further problem is that no testing was carried out blind, so test results and teacher's reports may have been influenced by the knowledge that all these children had been hospitalised for failure to thrive.

Another problem is that none of these studies used a measure of post natal weight gain alone, the diagnostic criterion being an attained weight in all studies. Attained weight is the sum of all weight gain since conception, so a low birth weight infant may remain below the 3rd centile for some time despite gaining weight at an expected rate post natally. Low birth weight was partially controlled for in the study by Glaser et al (1968), who excluded infants with birthweight of less than five pounds. Low birth weight is a risk factor for adverse cognitive and educational outcomes (Hill et al, 1984, Abel Smith and Knight-Jones, 1990, Mutch et al, 1993) and so poor cognitive outcomes may be explained in terms of pre-natal rather than post natal growth in those children.

Studies of cases referred to hospital, but with improved methodology using controls, or repeated measures, or a comparison of interventions, have also found evidence of developmental delay in cases when testing was carried out blind. However, there is disagreement about the causes of the delay and not all these studies adjusted for the effects on outcome of pre-natal growth.

In a controlled retrospective study of 19 infants diagnosed as failing to thrive three to four years earlier, significant associations were found between admission to hospital for failure to thrive after four months of age and developmental deficits at follow up (Chase and Martin, 1970). More cases remained below the third centile for height, weight and head circumference at follow up if they were admitted after four months. There was no significant difference in developmental quotients between infants who were admitted to hospital with faltering growth before four months and controls. Chase and Martin point out that poor growth was present from birth in all cases, and as those admitted after four months were more likely to be below the 3rd centile for weight and height they also had the most chronic slow weight gain.

Using a repeated measures design to investigate the relationship between weight gain and mental development, Field (1984) studied 17 infants below the 5th centile for weight, testing them five times from admission, using the Bayley scales, until 6 to 13 months after discharge. These children were developmentally delayed upon admission and subsequently the Mental Development Index improved to the normal range, but, there was continued delay in motor development. Age at onset did not predict Bayley scores at outcome, but there was a strong relationship between improved Mental Development scores and change in weight. Those children with a small head circumference remained smaller than other cases throughout the study, they gained weight in a similar pattern and showed no differences in mental quotient from controls. Field argued that the improvements in mental quotient and growth were the result of interventions (four of the children were placed in foster care, and continued contact with the mothers of children who remained at home encouraged them to maintain weight gain and levels of stimulation). Unfortunately testing was not carried out blind to the child's condition, and such frequent testing, albeit at a young age, raises the problem of practice effects.

Singer and Fagan (1984) reported that, on initial assessment, a visual memory test was easily performed by infants with failure to thrive of non-organic (NOFT) aetiology (below the 3rd centile for weight for conceptional age) and normally growing controls. Those infants with failure to thrive with an organic aetiology (OFT) performed the task at chance level. However, Bayley Mental Development Index (MDI) was significantly

lower in both organic and non-organic failure to thrive groups compared with controls at the time they completed the visual memory test (NOFT = 77.6, OFT = 67.7 and Controls = 120.2). They were also lower at 20 months old (NOFT = 80.5, OFT =72.2, and controls = 109.0) and the Stanford-Binet was lower at 3 years old (NOFT = 78.6, OFT = 67.7 and controls = 97.4). As children diagnosed as failing to thrive and controls had comparable results for the Visual Memory Test, Singer et al concluded that the impairment in children with non-organic failure to thrive was not therefore neurological but environmental. It would thus follow that cognitive development could be improved by environmental change. In a later intervention study of 29 children followed up at 3 years old, however, Singer (1986) again found depressed scores using the McCarthy Scales of Children's Ability, but they found no significant difference in IQ attributable to different placement and intervention strategies, suggesting a more limited effect of environment.

Drotar and Sturm (1988) tested 59 infants with decelerating weight gain and weighing less than the 5th centile of the National Center For Health Statistics (NCHS) reference population (Hammill et al, 1979). Bayley scale scores for these infants when first hospitalised had a mean of 99.6, whereas Stanford-Binet scores administered when the children were aged 3 years old had a mean of 85.4. This apparent decline may have been a result of the use of different psychometric tests which measure different cognitive domains, so these differences are hard to evaluate in the absence of a control group. The effect of three types of intervention, family centered, parent centered and advocacy were studied. As in the study of Singer (1986) no significant effect of intervention was found.

Drotar and Sturm (1988) also investigated the effects of certain characteristics of infant growth delay, such as age at onset (defined as the estimated age at which the child was first below the 5th centile), degree of wasting (percentage of weight for height of NCHS standards) and duration of failure to thrive (the time that elapsed between the age at which the child's weight first reached the 5th centile and the child's age at hospitalisation). Age of onset, in a regression analysis, accounted for 10% of variance of IQ; the earlier the onset the lower the predicted Stanford Binet IQ at 36 months. The degree of wasting and duration of failure to thrive did not correlate with

IQ once age at onset was controlled. However, as 87% of cases showed only a mild degree of wasting (7% were severely malnourished) and the age when they were hospitalised was between one and nine months, the range of these measures may have been too small to be a reliable predictor of developmental delay.

The controlled, repeated measures and intervention studies described above have relatively short follow up periods, most children being studied within two to three years of discharge. However, Oates et al (1985) followed up at 13 years of age a group of children studied earlier (Hufton and Oates, 1977). This time a control group was recruited from the child's school, matched for age, sex, social class and ethnic group. WISC Full Scale IQ deficits of 10 points were found in comparison with controls, but this just failed to reach conventional levels of statistical significance ($p = 0.06$). A clearly significant difference was found in Verbal IQ scores of 12 points. The cases continued to have significantly lower reading ages, eight being more than 36 months behind their chronological age in comparison with only one control with the same delay. The cases also scored lower on the Vineland Social Maturity scale (Doll, 1965). However, whilst the controls were matched on some variables, a number of other risk factors for poor cognitive development, such as psychiatric problems, previously reported to be present in the families of the case children (Hufton and Oates, 1977), and maternal education were not controlled for.

As there is little agreement about the timing or persistence of cognitive delay, caution should be exercised about studies which consistently find evidence of developmental delay, but at an unspecified age and under different conditions. Such inconsistency is indicative of practical and methodological problems which may have influenced the characteristics of the samples studied, such as the extent and type of cases lost to follow up, and strategies used to select cases.

High rates of attrition have been a problem with follow up studies. Singer (1986) lost 14 out of 39 subjects in a three year period and Oates et al (1984, 1985) lost 16 out of 30 after 12 years. Drotar (1990) expresses concern that the effect of this form of sample self selection is not adequately evaluated, for example, demographic variables of those lost are not often compared to ensure that they do not differ from the follow

up sample. It has been found that families of infants that drop out of studies have more problems than those who do not (Aylward et al, 1985), and parental permission for a child's participation in a study is less likely to be given for children who are less socially competent and have poor peer relations (Beck et al, 1984), so introducing a possible source of bias into the sample available for follow up. Furthermore, the small number of cases and controls followed up do not provide sufficient statistical power to reliably detect statistically significant differences in outcomes, especially if the effect is small. Small sample size increases the likelihood that statistically significant differences in psychological outcomes found between groups will be false or missed.

Problems have also been identified with selecting children to study who have been referred to hospital or specialist clinics, as they are not representative of all children who fail to thrive. In a study of the detection of failure to thrive in the community, Batchelor and Kerslake (1990) found that a filtering process begins with the initial diagnosis at the primary care level. Sociodemographic data for 39 children whose weight was found to be on or below the 3rd centile were examined. Eighteen were diagnosed as having failure to thrive of non-organic aetiology, eleven with failure to thrive with an organic cause and eight were regarded as just small. Two cases remained undetected. When compared with the remaining low weight children, the eighteen children diagnosed as having failure to thrive of non-organic aetiology had fewer parents who were owner occupiers or in employment, more of them were in single parent families, more had experienced family changes and more were developmentally delayed. Nine of the children who were low clinic attenders and eight who had been on the child protection register were amongst those diagnosed as having non-organic failure to thrive. By contrast, of the eight children also below the 3rd centile for weight, but described as just small, none were developmentally delayed, all were taken to the child health clinic regularly and most came from two parent homes with at least one parent working. Batchelor and Kerslake suggest that, once an organic cause of their growth failure has been discounted, social factors are used to determine diagnostic category instead of growth data, the latter only used to confirm a diagnosis of failure to thrive that has already been made. It is interesting to note that there were no differences in recovery rates between each diagnostic category.

In addition to the filtering process identified by Batchelor and Kerslake during the diagnostic process, Drotar (1990) identified a second filter operating during the referral process. In his discussion of sampling issues in failure to thrive research, Drotar (1990) argues that once a diagnosis has been made, referral may depend upon the severity of the condition, how difficult it is to diagnose, or how responsive it has been to outpatient management. The latter may depend on a physician's judgment about the ability of the caregivers to effect the child's recovery, so favouring referral of children from disadvantaged backgrounds. This combination of detection and referral bias as a result of using psychosocial criteria may explain the association found in clinic populations with family dysfunction, low social class, low IQ and poor educational outcome.

Improved methodology would require the screening of whole populations using growth criteria alone. The advantage of this is that any bias in the case sample, resulting from referral processes, is eliminated. A second requirement is that sample sizes should be large enough to detect the expected difference between the cases and a comparison group. Attrition should also be minimal and where possible those lost to follow up compared with those followed up. None of the studies reviewed so far meet these methodological requirements.

1.2.3 Population based studies of infants failing to thrive

To avoid the known methodological problems of studying referred cases, a number of recent studies have instead screened whole populations to identify cases. But even population based studies have tended to focus on deprived populations, with the expectation that the prevalence of failure to thrive would be higher (Mitchell et al, 1980, Edwards et al, 1990), thus making deprivation a criterion for selection by default. Drotar (1990) recommends that the population from which the sample is drawn should be described, and the effect of sample selection on findings should be considered.

The first study to screen a whole population for failure to thrive was that of Mitchell et al (1980). They argued that failure to thrive is often inferred to be the cause of the deficits identified in these studies and not 'the conditions underlying failure to thrive' and that a suitable comparison is with normal sized peers in the same social environment. They selected all children that met their anthropometric criterion for failure to thrive from a cohort of 312 two to five year old children attending three clinics in a poor rural community in central North Carolina. The cases were defined as children with weights under 80% of the Stuart standards (Reed and Stuart, 1959) for weight in the first 24 months if the previous weight was above 80%. Controls from the remainder of the cohort were matched on sex, age, mother's age and marital status. In addition a measure of life events for children (Coddington, 1972) was used to compare the psychosocial readjustment required to cope with events occurring in the family in a one year period. The cases and controls were compared between 3-6 years of age using the McCarthy Scale of Children's Abilities (McCarthy, 1972). No significant difference was found in cognitive ability between the cases and controls. For both groups 'social turmoil' was found to be a better predictor of the McCarthy General Cognitive Index and behaviour problems than failure to thrive.

Unfortunately, in this study only 19 of the 30 cases were sought for follow up, and only 12 completed a full examination. The high rate of attrition and the relatively mild criterion used for failure to thrive leaves the findings of Mitchell et al (1980) open to question. Nevertheless, Mitchell et al (1980) had made the key methodological advance of selecting samples by population screening and the finding that there was no significant difference in cognitive development clearly required replication. Since then five further studies have screened whole populations with differing results.

One such study was carried out by Dowdney et al (1987). Of 2145 live births registered with an inner city health district community paediatric service, 1868 were traced, of whom 25 children born at term were identified as failing to thrive. The last recorded weight for these children was below the 10th centile, they were still below the 10th centile for height and weight when traced at 4 years old and they remained below the 10th centile when the height of both parents was taken into account. Cognitive assessment of 23 cases was carried out using the McCarthy Scales (McCarthy, 1972).

Results were compared with those of a normally growing control group selected from the remainder of the cohort, matched for age, sex, ordinal position, ethnic origin, and birth weight. A 20 point difference in mean General Cognitive Index (GCI) scores was found (Mean GCI score for cases 77.1 SD 17.6, and controls 97.7, SD 15.2). Statistically, this was a highly significant difference.

This study followed all the design protocols required to produce convincing evidence, in that it was a controlled population based study, with a low rate of attrition, anthropometric selection criteria, and testing was carried out blind. The deficit was very large, especially as the control group were themselves 10 points below the mean GCI scores found in a sample of British children (Lynch et al, 1982). It is interesting to note that only four of the cases had been referred for investigation of failure to thrive, so the cases had not benefited from any intervention.

In this study the exclusion criteria were very rigorous. All those whose last recorded weight was below the 10th centile were regarded as potential cases. After premature babies, those with congenital defects and those whose current weight was above the 10th centile were excluded, 61 of the 138 potential cases remained. During 'paediatric examination and interview with the mother' (Dowdney et al, 1987) the remaining sample of 61 children was further reduced to 25 cases, as those with mid parental height below the 10th centile were excluded. It is unclear from the published paper if mid parental height was the only criterion used during the final interview for the selection of cases.

These results contrast sharply with those found by Mitchell et al (1980) even though, in both, the McCarthy Scales were used to assess cognitive ability. This difference between studies could have been explained by the more rigorous selection criteria used by Dowdney et al (1987) or the effect of selective attrition in the study of Mitchell et al (1980).

There is more consistent evidence to support the view that there are early deficits in development in younger children. This is provided by two community based studies of younger children.

A community based study (Wilensky et al, 1996) conducted in three neighbourhoods in Jerusalem found 55 children from a cohort of 1452 whose weight for age was below the 3rd centile compared with the NCHS reference population. Fifty cases and controls matched for birth weight, sex, age, ethnic origin, and parity, were assessed using the Bayley scales at 20 months. There was a significant difference between the cases and controls in Bayley MDI scales (mean score 99.7 and 107.2 respectively). The principal predictors of MDI in the case group were maternal education and the HOME inventory (Caldwell and Bradley, 1976) which evaluates the family environment. Despite mixed social class, average maternal education levels were poor in both the case and control groups and did not differ. Wilensky et al argue that susceptibility to sub-optimal conditions is greater in the biologically more vulnerable failure to thrive children. The slightly lower birth weights and smaller head circumference in cases at birth reflect greater biological vulnerability, although head circumference did not contribute to the explained variance of MDI. The cases also had more recorded medical problems in their first year.

These results were similar to those found by Skuse et al (1993, 1994a, 1996) in earlier work. They describe a study of 15 month old infants selected from a cohort of 2510 children born in an inner London health district. After attrition and the exclusion of preterm and severely growth retarded neonates, and children with organic disease, there were 49 children who, by 12 months of age, were found to have had a weight Z score of less than -1.88 standard deviations (equivalent to the 3rd centile) for three months or more. Bayley MDI scores were 98.2 for cases and 108.5 for controls with PDI scores of 96.7 and 103.6. The difference between cases and controls in MDI and PDI scales when compared separately failed to reach conventional levels of statistical significance ($p = 0.07$ for both), but together they were significant at the $p = 0.007$ level.

Comparisons were made between children with early and later onset failure to thrive (Skuse et al, 1993). Early and late onset were defined as, 'the difference in weight between birth and 6 months was greater than or less than half the difference between birth and one year'. Skuse et al say that if the difference in weight in the first six

months was greater than half the difference between birth and one year this was indicative of early growth faltering. In order to make sense of the definition, Skuse et al must be referring to weight standard deviation scores which are comparable across ages irrespective of the higher rate of weight gain early in the first year, and not weight as stated. A greater difference in SD score in the first six months than the latter six months would indicate earlier growth faltering.

Interesting differences emerge between the two categories. Bayley scores for the late failure to thrive group do not significantly differ from controls, despite their mothers being more depressed, having less social support and having lower IQ scores than the mothers of early onset cases. However, those infants whose weight gain was lowest in the first six months showed significantly larger Bayley MDI and PDI deficits than the late failure to thrive group. In an analysis of variance of the effect of early or late weight gain on Bayley MDI and PDI scales, the early/late variable explained 21% of the variance of Bayley MDI and 15% of Bayley PDI. One explanation for this difference is that the mean birth weights of the late onset cases was 0.63 of a standard deviation below that of early onset cases, so their total fall was less.

An alternative view is that the first six months is a sensitive period of development when poor growth has more adverse consequences for long term development. To test for this Bayley scores at 15 months were related to standardised weights for age from birth and at ages 4 and 6 weeks, 3, 6, 9, 12,and 15 months (Skuse et al, 1994). A significant correlation between the two variables was found at 6 weeks ($r = 0.31$), three ($r = 0.45$), six ($r = 0.42$), and nine ($r = 0.33$) months, but not at 12 or 15 months. From this Skuse et al (1994) develop a complex statistical model with age at onset, duration and severity of growth faltering, the effect of cognitive stimulation and the effect of congenital malformation as predictor variables for Bayley MDI. The model accounts for 37% of the variance of Bayley MDI and PDI scores and predicts that infants whose weight falls by 2 standard deviations between birth and 6 months will have a Bayley Scale deficit of 10 points, whereas the same fall between four and ten months results in the loss of only 3 points, and there is no deficit when the fall occurs after eight months.

There are two main criticisms of the statistical model devised by Skuse et al (1994), both of which relate to the range of scores used for comparison at different ages. The strength of a correlation is dependent on the range of scores available in a data set, so that the greater the range of scores the stronger the correlation likely to be found between the variables. Skuse et al (1994a) use weight SDS at different ages as a covariate for Bayley scores at 15 months. But, the range of scores for weight SDS for all case children studied, varies with each age band, increasing from birth then getting progressively smaller as the children get older, but of course the range of Bayley scores at 15 months remains constant. The difference of weight standard deviation score between two time points is also used as a predictor. But the fall in weight standard deviation scores is not consistent from age to age, the largest fall in relative weight occuring between 6 weeks and 6 months after which there is a bottoming out effect from nine months of age. The differences between later weight SDS, which included weight standard deviation scores after nine months, will be smaller and will have a smaller range of scores than differences between early weight standard deviations. Thus the strongest associations with Bayley scores at 15 months will be identified at the ages where there is the largest range of weight SDS and between the ages with the largest fall.

With these limitations on the model it would be important to show that earlier weight variables have a significant additional impact on cognitive outcome after current weight had been entered into the regression. In their study of nutritional supplementation of undernourished children in Guatemala, Lasky et al (1981) found that growth spurts and behavioural development covaried, but once current length and weight were entered into the regression, previous measures did not account for a significant amount of the variance of developmental scores at 15 months. Furthermore, deficits in Bayley scores comparable to those found by Skuse et al (1993, 1994) were also found by Wilensky et al (1996), but in the latter study very few cases (n = 6) had reached the 3rd centile by 6 months and no effect of age was found.

One of the questions raised by Lasky et al (1981) and subsequently investigated by Pollitt and Mueller (1982) was whether the relationship between developmental scores and anthropometric measures found in an undernourished population would also be

found in a well nourished population. In a sample of well fed children in the United States, aged between 3 and 6 years of age, it was found that two significant factors correlated with IQ, mother's education and weight for height. Pollitt and Mueller argue that weight for height is a measure of physiological maturation, whereas in an undernourished population height is the most important covariate.

The weight of evidence from the four population based studies discussed suggests that there is indeed an effect of failure to thrive on cognitive development, but that the anthropometric correlates of poorer scores for psychological tests should be investigated. However, it is not possible to dismiss the findings of Mitchell et al (1980) of no significant difference in the the oldest group of children studied. In a follow up of the same children originally studied by Skuse et al (1993, 1994, 1996) at age six years (Boddy, 1997), a difference of only 3.9 McCarthy GCI points was found between the groups, and this was also not statistically significant, thus replicating the results of Mitchell et al (1980). It is possible that the findings of this later study can be explained in terms of the ceiling effect sometimes found in children tested after the age of seven years (Kaufman and Kaufman, 1977) as the control group scored above the mean for Bayley MDI in the first study, and scores for the two tests have been found to correlate (Kaufman and Kaufman, 1977). But the mean age of the children was six years, which is below the age when a ceiling effect is clearly evident. An alternative explanation for these findings is that deficits found early diminish over time.

In order to ascertain whether failure to thrive is associated with more enduring deficits and what the relationship is between anthropometric measures and IQ, a long term follow up study of an annual cohort (n = 306) of children born in a deprived area of Newcastle Upon Tyne was carried out by the author (Corbett, 1994, Corbett, Drewett and Wright, 1996). Fifty two infants in the cohort who had fallen across two or more major weight centile lines for a month or more were identified as cases and originally studied by Edwards (1988). At 6-7 years of age 89% of the cases and controls were followed up and tested at school using the WIPPSI-R (Corbett, 1994, Corbett, Drewett and Wright, 1996). Parents and teachers were asked to complete the Achenbach Child Behaviour Checklist of problem items. All tests were carried out blind to the case status of the child. IQ scores for all children were low, so no ceiling

effect was found, but no significant difference was found between the groups on any of the outcome measures.

A problem with this study was that the anthropometric criterion for selection of the cases, a fall across two or more major centiles, was over inclusive, as a fall of this magnitude is not unusual in larger infants. However, an association was found between IQ and the severity of failure to thrive within the case group, so that, for every one standard deviation a child was below a weight standard deviation score predicted after nine months of age from their six week weight standard deviation score, their IQ was on average 5.8 points lower. The findings do not support the selection of a 5% threshold to define caseness, as there was a linear relationship between the covariates. Nevertheless, since infants with a Thrive Index value below the 5th centile could be expected to have an average deficit of at least 7 to 8 IQ points (or half of a standard deviation) at follow up, this threshold for infant weight gain reasonably delineates the degree of growth faltering related to IQ deficits which would be reliably detectable, and is clearly of educational importance. This association warrants further investigation. However, the association was only tested for within the case group, so it could be argued that this merely confirmed the association between physical maturation and IQ found in a normal population (Pollitt and Mueller, 1982), although, in the Newcastle study, there was no concurrent relationship between IQ and heights retrieved from school records at follow up.

1.2.4 Summary of studies of outcomes of failure to thrive

Previous studies have been reviewed in two parts, those based on samples of cases referred to hospital or specialist clinics, and those based on whole population screening. The methodological and design problems of earlier studies of referred cases have resulted in a number of inconsistencies in findings, but even recent population based studies which have utilised improved methodology fail to agree on the effect of failure to thrive on cognitive development.

Uncontrolled studies of children originally referred to hospital found that, at follow up, most cases had an IQ within the normal range (Glaser et al, 1968, Hufton and Oates, 1977). These studies also report profound difficulties at school, such as poor speech development and reading ability, but these school problems may be better explained by factors other than failure to thrive which have not been controlled for, such as social class or family characteristics.

Controlled, repeated measures and intervention studies found differences in Developmental Quotient in cases dependent on age or the effect of intervention. But these associations were not consistent. The lack of consistency of these findings may be attributable to a reliance on cases referred to hospital, which may result in highly selected and non-comparable samples. However, even where the methodology has been improved by screening whole populations using anthropometric criteria alone, there are contradictory findings. For example, in the two population-based controlled studies both using the McCarthy scales at follow up, one found no significant difference in GCI (Mitchell et al, 1980), another found a 20.6 point deficit (Dowdney et al, 1987).

Another explanation for inconsistent results is that, the small samples used do not have the statistical power to detect quite large effects at any conventional level of statistical significance (Cohen, 1992), especially as variability between samples of biologic, environmental and psychological risk factors produces sample heterogeneity, so increasing the variability in scores. The problem of small sample size is compounded when high rates of subject attrition have reduced the size and the representativeness of the samples studied.

Some consensus has been found in respect of Bayley MDI subscales which were reported to be between 7 and 10 points lower in two studies (Skuse et al, 1994a, Wilensky et al, 1996). But in a further follow up at age six of those children studied earlier by Skuse (Boddy, 1997), these deficits were not found to be persistent, as no statistically significant difference was found in McCarthy GCI. One explanation for this is that deficits found at an early age diminish over time.

Another possibility is that specific characteristics of infant weight gain are predictive of long term cognitive deficits, making a subset of infants more vulnerable. For example, in a study conducted in Newcastle, no statistically significant difference in IQ was found between the case and control groups when the children were aged 6 to 7 years old when tested with the WIPPSI-R (Corbett, 1994, Corbett, Drewett and Wright, 1996), but there was evidence of a persistent deficit as an association was found between IQ and severity of growth faltering in the case group. Other characteristics, such as age at onset of failure to thrive may be associated with cognitive deficits. The evidence for this is confusing, association with age being reported in some studies and not others.

If cognitive deficits were a clear consequence of failure to thrive, then it would be possible to find consensus on the extent of the deficit, the characteristics of infant growth which predict cognitive delay and the timing of such a delay. But the key studies fail to agree on any of these. Clearly some inconsistencies in outcomes could be expected where individual differences in susceptibility and environmental circumstances determine the extent to which a child is vulnerable to early adversity. However, if the criterion for the identification of cases accurately sampled the population of infants failing to thrive and the methodology did not introduce bias, then individual differences would not prevent consensus across studies. It is essential to identify all the methodological differences in these studies which result in these contrasting findings.

1.3.1 Methodological issues

A number of methodological problems have been identified in the review of early studies of cognitive outcomes of failure to thrive. These include poorly defined selection criteria, small sample sizes, lack of comparison groups, reliance on referred cases for sample populations, high rates of attrition, and lack of blind testing at follow up. Methodological improvements have been incorporated into the design of more recent studies which have used carefully defined criteria for the identification of cases

in primary care populations, recruited control groups, achieved very high follow up rates and have carried out all testing blind to the case status of the children.

Two other sources of variation in study samples which may account for the lack of consensus in findings from previous studies, need to be given further consideration. These are the effect of additional exclusionary or inclusionary criteria and the impact of treatment services.

1.3.2 Exclusionary and inclusionary criteria

Drotar (1990) argues that exclusionary criteria restrict the sample and limit the generalizability of findings. It is important though, to propose some limits to make studies comparable. For example, the age in which failure to thrive is deemed to occur is usually the first two years of life, although studies of abused children have included children diagnosed as failing to thrive after this period (Money et al, 1983). Some exclusionary criteria, however, exclude important groups most at risk of failing to thrive, such as low birth weight infants.

Frank and Zeisel (1988) point out, though, that infants with birth weights less than 2500 g are usually excluded, even though more low birth weight than normal weight infants fail to thrive. Low birth weight infants who also go on to fail to thrive have an accumulation of risk factors. In a study of the post natal weight gain of low birth weight infants (Kelleher et al, 1993), all scores for Stanford-Binet IQ at three years were lower than for a standardisation population, but IQ was significantly lower for those low birth weight infants who also failed to thrive post natally, than those who did not.

Another exclusion criterion which is difficult to justify is the distinction made between non-organic and organic failure to thrive, where the aetiology of the latter is attributed to an identifiable organic condition. Infants with poor growth of organic aetiology are often excluded (i.e. Drotar et al, 1985, Dowdney et al, 1987) or provide a separate comparison group (Singer and Fagan, 1984).

A clear distinction between organic and non-organic failure to thrive, however, could not be made in 19 out of 82 children studied by Homer and Ludwig (1981). Failure to thrive, in these cases, had both an organic and non-organic aetiology. That is, a number of children with organic failure to thrive shared some of the characteristics of environmental deprivation generally associated with the non-organic group, such as family dysfunction and evidence of neglect (Homer and Ludwig, 1981). Furthermore, Sills (1978) classified a number of children as having failure to thrive with undetermined aetiology who did not fit either diagnostic classification, that is, they were not sick and there was no evidence of psychosocial deprivation.

Thus, the organic and non-organic categories are not mutually exclusive as there may be multiple causes or a yet undiscovered aetiology, and the use of psychosocial deprivation as a diagnostic criterion for non-organic failure to thrive is likely to result in excluding cases from caring backgrounds and adequate homes. The distinction made between organic and non-organic failure to thrive is merely an attribution which incorrectly implies subgroup homogeneity for the cause of failure to thrive.

As categorisation into non-organic or organic failure to thrive cannot be made with confidence at the time of diagnosis, Drotar (1990) has suggested that diagnosis may be confirmed retrospectively, after the benefit of medical intervention. Evidence of weight gain in a controlled environment, for example in hospital, has been used as confirmation of a diagnosis of failure to thrive (Drotar et al, 1985). However, Frank and Zeisel (1988) point out that when treatment is given, improved growth occurs within a year of diagnosis in only 50% of cases, with 40% stabilising at a lower centile position and 10% of children continuing to deteriorate. The use of catch up growth as a criterion for inclusion would therefore exclude half the diagnosed cases.

It has been argued that evidence for behavioural or developmental deficits should not be used as a criterion for inclusion as it would confound follow up measures (Drotar et al, 1985). At the same time, exclusion of infants with developmental delay at intake precludes studies where improved development quotients can be expected after interventions (ie. Field, 1984).

Exclusionary criteria used in previous studies cannot be justified in terms of identifying a true failure to thrive population. The case for excluding infants on the grounds of low birth weight, organic condition, adequate psychosocial background, or subsequent response to treatment cannot be supported. Infants who fail to thrive are diverse. One recommended way of allowing for this diversity is to analyse subtypes categorised according to risk factor (Drotar, 1990). Careful selection of the sample population is needed to avoid potential sources of bias and larger numbers of cases need to be studied so that statistical power is sufficient despite the variability of aetiological factors.

1.3.3 Treatment services

Drotar (1990) argues that treatment services available in the area studied should be described in order to evaluate possible treatment effects on outcomes, but these may have a limited effect because of poor detection of failure to thrive at the primary care level and uncertain effectiveness of interventions. In the study by Batchelor and Kerslake (1990) they found that 36% of infants below the 3rd centile were not detected as cases of failure to thrive, and in the study by Dowdney et al (1987) only four out of twenty five children had been referred to hospital. Intervention studies have not produced a uniform improvement in cognitive outcome. Field (1984) attributed the improvement in Bayley Scale scores found in her study to multiple interventions and the regular contact maintained with the families. Yet more clearly defined interventions in studies by Drotar (1985, 1988) and Singer (1986) failed to produce an improvement or significant differences in cognitive development between intervention groups.

1.3.4 Summary of methodological issues

A continuing source of variation between studies is the criteria used to exclude cases. Many exclusionary criteria lead to the omission of important groups of infants who

otherwise fit the anthropometric and age criteria of failure to thrive, such as low birth weight infants. Thus samples studied are not representative of all children who fail to thrive, leading to inconsistencies in the findings and factitious associations between failure to thrive and other variables.

Exclusionary criteria used often do not clearly identify categories of cases. An example of this are cases excluded on the grounds that they have failure to thrive with an organic aetiology, even though the organic and non-organic categories are not mutually exclusive. Other criteria have been applied retrospectively, in order to confirm a diagnosis of organic failure to thrive, such as an improvement in rate of weight gain or developmental quotient following intervention. These cannot be justified either, since no treatment has shown significant improvements in all cases and this must lead to the misclassification of large numbers of children who fail to improve. In any case, a number of cases are undetected by conventional methods and thus remain untreated.

It is clear that infants who fail to thrive constitute a heterogenous group increasing the variability in outcome measures. In order to take account of such diversity, a large study population is required before differences in outcomes can be reliably detected.

1.4.1 Psychological and educational outcome measures

So far we have discussed in some detail the different definitions of failure to thrive used in different studies, the cognitive and educational deficits found and the methodological problems in previous work. We now need to evaluate the psychological outcome measures previously used.

The main psychological outcome in previous studies has been either a developmental quotient using a standardised infant test, such as the Bayley scales (i.e. Skuse et al, 1993, Wilensky et al, 1996), or, for older children, the McCarthy scales (i.e.Dowdney et al, 1987, Boddy, 1997) or an intelligence test such as the WISC (Hufton and Oates,

1977). Measures of educational ability or attainment and behavioural ratings have also been used (i.e.Elmer et al, 1969, Hufton and Oates, 1977).

In longitudinal studies, Bayley (1970) found the predictive validity of standardised infant tests disappointing, leading her to conclude that the Bayley Scales were most appropriate as a measure of a child's relative performance at the time, not as a predictor of later performance. In an analysis of a sample of longitudinal studies, McCall (1979) showed that the median correlations for a number of traditional psychometric tests of infant performance (e.g.Gesell, Bayley, Cattell and Griffiths) with childhood IQ (e.g.WISC and Stanford-Binet), increases with the age at which the infant is tested, so the later the infant test is given, the better the prediction of childhood IQ. At the same time, the later the measure of childhood IQ the lower the correlation with infant test scores. Their results are reproduced in table 1.4.1.

Table 1.4.1 Median correlations across studies between infant test scores and childhood IQ (after McCall, 1979).

Age of childhood test in years	*Age of Infant test in months*			
	1-6	*7-12*	*13-18*	*19-30*
8-18	0.06	0.25	0.32	0.49
5-7	0.09	0.20	0.34	0.39
3-4	0.21	0.32	0.50	0.59

Data taken from Anderson (1939), Bayley (1933), Bayley (1954), Birns and Golden (1972), Cattell (1940), Cavanaugh et al (1957), Elardo et al (1975), Escalona and Moriarty (1961), Fillmore (1936), Goffeney et al (1971), Hindley (1965), Honzik et al (1948), Ireton et al (1970), Kangas et al (1966), Klackenburg-Larsson and Stensson (1968) , McCall et al (1972), Moore (1967), Nelson and Richards (1939), Werner et al (1968).

This illustrates the simple truth that the closer the age between two tests, the higher the correlation. But studies of children tested and retested in later childhood do not show such a steep gradient in age to age correlations and the correlations are much stronger. McCall (1979) illustrates this point with data from the Fels and Berkeley Longitudinal studies where the correlations of Stanford-Binet at ages 10, 11 and 12 with Stanford Binet tests at age 9, are 0.90, 0.82, and 0.81 respectively.

McCall (1981) explains the relatively poor predictive power of the infant tests reviewed in terms of minor and transitory individual differences during infancy when normal development is canalised. After the first two years individual differences increase as a result of environmental and genetic influences, stabilising across ages and so correlating more strongly. But part of the problem with the predictive validity of infant development tests is that they do not test the same constructs as childhood intelligence tests: for example, a test of sensori-motor performance may not be expected to correlate with verbal or performance skills as measured in a standard IQ test. McCall (1979) also points out that specific items in infant tests which may be highly predictive of later IQ, such as vocalisation in the first year of life in females, are often included in total scores for infant tests which include a number of items with lower predictive validity. So it is uncertain whether a deficit found in developmental quotient using Bayley Scales has any lasting significance.

There is evidence, though, that the predictive validity of infant tests used in severely impaired clinical populations is better. Largo et al (1990) regularly tested 119 term, 118 preterm and 78 retarded children from nine months of age to nine years. They used a number of different developmental assessments at different ages, such as the Griffiths test (Griffiths, 1954) in infancy and the Wechsler Preschool and Primary Scale of Intelligence (WIPPSI). Interage correlations for subtests of developmental assessments were all higher for the retarded children ($r = 0.26$ to 0.82) than for preterm ($r = 0.17$ to 0.57) or term children ($r = 0.0$ to 0.37). These are difficult to compare as different assessments were used for the retarded children at the same chronological age, but it is also interesting to note that social class was a significant covariate in the term and preterm groups, but not for the retarded group, indicating the relative lack of impact of sources of variability on the retarded group. So the

predictive value of infant tests in the present context depends on the extent to which failure to thrive causes permanent damage which results in a major cognitive deficit.

However, there is evidence that potential IQ deficits in early childhood are diminished in vulnerable groups of infants given preschool education. Zeskind and Ramey (1978) studied children who were foetally malnourished. These infants were observed to have a number of behavioural characteristics similar to the children failing to thrive studied by Wolke et al (1990). They made poor use of available stimulation, had deficient interactive behaviours and a higher pitched cry than normal infants. They were generally described as apathetic, unresponsive and irritable when aroused. In the long term, foetally malnourished infants have also been shown to have IQ deficits and require special education (Sameroff and Chandler, 1975), the severity of the deficit being related to economic deprivation.

In a randomised intervention study of foetally malnourished infants (Zeskind and Ramey, 1978), intervention and non-intervention groups were matched with normally growing controls. From 3 months of age all the infants were given nutritional supplementation, but only half were sent to a day care programme designed to stimulate intellectual growth. At follow up, at 24 months of age, the infants attending day care maintained their earlier score on Bayley mental development index, but both cases and controls who did not receive the day care programme showed a decline in Bayley MDI, the decline being most marked in the non-intervention case group. At 36 months the mean IQ, as measured by Stanford-Binet, of the intervention cases and controls did not significantly differ, but the intervention case group scores were 18 points higher than the non intervention cases (mean IQ of intervention case group = 96.4. Mean of non-intervention case group = 70.6), and the intervention controls scored 11 points higher than the non-intervention controls (Zeskind and Ramey, 1981).

Deficits found in IQ tests, such as the WISC or the Stanford-Binet, administered after infancy, in childhood, are indicative of more enduring impairments. Despite some individual variability (Skodak and Skeels, 1949, Wechsler, 1981), these tests are moderate to strong predictors in child and adult populations of academic achievement, occupation and earnings potential (Zimmerman and Woo-Sam, 1972, Matarazzo,

1972, Herrnstein and Murray, 1994). So deficits found in childhood may have long term consequences.

However, Cattell (1963) argues that in addition to a biological component of intelligence, called fluid intelligence, which is innate, primarily non-verbal and culture free, there is a component of intelligence, referred to as crystallised intelligence, which is based on experience and learning, and is reflected in the extent to which a person has accumulated knowledge. Clearly there is an interaction between the two components, but because of the cumulative nature of crystallised intelligence, IQ test scores are stable and have a high correlation with previous IQ tests. This implies that deficits as a result of a discrete period of early growth faltering would be minimised as new experiences accumulate, and only persistent adverse factors would produce continuity of the deficit. Evidence that the effects of early adversity diminishes is found in a study of children with failure to thrive with an organic aetiology compared with healthy siblings (Lloyd-Still et al, 1974) where there were large deficits in IQ in the younger cases, but very small non-significant deficits where older cases were compared with their siblings. Nevertheless, the stability of IQ suggests that a deficit found in IQ in older children, after they have started school, would provide more convincing evidence of significant cognitive delay with more worrying long term implications.

It must be remembered that intelligence tests measure aptitude for learning, and are not a direct measure of actual achievement. Previous studies of children who failed to thrive who were followed up at school age have found that although most children had an IQ within the normal range (Hufton and Oates, 1977, Oates et al, 1984) they nevertheless had poorer verbal ability, delayed reading ages and increased levels of educational failure. Elmer et al (1969) found that six out of seven of the children of school age studied by them were in special education classes. In these studies direct measures of relevant educational skills and attainment found that children who had failed to thrive were experiencing more serious problems at school at follow up than were indicated by IQ tests alone. Furthermore, Kaufman and Kaufman (1977) found that learning disabled children scored 15 points lower using the McCarthy Scales than the Wechsler Preschool and Primary Scale of Intelligence or Stanford-Binet, suggesting a lack of sensitivity in standard IQ tests as predictors of learning difficulties

in young children. Therefore, another measure of a specific school ability, such as the Graded Reading Vocabulary Test (used in the study of Hufton and Oates, 1977), has been useful as a way of identifying those children with specific problems at school.

In addition to the characteristics of the child being studied, it is important to evaluate parental and familial characteristics that may explain IQ outcomes. Maternal education and social class are often used to take account of the effect of parental influence on outcomes (e.g. Singer and Fagan, 1984, Singer, 1986, Wilensky et al, 1996), along with a number of other measures, such as personality tests and knowledge about their child's education (e.g. Oates et al, 1984, 1985). Wright and Deary (1992), in a critique of the study of the effects of breastfeeding on the intelligence of preterm babies carried out by Lucas et al (1992), argued that in their own work maternal education and social class were found to be poor substitutes for a direct measure of maternal IQ, accounting for only 35.5% of the variance of childhood IQ when IQ shows up to 70% heritability. In a regression of IQ scores for 870 adults onto the mean of their parent's IQ, the regression coefficient was 0.60 (Bennett et al, 1985).

Although mean parental IQ was used in the study of Bennett et al (1985), there is evidence that mothers have a stronger effect on the expression of intelligence than fathers. Reed and Reed (1965) found that low IQ mothers with normal IQ fathers had retarded offspring two and a half times as often than in offspring of mothers with normal IQ and fathers with low IQ. In a longitudinal study of normal children, the IQ of biological and adoptive children was compared (Scarr et al, 1993). The IQ correlation between adoptive mothers, fathers and their biological children at age 7 years was 0.40 and 0.25 respectively, and at age 17 was 0.45 and 0.13. One explanation of these maternal effects is that mother's usually spend more time with their children than father's, creating the intellectual level of the environment (Scarr, 1981). For example, Ruddy and Bornstein (1982) found that maternal stimulation at 4 months is correlated with the size of vocabulary at 12 months, and Ho (1987) found that the mother's verbal responsiveness was correlated with cognitive performance at 2, 3, and 4 years. In order to rule out the possibility that the IQ of children failing to thrive is attributable to factors influenced by the mother, a direct measure of maternal intelligence is most useful.

Maternal IQ scores have been used by Chase and Martin (1970) and Skuse et al (1994) to evaluate their contribution to children's cognitive development in studies of failure to thrive. Chase and Martin (1970) found no significant difference between case and control mothers IQ, and Skuse et al (1994) found that mothers of late onset failure to thrive cases had a significantly lower IQ than early onset cases. In the latter study, maternal IQ when entered as a covariate only accounted for 8% of the variance of Bayley MDI and was not a significant covariate for PDI. But, as with the correlation between infant tests and later IQ, Skodak and Skeels (1949) in their classic study of adoptive families, found little correlation between the biological mother's IQ and test scores for infants aged two, but an increasingly strong correlation up to age 14. This increasing correlation with parents IQ was also observed by Defries et al (1987) and Scarr et al (1993). To my knowledge, no outcome studies of failure to thrive which have tested children after infancy have adjusted children's IQ for the effect of maternal IQ.

1.4.2 Anthropometric outcome measures

In addition to psychological and educational outcome measures, anthropometric outcomes have been used to assess the characteristics of growth since infancy. A child who failed to thrive in infancy and was found to be short and underweight at follow up would have a chronic form of growth failure and, like those children studied by Dowdney et al (1987), cognitive development could be expected to be more severely affected. In studies in Jamaica it was found that stunting rather than wasting was most predictive of cognitive deficits (Grantham-McGregor et al, 1989, 1991). In a Newcastle cohort (Wright, 1996) proportionate short children had more adverse home environments. Long term growth can be characterised using height and weight at follow up.

Head circumference is thought to be indicative of brain growth in the first year of life as the early increase in cell number and its later levelling off is related to a parallel increase and levelling off of head circumference (Winick and Rosso, 1969). In 1969,

Winick wrote of the growing concern that 'a critical period of brain growth may exist during which malnutrition, even in a mild form and even for a short time, may produce irreversible damage.' He reviewed evidence for this in a number of animal studies and studies of malnourished children in the third world (Winick, 1969).

Gunston et al (1992) take a more optimistic view. In a study of children recovering from kwashiorkor, they used magnetic resonance imaging to study the brains of twelve infants in South Africa. Over a period of 90 days the shrinkage of the brain observed in the early stages of treatment was completely reversed in 9 of the 12 children. Gunston and colleagues attributed the shrinkage of the brain to reversible changes in osmolality.

Smaller head circumference has been reported in children who fail to thrive (Chase and Martin, 1970, Field, 1984, Wilensky et al, 1996), although no correlation between cognitive outcome and head circumference has been found (Field, 1984, Skuse et al, 1994, Wilensky et al, 1996).

1.4.3 Confounding of growth, IQ and other variables

Growth and cognitive development are mutually associated with a number of variables which must be controlled in order to avoid confounding. Confounding is when a variable which is independently associated with both of two variables under investigation gives rise to a correlation between them.

Kuh and Wadsworth (1989) studied the parental heights, childhood environment and subsequent adult heights of a national birth cohort of 5362 children born in one week in 1946. Parental height explained 26% of the variance of adult male and female heights. After adjusting children's height for parental height, they found that men were on average 3.8 cm and women 4.7 cm taller in social class 1 than in social class 5. First born children were taller and height declined according to the number of later born siblings. The children of parents educated to secondary education level or higher were taller than those whose parents finished school at 14. Overcrowded homes with

poor amenities, poor standards of maternal care and serious childhood illness all significantly contributed to shorter adult stature. Similarly, Seigel (1982) reported that socioeconomic status, parental education, birth order, quality of parenting and the home environment as measured by the HOME inventory (Bradley and Caldwell, 1976) and reproductive risk factors such as smoking and perinatal illness were predictive of Stanford Binet IQ and Reynell Language expression and comprehension at three years of age in preterm infants. Sameroff et al (1987) found a number of these same environmental factors (maternal education, social class, family size) were independently associated with verbal IQ, as measured by the Wechsler Preschool and Primary Scales of Intelligence, in four year olds. So stature and IQ may be associated through a number of socioeconomic, familial, and medical factors.

One way to control the effects of confounding variables is to recruit a comparison group matched for those variables. Where a comparison group has been used in population based studies of failure to thrive, a strategy used to control for socioeconomic status has been to select controls from the same geographical area (Mitchell et al, 1980) or clinic attended (Dowdney et al, 1987, Edwards et al, 1990, Wilensky et al, 1996). In the study by Wilensky et al (1996) controls were also matched for parental education. This information is not usually routinely collected and therefore is not normally known before the children are studied, so another strategy used to control for its effect is to gather the data at follow up, and then to test for an association, as did Chase and Martin (1970) and Singer and Fagan (1984). The effect of characteristics of the family structure such as family size, birth order and time between birth of siblings have also been evaluated. In the study by Chase and Martin, (1970) familial characteristics, such as more younger siblings and a shorter time between births in the case group, were reported after follow up. However, instead of adjusting for the effects of family structure, Dowdney et al (1987), Skuse et al (1993, 1994) and Wilensky et al (1996) attempted to control for this by matching their cases with controls of the same ordinal position.

Whilst recruiting a comparison group matched for confounding variables appears to be a less complicated strategy in theory, it presents a major practical problem. It would be difficult to find sufficient pairs of children who could be matched on all possible

relevant variables, matching becoming less accurate with each additional variable. For example, Dowdney et al (1987) matched cases and controls on sex, age within three weeks of the case, gestational age, ethnic origin, birth weight and ordinal position. Despite selecting controls from a very large cohort, they reported that they were not entirely successful in matching for ordinal position.

An alternative design would be to adjust outcomes for unevenly distributed potential confounders identified at follow up, using multivariate analysis. A problem with this is that potentially confounding variables need to be carefully identified and measured. Injudicious selection of variables to enter into a multivariate analysis would lead to considerable loss of statistical power. Cohen (1992) argues that to achieve 80% probability of detecting a moderate size effect at the 0.05 level using linear regression requires a sample size of 67, and this would need to be increased by approximately nine subjects for each additional uncorrelated covariate entered into the analysis. None of the controlled population based studies investigating the cognitive development of children failing to thrive carried out to date, have had sufficient statistical power to adjust for confounding variables.

As careful selection of variables is important, risk factors for failure to thrive with the potential to confound psychological outcome measures need to be identified. In a review of the literature on the process of parenting in failure to thrive, Boddy and Skuse (1994) criticise previous work for the use of poorly validated measures and on methodological grounds already raised in this review. They argue that the findings across studies are not consistent and fail to take into account the direction of causality; for example it is not clear whether stress is the cause of failure to thrive, or the effect on the family of hospitalising a child for failure to thrive. Because many studies of failure to thrive have been retrospective it is also not clear which factors are directly associated with the syndrome and which are associated with the reason for referral or a consequence of the diagnosis.

To address the problems identified in retrospective studies, Altmeier et al (1985) studied antecedent risk factors for failure to thrive prospectively. Mothers were interviewed during their first prenatal visit to a clinic in a large inner city hospital in the

USA. For eighteen months after delivery, growth charts were completed during routine attendance at a child health clinic. Case ascertainment was by blind review of growth charts. Of 274 low income families 15 cases were subsequently identified and the data from the prenatal interview compared with 86 randomly selected children. Life stresses for mothers such as more arguments, separation and reconciliations with the father, and the death of a friend, were predictive of failure to thrive. For fathers the predictors were job loss or arrest. Mother's perceptions of childhood nurture were also significant predictors. High levels of antenatal and perinatal problems were identified, mothers of cases tending to gain less weight during pregnancy and have more pregnancy complications. Infants had shorter gestation and were shorter in length at birth, but birth weights were not significantly different. Some of the risk factors for failure to thrive which were identified by Altmeier et al, such as lack of family support and more life events were also found to be significantly associated with lower scores for verbal IQ in the study of four year olds by Sameroff et al (1987).

Where data about potential confounders are gathered retrospectively, recall of past events can be affected by current mood and circumstances (Baddeley, 1990), and data gathered about current life stresses or levels of depression and anxiety may not reflect circumstances at the time the infant presented with failure to thrive. The relevant variables are therefore difficult to verify retrospectively as they have not been selected concurrently with the episode of failure to thrive.

Some early problems may be better documented and verifiable, reducing the need to rely on reported problems. A good example of this is a child's medical history. Childhood illness was found to affect adult height (Kuh and Wadsworth, 1989) and children who fail to thrive have been reported as having a higher incidence of illness in the first few months of life (Sherrod et al, 1984, Wilensky et al, 1996). Some minor recurring illnesses can adversely affect a child's test performance, for example, Mitchell et al (1980) found a higher incidence of ear infection (Otitis media). This can result in periodic hearing loss. The WISC-III^{UK} manual (Wechsler, 1992) discusses the need to take such impairments into account when evaluating test performance.

More difficult to evaluate are the findings of Ramsay et al (1993). They compared feeding skills of non-organic failure to thrive children with children with failure to thrive with an organic origin. Feeding impairment and interactions between the mother and child were similar in both groups, although the onset of failure to thrive was later in the non-organic group. These similarities were attributed to an occult neurophysiological disorder which pre-existed the symptoms of failure to thrive, leading to feeding skills disorder. As maternal recall of infant feeding problems may be affected by subsequent experience with the child and can only be checked in medical records if the child has been referred, oral motor problems may be more easily identified at follow up as speech disorders or poor verbal ability. These have been found to be associated with failure to thrive (Bithoney, 1986, Chase and Martin, 1970, Dowdney et al, 1987, Elmer et al, 1969, Oates et al, 1985).

1.4.4 Summary of outcome measures and data required at follow up

Developmental tests used as outcome measures in infants with failure to thrive have been shown to have poor predictive validity, so it is not clear that deficits found in studies of infants reveal enduring impairment. It is, therefore, important to study children longitudinally.

In the studies reviewed, where children who failed to thrive have been tested after the age of six years, they have not been found to differ significantly from their control group in overall scores, although some subtests have shown a significant difference, and delayed reading ages and high school failure rates have also been found. In these longitudinal studies children's IQ scores were not adjusted for maternal IQ, which has been found to be a strong predictor of children's IQ, maternal education or social class being used instead as a proxy. It would be important to show whether case children's IQ is less than that expected given the mother's IQ.

Anthropometric measurements, particularly height, have been found to be associated with cognitive outcomes in undernourished populations. However, these associations between outcomes with undernutrition are not clearly demonstrated with failure to

thrive, as head circumference, which has been found to be smaller in undernourished populations and children who fail to thrive, was not found to be associated with cognitive outcome in children who fail to thrive.

A number of potentially confounding variables need to be considered. However, these variables are difficult to identify and measure accurately. Information can be gathered at the time of testing, and may be informative about a child's test performance, but, where risk factors for early growth problems are assessed using current information or recall, they may not accurately reflect those at the time the child was identified as failing to thrive.

To control for potential confounders which may not have been identified at the time of recruitment, two population based studies have used control groups matched on a limited number of known variables, adjustment being made for the effect on outcome of confounding variables identified at follow up using regression methods. But despite the large populations from which cases originated, both studies were statistically underpowered for this type of analysis.

In general, a simple association between poor growth in infancy and cognitive development should be treated with caution. Measures of cognitive development may have different long term implications at different ages. Any association found between the covariates may be better explained in terms of a third variable, which is difficult to identify or can only be measured at follow up.

Finally, in order to provide sufficient statistical power to detect a difference in outcome and to adjust for confounding variables, relatively large numbers of cases and controls need to be studied. Despite large scale population screening in recent studies, the numbers of cases identified and studied have not generally provided sufficient power for this type of analysis.

1.5.1 Conclusion

Failure to thrive is a common problem in infancy, is mostly detected by routine health checks and growth monitoring, and may result in referral to hospital. Whilst causing concern at the time of diagnosis, the long term sequelae of failure to thrive have yet to be established. Research which has investigated the psychological outcomes of failure to thrive has produced inconsistent results. In some studies deficits in cognitive development and high rates of educational failure have been found and in others there were no significant differences between cases and controls. Thus, there is no consensus about the extent of cognitive deficits in children who failed to thrive. A number of problems have been identified in previous work which could account for this inconsistency in results.

A problem with many early studies is that they have relied on cases referred to hospital or specialist clinics. These samples are unrepresentative of children who fail to thrive in general, as the detection of cases at the primary care level during routine health checks, and the subsequent decision to refer them during clinical interview, are influenced by psychosocial factors. To overcome these problems, recent studies have identified cases by whole population screening, using anthropometric criteria alone.

Although population screening resolves one difficulty with the selection of cases, it focusses on another. In previous research no consensus has been reached for the anthropometric criteria which should be used to define failure to thrive, raising the possibility that non-comparable samples of cases have been identified. Definitions have tended to focus on attained weight for age without adequately allowing for an individual child's earlier weight. The effect of this would be to identify, as cases, small infants with weight gain within normal limits, and possibly miss infants who have shown a large fall in relative weight, but who remain above the screening threshold. An alternative approach to using an attained weight is to use a velocity based measure to compare weight gain over a period of time with that of a standardisation population. However, this cannot be done by simply observing the relative fall of an individual's weight across centile lines on a weight chart, as this does not take into account regression to the mean.

Sample sizes in most studies have been too small to reliably detect a statistically significant effect or to adjust for confounding variables, and this is so even for a number of recent studies, where cases have been identified by screening large populations. Nevertheless, there is strong evidence from two population based studies testing infants using the Bayley Scales that there is significant developmental delay in infancy.

The long term significance of early developmental delay, in all except high risk infants, has not been established, as infant tests such as the Bayley Scales are not reliable predictors of later intelligence quotients. In order to ascertain if poor growth in infancy has any lasting effect on cognitive development, a long follow up period is required, in order that tests administered at a later age, which are more predictive of academic ability and adult intelligence, can be used.

Most longitudinal studies of children who failed to thrive have been affected by high rates of subject attrition, and their results have not found consistent evidence for deficits in cognitive development, as measured by psychometric tests. Some evidence exists for poor educational attainment in case groups where IQ was not significantly different from controls, but, in these studies, testing was not carried out blind to the children's case or control status. Moreover, none of these studies have taken account of the most important explanatory variable for children's IQ, that is, the IQ of their mother.

Thus a great deal of time is devoted to monitoring infant growth at the level of primary health care, without a full understanding of its long term significance. Previous studies, as a whole, have produced ambiguous results, failing to agree upon the characteristics of infant growth which are associated with cognitive development, at what age an effect is found, or indeed whether there is an effect at all. As this is an issue which may have significant public health and educational implications, further study of psychological outcomes of failure to thrive is needed to investigate the association between poor growth in infancy and later cognitive deficits.

Chapter Two

Aims and method

2.1.1 Aims

In view of the existing strong evidence that developmental delay in infancy is associated with failure to thrive (Skuse et al, 1993, 1994, Wilensky et al, 1996), it would be important to show if this association is also found in children of school age, when intelligence tests are more predictive of adult intelligence (McCall, 1979) and academic ability (Matarazzo, 1972). Although IQ correlates with academic ability, a direct measure of the abilities required for optimal school performance is essential as previous studies have identified specific deficits, such as delayed reading age (Hufton and Oates, 1977, Oates et al, 1985). Reading, in particular, is an essential educational skill and the foundation of educational progress. Thus, the aim of this study, was to investigate the association between poor weight gain in infancy and later intelligence and reading ability.

To provide a strong test for such an association, the design issues raised in the review of previous work needed to be considered. The aim was to use a theoretically correct anthropometric criterion to screen an unreferred population. The population sampled was to be large enough to yield sufficient children to provide the statistical power to reliably detect a difference in IQ between cases and controls, and a number of potential confounders were identified so these could be taken into account. Children were to be followed up after they had started school, between the ages of seven and nine.

2.2.1 Subjects

It has been argued that a study of outcomes of failure to thrive should be based on all cases identified in a community in order to avoid referral bias, and that the sample of cases generated should be large. A survey of the weights recorded for an annual cohort of children still resident in the Newcastle Health Authority area in 1989, and born between April 1987 and March 1988 was conducted in Newcastle Upon Tyne by Wright et al (1994a, 1994b). This work provided the population base for the present study from which cases of failure to thrive were identified.

The birth weight and up to ten subsequent weights to 24 months of age were collected for an annual cohort from child health records and analysed. These data were used to establish the variation in weight gain in a population of term (>37 weeks) infants and the distribution of growth delay in Newcastle Upon Tyne (Wright et al, 1994a, 1994b).

The cohort consisted of 3653 children of whom 235 were born preterm (at less than 37 weeks gestation). For the remaining 3418 infants a birth weight was available for 89%. At least 92% had one recorded clinic weight, 83% a six week weight and 79% a last recorded weight between 9 and 24 months. In the analysis weights retrieved for infants at each age were grouped into six age bands (6 weeks, 3, 6, 9, 12, 18 months) with only the weight nearest to the centre of the band used. At least one weight was available for every age band for 32% of the children and 71% had weight recorded at six weeks and between 9 and 24 months. The numbers available for analysis in each age band are shown in table 2.2.1.

Table 2.2.1 Number of weights recorded at six weeks and with a weight in a later age band (after Wright, 1994a).

Age	*Number of weights*	*Percentage of term children in cohort*
6 weeks	2836	82.9%
3 months	2550	74.6%
6 months	2479	72.5%
9 months	1964	57.5%
12 months	2040	59.7%
18 months	1553	45.4%
Between 9-24 months	2432	71.2%

In order to identify infants whose weight gain was subnormal the relationship between early and late SD scores for all term infants in the cohort at all available time points was first defined. A baseline early weight SD score (SDS_1) was calculated for each child in the cohort by taking the mean of all the weight SD scores available between birth and three months (1-3 values). The original formulation of the Thrive Index used weight at six weeks as the baseline weight (Wright et al, 1994). Weight at six weeks rather than birthweight was used because birthweight is influenced by maternal stature and conditions during pregnancy, leading to poorer correlations between birthweight and later weights than between weight at six weeks and later weights. In this study the mean of the SD scores over the first two months was used, so as to reduce effects of measurement and recording error. Expected weight in SD scores in each age band between 3 and 18 months (SDS_2), conditional upon SDS_1 was then estimated using linear regression:

$$\text{Expected } SDS_2 = r^*SDS_1$$

The regression takes this simple form, with no constant and with r the correlation between SDS_1 and SDS_2, because the scores are standardised. The correlation coefficients (r) were calculated for the baseline score (SDS_1) and weight SDS at 3, 6, 9, 12 and 18 months were 0.82, 0.70, 0.62, 0.63 and 0.60 respectively. Over each of the five time intervals the expected SDS_2 was then used to calculate the Thrive Index, the difference between actual and expected weight (actual SDS_2 – expected SDS_2) for each child in the cohort. The lower 5th centile for the Thrive Index was calculated empirically for each age band (3, 6, 9, 12 and 18 months) and was -0.95, -1.19, -1.36, -1.33, -1.46 respectively. Children were defined as cases if they had Thrive Index values below the threshold in two or more of the five predefined age bands.

This analysis was undertaken by Dr. Charlotte Wright, so that the case or control status of subjects was not disclosed to the author.

2.2.2 Selection of controls

A paediatrician (Dr. Charlotte Wright), and a research nurse (Jane Callum) selected the controls. The cases were stratified by age, sex and the deprivation indices of the area in which they were resident between 18 and 30 months old. The deprivation levels in the city had been mapped and classified as deprived, intermediate and affluent as described by Wright et al (1994b). Previous classification of administrative health wards (Townsend et al, 1988) using census data on car and home ownership, overcrowding and unemployment, did not take into account the social diversity found in discrete pockets in adjacent areas within the same ward. Consultation with health workers throughout the city produced a more precise mapping of neighbourhoods with homogenous levels of deprivation. Each of these neighbourhoods was then given a deprivation classification of affluent, intermediate and deprived, based on the aggregate of its constituent environmental deprivation scores (Townsend, 1989).

Control children whose actual weight SDS was within one standard deviation of their predicted SDS were then chosen from the same sex, age and deprivation strata as cases in equal numbers. In most cases the next child on the GP's register of the same age (within one month), sex and area deprivation classification was selected. Where no control was available from the same GP practice, a control that fulfilled the criteria was selected from a GP practice that served an area with similar levels of deprivation. All children when aged 7 to 9 years old were traced from child health records, and their details entered onto a database. If a case had moved out of the area, the control was retained (n = 4). If a control had moved out of the area, a new control was selected (n = 14).

The case-control status of all children was withheld from two psychologists (Sally Corbett and Kathryn Parkinson) who were to administer intelligence and reading tests to the children. This was to ensure that all psychometric testing was carried out blind to the group membership of the child.

2.3.1 Materials

Children's IQ and reading ability were tested. Their height, weight and head circumference were measured. Mothers were also given an IQ test, and their height was measured. Mothers were also interviewed using an interview schedule. The tests and measures used are described below.

The children were all tested, in school, using the WISC-III^{UK} (Wechsler, 1992) and the Wechsler Objective Reading Dimension (Psychological Corporation, 1993).

The WISC-III^{UK} for children aged 6 years to 16 years is the latest edition of the WISC-R^{UK}. The test incorporates a number of changes of importance to this study. For example, poorer test performance was anticipated for cases, and although the younger children seen were 7 years of age, a year older than the youngest age recommended for

use with the WISC-III^{UK}, there was concern that there would be a floor effect. Easier items have been introduced to the Arithmetic, Picture Arrangement and Block Design tests, such as picture items in the Arithmetic test to assess counting and number concepts.

As the WISC-III^{UK} is new there is a paucity of reliability and validity studies. However, the WISC-III^{UK} was standardised on a British population of 824 children and a comparison was made with the US WISC-III data of scaled scores derived from raw scores, which were found to be similar, so findings from studies of the WISC-III can be generalised to the WISC-III^{UK}.

Correlations between the Verbal, Performance and Full scale scores for the WISC-R and WISC-III administered in counterbalanced order to a sample of 206 children were 0.90, 0.81, and 0.89 respectively. Although correlations were high there were absolute differences in the point of origin of IQ scores. Verbal, Performance and Full Scale scores for the WISC-III were 2, 7, and 5 points lower than the older WISC-R as the WISC-III has been restandardised to reflect secular changes in IQ scores.

The high test re-test reliability of the WISC-III is widely recognised. The stability co-efficients for Verbal, Performance and Full Scale scores found in a study of 353 children divided into three subgroups are 0.90 to 0.94, 0.86 to 0.88, and 0.92 to 0.95 respectively (Wechsler, 1992).

As all IQ tests were administered by the author, interscorer reliability was not a problem. In those subtests which require more judgement (Similarities, Vocabulary and Comprehension) reasons for scoring decisions were recorded and used as the basis for later scoring decisions to ensure consistency over time.

The WISC-III^{UK} was administered omitting the supplementary and optional subtests. These subtests were omitted to ensure that testing could be completed with minimal disruption to the school, as break times were about an hour and a quarter to an hour and a

half apart and testing usually took about one and a quarter hours for each child. It was also important to retain the child's interest and enthusiasm for the whole test. If testing had taken longer, it would have been necessary to give the child a break or return to complete the test on a second visit.

The most important justification for using the WISC-III^{UK} in an educational context is its moderate to strong concurrent validity for a range of achievement tests, its moderate correlation with Teacher Ratings and its ability to discriminate children with moderate and specific learning deficits from control groups, as reported in the manual (Wechsler, 1992).

However, although the WISC tests are a well known predictor of academic achievement with concurrent validity coefficients of 0.50 to 0.60 for academic achievement tests (Zimmerman and Woo-Sam, 1972), a relevant measure of actual achievement in a vital educational skill, reading, was also considered important in order to establish the direct educational relevance of findings, especially as previous studies had reported educational failure amongst children who fail to thrive (Elmer et al, 1969, Hufton and Oates, 1977, Oates et al, 1985).

The Wechsler Objective Reading Dimension (WORD) is an individually administered test for the assessment of reading in children aged from 6-16 years of age, which takes up to half an hour to administer. Developed in the United States and tested on a sample of 4252 children there, WORD was standardised alongside the WISC-III^{UK} during the validation study carried out on 824 UK children (of whom 794 completed WORD as well). An additional 850 tests were carried out to provide data from UK children for the adjustment of US transformation tables (Psychological Corporation, 1993).

The choice of the WORD depended on two key features of the test. Firstly, it had been standardised with the WISC-III^{UK} allowing analysis of a potential discrepancy between the means of IQ and reading scores. This may be important as other studies have found poorer reading scores for cases than might be expected given IQ (Oates et al, 1985). A

second reason for choosing WORD is that unlike many reading tests, such as the Neale Analysis of Reading Ability (Neale, 1989), scores are available for the poorest 5% of readers in the 7 year old age range. For example, the minimum standardised scores for the three WORD subtests are 64, 69, 60 for Basic Reading, Spelling and Reading Comprehension respectively. These scores would be expected given an IQ score of 40.

Each test comprises three subtests, Basic Reading, Spelling and Reading Comprehension, from which a Composite reading score can be derived. However, the Reading Comprehension subtest is not given if Basic Reading and Spelling Scores do not exceed a total of 8, so in those cases a Composite score cannot be calculated. Instead subtest scores can be used in statistical analysis of results in order to ensure that data from all children, especially poor readers, are included.

The manual reports high reliability coefficients. Split half reliability coefficients for seven to nine year olds were 0.95 to 0.96 for Basic Reading, 0.93 to 0.96 for Spelling and 0.91 to 0.93 for Reading Comprehension. Test-retest reliability coefficients after a mean test interval of 17 days were 0.94 for Basic Reading and Spelling and 0.85 for Reading Comprehension.

The Reading Comprehension test is the subtest requiring most judgement and therefore the main source of scorer error. A study, reported in the manual, to test interscorer reliability for the Reading Comprehension test, using 50 testers, found a reliability coefficient of 0.98. Although another psychologist carried out most of the reading tests, 17 were carried out by the author. In order to improve interscorer reliability, both psychologists were trained in the use of the reading test together and tested a number of children, separately and together, prior to the study. They also conferred frequently where there was some ambiguity in the score to be given to a response and a list of these scoring decisions was maintained.

The manual provides evidence of content, construct and criterion related validity. Content of the test was examined during its development to reflect current curricular trends, the content of school texts and other achievement tests. This was reviewed throughout the development of the test. Construct validity of the WORD is supported with evidence from intercorrelations among the subtests, correlations with the WISC-III and multitrait-multimethod study of WORD and other achievement tests. Moderately strong correlations between the subtests are reported in the manual ($r > 0.65$) for seven to nine year old children. Correlations with the WISC-III^{UK} in this age range were highest for Verbal IQ and the subtests ($r = 0.36$ to 0.77) in particular Basic Reading ($r = 0.48$ to 0.76) and Reading Comprehension ($r = 0.49$ to 0.77). Correlations between measures from tests such as the Richmond Test of Basic Skills and the WORD which measure the same constructs were between 0.71 and 0.78. Evidence for criterion related validity consists of comparison of WORD scores with scores from other achievement tests and the test's ability to correctly discriminate children who are gifted or those with learning difficulties. High correlations were found ($r = 0.79$ to 0.87) between the Basic Reading and Spelling subtests and the equivalent tests in other batteries (BASIS, K-TEA, WRAT-R, DAS) and the mean scores of the corresponding subtests are almost equal despite different standardisation populations.

Data are presented in the manual which shows that mild learning difficulties were detected in 91 children with a mean WISC Full Scale IQ of 89. Given the IQ, expected mean standardised scores for the subtests were 93, 94, and 93. However, the actual mean scores were 79.8, 76.8, 83.7. Discrepancies between IQ and standardised subtest scores for Basic Reading of 9.75, for Spelling of 10.18, and for Reading Comprehension of 10.60 are significant at the $p = 0.05$ level in the 7 to 9 years old age group. This suggests sufficient sensitivity to detect mild learning difficulties amongst the children in the present study, most of whom are attending main stream school and have not been referred for assessment for special educational needs.

2.3.2 Anthropometric measurement of children at follow up

Height, weight and head circumference were measured by a research nurse on another visit to the children at school. Height was measured accurate to 0.1 cm using the Leicester Height Measure following a standard procedure. The children removed their shoes and stood with feet in the position marked on the measure. They were asked to inhale, then allow their shoulders to drop, their position being maintained by supporting the jaw as they exhaled. The head was positioned in the Frankfurt plane so that the cross bar rested flat on the head. Head circumference was measured accurate to 0.1 cm using the tape for the Leicester measure positioned above the supra orbital ridges. Weight was measured accurate to 100 g using portable electronic scales (SECA Scales, Model Number 835). The children were asked to wear light clothing (T shirts and shorts) and were weighed without shoes. On occasion, when the weather was cold or school heating was not working, the children were not asked to remove warm clothing, but the type of clothing worn was coded. The child's weight was then adjusted for the weight of clothes worn, the weight of clothes being estimated from test weighing of similar types of clothes.

2.3.3 Interview on the home visit

An interview schedule was completed with the mothers at home, their height measured, accurate to 0.1 cm with the Leicester Height Measure and their IQ tested using the WAIS-R^{UK} (Wechsler,1981 and Lea, 1986).

The Interview schedule was constructed using the EPED editor in EPI INFO Version 5 (Dean et al, 1990). The interview had three purposes; to check information already available from child health records, to gather data about variables that might be associated with the children's IQ and educational outcomes, and to gather data about variables associated with physical growth. Closed questions were mainly used. In order to reduce the likelihood of recording error responses were mainly recorded in alpha numerics such

as <Y> for 'yes' and <N> for 'no', but some responses were recorded numerically, the numbers referring to predesignated categories. A paper version of the Schedule was produced using Word for Windows (Appendix II) which included open questions designed to give some additional information on health and feeding problems.

Information was gathered under key headings:

IDENTIFIERS

SECTION A - General information about your child

SECTION B - About your child's birth

SECTION C - About your child's health

SECTION D - About child care

SECTION E - About you

SECTION F - About your family

The rationale for items included in the schedule is described below.

IDENTIFIERS

1. Study Number

A study number was allocated which matched the record number on the EPI INFO data entry programme. This number indicates the approximate chronological order in which tests were carried out and made administration and double data entry easier to manage.

2. Identifier

The identifier was the original number allocated to each child in the annual birth cohort during the conduct of the earlier population study of infant growth (Wright et al, 1994a). Children were listed by GP practice, then in birth order for each practice and given sequential identity numbers. This number is the common identifier for all the study children and is used to identify subjects in all data files.

3. Date of interview

With mother's date of birth this variable could be used to calculate the exact age of the mother at the time of the interview.

SECTION A - General information about your child

This section recorded the current name and address of the child, either of which may have changed as a result of remarriage, and the child's date of birth, age and sex. These data were used to update the database, and to check and cross reference data held in different files.

The main aim of the home visit was to test each child's mother's IQ, so the carer was asked what their relationship to the child was to confirm that the person interviewed was the natural mother, or to record the biological relationship of the carer interviewed if it was not possible to see the mother.

Information about the carer was also collected to ensure that the person being interviewed had relevant knowledge of the child's background. They were asked whether they had always been the only carer or if not, who was the main carer and for how long.

SECTION B - About your child's birth

Mothers were asked if there had been any medical problems in pregnancy and if so, to describe them. They were asked if it had been a difficult birth and how long the second stage of labour had lasted, whether the baby had been delivered by caesarean section, or whether forceps or vacuum extraction had been used, and were asked to describe any other complications. Mothers were also asked whether the baby had arrived early or on time, and if early, by how many weeks.

Mothers were asked if they smoked. This question referred only to their current status as smokers and no questions were asked about the number of cigarettes. It was felt that a

more detailed account of smoking may have jeopardised the co-operation we sought to maintain.

SECTION C - About your child's health

These questions related to three issues. The first was feeding behaviour, in particular whether the child was ever breast fed and if so for how long. Mothers were asked if there had been problems feeding their child and to describe these difficulties in detail. The second issue was whether the child had been treated in hospital for a medical condition. All serious problems that lead to hospital inpatient or outpatient admissions were recorded and described. The age of the child when seen in hospital was recorded for up to two admissions, although all admissions were noted on the schedule. Where a mother reported medical problems during pregnancy, or parturition, or problems with their child's health, the name of the hospital and the consultant was recorded and an additional consent was sought (Appendix III) to consult medical records in order to ascertain if the condition was likely to affect growth or performance on psychological or educational tests. The third issue was to determine whether the child had any disability which would affect test performance. Mothers were asked about their child's sight, hearing and speech. In addition to a number of direct measures of speech and language ability, Bishop et al (1995) used simple verifiable criteria such as referral to a speech therapist and length of time the child had received speech therapy to ascertain if there was a history of severe speech impairment for participants in their study. In order to distinguish speech problems likely to have a large influence on IQ test results from relatively minor problems, the mother was asked if their child had ever been referred to a speech therapist, and if so, for how long the child was seen by the speech therapist. Mothers were also asked about any history of learning difficulties in school and the assessment stage reached according to the Department of Education Code Of Practice on the identification and Assessment of Special Educational Needs (Code of Practice, 1994). They were asked if anyone in the family has speech or reading difficulties, and if so who in the family is affected as Bishop et al (1995) found a significant degree of heritability in their kinship study of speech and language disorders

SECTION D - About child care

This section determined how much pre-school education the children had before starting main stream school. Mothers were asked if their child had attended playgroup or nursery school before going to school, how long their child attended playgroup or nursery school and how many days per week they attended.

SECTION E - About you

Ethnic origin was recorded. To assess maternal education mothers were asked if they left school at 16 years old, where they completed their full time education, and what qualifications they gained.

SECTION F - About your family

Factors relating to the structure, size and level of economic deprivation of the family are recorded in this section. Data were gathered about the siblings of the study child. The date of birth, age and sex of each sibling for up to ten children, the number of children in the family and birth order of the study child were recorded. It was also noted if the child was a twin. To ascertain the level of economic deprivation, data were collected about three main indicators of wealth: employment, home and car ownership.

2.3.4 Anthropometric measurement and assessment of parents

The person interviewed, usually the mother, was asked to estimate the natural father's height in feet and inches. The height of the mother was measured using the Leicester Height measure following the standard procedure described for measuring the height of children. The author was trained in this procedure by a research nurse.

2.3.5 Assessment of maternal IQ

A four subtest short form (Silverstein, 1982) of the WAIS-R^{UK} (Psychological Corporation, 1986) was administered to the mothers after completion of the interview schedule. The four tests given were Picture Arrangement, Vocabulary, Block Design and Arithmetic. Using data from the WAIS-R manual each sum of scaled scores was transformed to a standard score of 100 with an SD of 15 to provide an estimate of the Full Scale IQ (Silverstein, 1982). This combination of subtests is highly correlated with the Full Scale WAIS-R ($r = 0.92$ to 0.95, depending on age group) with reliability coefficients of between 0.90 to 0.95 depending on age group (Silverstein 1985).

A two subtest short form (Silverstein, 1982), using Vocabulary and Block design, was used for four carers who were concerned that they did not have time for the four subtest Short Form. Reported reliability coefficients, and validity coefficients as measured by correlations between the Short Form and the Full Scale are > 0.89 and > 0.87 (Silverstein 1985).

Both short forms are recommended by Groth-Marnat (1990) as the best combination of subtests as Vocabulary and Block design are both stable measures of g and the addition of arithmetic and picture arrangement 'provides an assessment of auditory attention [...] how effectively a person functions in the real world [...] a person's knowledge of sequencing and his or her relative perceptiveness regarding common social situations'. Silverstein (1985) calculates the standard error for each age group for the estimate of IQ for the four and two subtest short forms as between 4.97 and 6.61, and 6.07 and 7.59 IQ points respectively.

For this study, training in the administration and scoring of the WISC-III^{UK}, WAIS-R^{UK} and WORD was given by a Clinical Psychologist. He observed a number of trial tests on children and adults and checked the scoring. He also advised how to explain test scores clearly to interested parents so as not to cause unnecessary anxiety about their child. In

the case of parents who had well founded concerns about their child's progress at school and test performance, it was important to ensure that they knew who to discuss these concerns with, and assure them that access would be given to the child's results to the appropriate professional with the parent's permission.

2.4.1 Design

The Newcastle cohort has provided a rare opportunity to follow up a group of children in the long term who have well documented data on weight and levels of deprivation (Wright et al, 1994a, 1994b), but have not been influenced by previous attention. The data used for these studies were retrieved from child health records and primary health care professionals and at no stage were any of the children or their families contacted in the conduct of this earlier research. The cases had never been physically examined, or referred for treatment, or even seen by the paediatric medical personnel involved in this research; and many cases had not even been detected by primary health care professionals during routine assessment either. As testers were also blind to the clinical status of the children, in many cases the study has the strength of a double blind design.

The design is a prospective follow up study of a cohort born at term. Routinely recorded weight data for the infants were screened using a new longitudinal criterion for failure to thrive. Those children with a rate of weight gain in the lowest 5% on two or more occasions were identified as cases. As all those who met the growth criterion in the sample population were classified as cases, no other clinical judgement was made in the sample selection, thus eliminating the effect of referral bias. The 5% threshold was selected for two reasons. Previous population based work in Newcastle (Corbett, 1994) suggests that a fall of this magnitude is associated with significant cognitive deficits and the threshold observes a recommended convention (Peterson et al, 1985, Drotar, 1990).

Power calculations were made to ensure that the yield of cases from the sample population had sufficient power to make a type two error unlikely, even after expected attrition. A 5% threshold used to define abnormal weight gain, when applied to the annual birth cohort with the required number of weights studied (n=3418) would by definition identify 170 children as cases, with slightly fewer using the stricter criterion of weight SDS below 5% on two or more occasions. Power calculations detailed in Table 2.4.1 show that in order to have an 80% chance of detecting a statistically significant difference of five IQ points, in a one tailed test at the $p<.05$ level, 120 subjects would be needed in each of the case and control groups. This allowed for some attrition.

Table 2.4.1 Power calculations according to Kraemer and Theimann (1988)

Effect size $\delta = (\mu_x - \mu_y) / \sigma$	$\delta = (5-0) / 15$ $\delta = 0.33$
Critical effect size $\Delta = \delta / (\delta^2 + 1 / pq)^{1/2}$	$\Delta = 0.33 / (0.33^2 + 1/ (0.5 \times 0.5))^{1/2}$ $\Delta = 0.33 / (0.1089 + 4)^{1/2}$ $\Delta = 0.33 / 2.027$ $\Delta = 0.16$
Number of subjects required for a one-tailed test at the 5% level with 80% power, where v is tabled against Δ	$n = v + 2$ $n = 239 + 2$ $n = 241$
Therefore number of subjects required for each arm of the study	n cases = 120 n controls = 120

Cases were stratified by age, sex and level of deprivation and controls recruited in equal numbers in each stratum. Weight SDS of the controls had remained within one standard deviation of predicted weight standard deviation. Thus the distinguishing feature between the groups is the presence or absence of poor growth in infancy as a risk factor.

At follow up, outcome measures are IQ, reading ability, height, weight and head circumference. Demographic, medical history, parental anthropometry and maternal psychological test data were to be collected in order to control for potential confounding.

2.5.1 Procedure

In order to contact the children and to examine their medical records it was necessary to obtain ethical approval, and consent from parents and school head teachers. After approval and consents were obtained, parents were visited at home before testing began. Testing was carried out in schools and on a second home visit. Relevant medical records were then examined, data entered into a spreadsheet and analysed. Finally, parents and schools were debriefed. Each of these stages is descibed below.

Approval was sought and gained from the joint Newcastle Health Authority/University of Newcastle upon Tyne Ethics Committee to assess the children and their mothers using standardised tests and a questionnaire. The Director of Education of the City of Newcastle upon Tyne agreed to allow permission to be sought from Head Teachers to test the children in school. Head Teachers were asked for permission to carry out testing in school (Appendix IV).

A letter was sent to the mothers (Appendix V) to ask consent for their participation in the study. If they agreed, they were visited briefly at home in order to explain that the study was designed to look at the relationship between patterns of growth in infancy and current

growth and development. The wording was important as many mothers were unaware that their child had failed to thrive by the criterion used in this study. The nature of the testing and the time when the child would be seen at school and possible times for a home visit were discussed. A form to consent to the child being visited at school was signed by a parent, usually the mother (Appendix VI).

If no letter had been returned, a visit to the last known address was made. Quite often parents had received a letter and were waiting for a visit, in which case they were given the usual explanation and asked for the second consent. If families had moved and they had left no forwarding address, they were retraced through the school nursing service, or medical health records and a letter was sent to their new address and the same procedure followed.

After signed consents had been obtained the children were visited at school on three occasions, firstly by the author to administer the WISC-III^{UK}, and on the second visit by Kathryn Parkinson, who administered the Wechsler Objective Reading Dimension. On the third and final school visit their height, weight, head circumference and blood pressure were measured by Jane Callum (the research nurse). The anthropometric measurements were carried out as part of a separate study of the growth of the original cohort which will be reported elsewhere. However, the relevant anthropometric data has been made available for analysis by Dr. Charlotte Wright.

Arrangements were made with schools to use a quiet room or workspace. Each child was asked if they minded helping to find out how children grew. At the end of the IQ test they were told that Kathryn would come to see them soon to do some reading with them. No child expressed reservations about the IQ test; Kathryn Parkinson offered to discontinue a reading test with one child who seemed apprehensive, but the child decided to complete testing. Tests were administered using standardised procedures as described in the manuals. Children were encouraged by praising their effort, not their responses, as recommended by Sattler (1988), and after the IQ test the child was thanked for their help

and given a letter of thanks addressed to them, and after the reading test they chose a sticker.

After the children had been tested at school the mothers were contacted by letter (Appendix VII) and a home visit was arranged. During the visit the mothers were interviewed using the interview schedule and where necessary were asked for permission to consult their child's medical records. Their height was measured using the Leicester Height measure, following the standard procedure previously described. Finally, the mothers were given the four subtest short form of the WAIS-R^{UK}.

All mothers who took part whose first language was not English completed the interview schedule, sometimes with the help of an interpreter who was usually a close relative or friend, but they were not asked to complete the WAIS-R^{UK} as it would have been too difficult to standardise the interpreter's instructions to them.

Mothers often used this opportunity to discuss the results of their child's test at school. The strategy used was to encourage the mother to talk about their child's school performance first. In most cases where an IQ was low, mothers were already aware of a problem and the results of a test were explained with the qualification that the tests only measure some abilities and may vary with circumstances on the day of testing. It was also made clear to the mothers that the author was not professionally qualified to interpret the results of the tests for them. Where problems were identified by the mother, possible sources of help were discussed and the author offered to convey test results to any professional the mother chose to consult, with the proviso that they must be shown to a Clinical or Educational Psychologist. Only one mother availed herself of this offer.

2.5.2 Coding of medical history

As failure to thrive can sometimes be due to an organic condition which may affect psychological outcomes as well, a distinction needed to be made between failure to thrive of non-organic and organic aetiology. Often infants with organic failure to thrive are excluded from further study. However, this distinction between 'organic' and 'non-organic' failure to thrive cannot easily be maintained as Homer and Ludwig (1981) found that the categories were not mutually exclusive, some children having failure to thrive with both an organic and a psychosocial origin.

Since the distinction between organic and non-organic failure to thrive is blurred, an alternative approach is to include all children who fail to thrive and make a statistical adjustment to outcomes by non-organic or organic attribution. This adjustment of outcomes can be made at two levels; where there is clear evidence of an organic condition which causes poor growth, and where the evidence is less clear. It would also be important to review and categorise conditions likely to affect cognitive outcomes as well as growth. For this reason the medical history of all the children studied was reviewed.

Following completion of data collection, using the study number known only to the author as an identifier, information was abstracted from the interview schedule about health problems during pregnancy, parturition and the child's early years. Each case was then reviewed blind by Dr. Charlotte Wright to ascertain whether a medical problem leading to hospital admission was likely to have resulted in poor growth or poor cognitive performance at follow up. Hospital records were then consulted for all children with a condition relevant to growth and cognitive outcomes. The medical data were coded according to the certainty (definitely or possibly) that the condition would directly or indirectly affect growth (such as an endocrine disorder or undernutrition) or cognition, as shown on Tables 2.5.2a and 2.5.2b. The codes ensure that all information from the interview schedule was coded. The medical history and coding schedule is included in the appendix (Appendix VIII)

Table 2.5.2a Codes for medical condition related to growth

Code	*Classification*	*Description*
1	DEFINITE	Major condition likely to cause stunting or major undernutrition
2	POSSIBLE	Recurrent episodes of relevant illness sufficient to cause chronic ill health which may cause stunting or poor nutrition
3	IRRELEVANT	Any admission for a single episode of illness
4	NO RECORD	Hospital admission reported by mother but no record found
5	RECORDS NOT CONSULTED	Hospital admission reported by mother unlikely to be significant. Records not consulted
6	NO ADMISSION	No hospital admission reported by mother

Table 2.5.2b Codes for medical condition related to cognitive outcomes

Code	*Classification*	*Description*
1	DEFINITE	Major condition likely to cause cognitive impairment.
2	POSSIBLE	Recurrent episodes of relevant illness sufficient to cause chronic ill health which may cause cognitive impairment.
3	NO HISTORY	No history of relevant illness

2.5.3 Data entry and databases used for analysis

All psychometric and interview data were entered onto an Epi5 database (Dean, 1990) by the author, then independently by an administrative assistant and checked for errors using the Epi5 validation facility. The anthropometric data were entered in a second file by the research nurse and the same double entry procedure was used. This second file also contained the case/control status of the child. The separate data files were merged for analysis.

Analysis of cross sectional follow up data was carried out using Epi5 by running programmes for data analysis written in EPED the word processing package. Data were exported to a file in SPSS-X format for further analysis.

2.5.4 Debriefing of study participants

Following analysis of the data a letter was sent to all Head Teachers (appendix IX) and mothers (appendix X) to thank them for their help and to convey the overall findings of the study.

Chapter Three

Demographic characteristics, attrition and medical history of the study sample

3.1.1 The sample selected

Using the procedure described in chapter two, 136 infants, whose rate of weight gain fell below the 5% threshold on two or more occasions, were screened in as cases. Of those selected 54 were boys and 82 were girls. One hundred and thirty six controls of the same age, sex and level of deprivation of their area of residence at eighteen months of age were selected from the remainder of the cohort.

As described, three separate data files were created during the course of this research. Matching of files was by an identity number common to all files and records were checked for errors by also matching for sex and date of birth. Date of birth in the first file was supplied by the mother, and in the second file was transcribed from medical records held by the GP. Five discrepancies were found in the date of birth recorded. Only one was for more than a few days, a discrepancy of one month for a case. However, retrieval of subsequent weights from the records for this case showed that the later ages recorded independently by health visitors corresponded with the date of birth entered onto the computer on the day of the child's birth. It is unlikely that such a discrepancy would not have been detected by the health visitor during clinic visits where dates are checked for the purpose of recording weights and sending appointments for vaccinations and health checks at specific ages. Inaccuracy in the date of birth recorded during the interview may be explained by incorrect information given by the mother or incorrect recording of the information given. Thus the date of birth recorded on the child health computer was used to calculate age of the child for weight SD scores and WISC-III^{UK} and WORD tests.

Analysis of the gestational age reported by mothers revealed that five children, one case and four controls, were reported as having a gestational age of less than 37 weeks, and did not apparently meet the selection criterion of having a term birth. Three of these children (one case and two controls) had gestations recorded in child health records of 37 weeks and were found to have birth weights within one standard deviation of the population mean making preterm delivery unlikely. As the evidence suggests a term birth, and there is often disagreement between the mother and her obstetrician about the expected date of delivery, in these instances the gestational age entered onto the child health record was accepted as correct. In two instances the gestational age had not been retrieved from the original records and so these children (two controls, one girl and one boy, born at 32 and 36 weeks) were excluded from further analysis. Thus the analyses were based on data for 136 cases and 134 controls.

3.1.2 Subject attrition

Children lost to follow up were identified and classified according to information held on the project database indicating whether they had been traced or not and the reasons recorded for non participation. A comparison of those lost and those followed up was then carried out.

After exclusions for previously unrecognised prematurity, of the remaining 270 children selected, 224 (82%) children were studied when aged between 7 and 9. One child had died. Those lost to follow up were 14 children who were untraced, and 4 cases and 14 controls who had moved from the area. The fourteen controls who had moved were replaced by the next child on their GP's list that matched the selection criteria. Twenty seven parents refused permission for their child's participation in psychological testing, although some of these parents had separately agreed for their child's height and weight to be measured. Attrition is summarised in Table 3.1.2.

While a similar proportion of case and control boys were lost to follow up, the number of girls lost to follow up in the case group was disproportionately more than for those lost in the control group (ratio of 2.2 to 1). In the case group 11 boys (20%) and 18 girls (22%) were lost to follow up, and in the control group 9 boys (17%) and 8 girls (10%).

Table 3.1.2 Loss to follow up for psychological testing

	Cases	*Controls*	*Total*
Total number of children identified	136	136	272
Short gestation (<37 weeks)		-2	-2
Died	-1		-1
Eligible for follow-up	135	134	269
Untraced	-9	-5	-14
Moved away	-4	-14	-4
Replaced		+14	
Consent withheld	-15	-12	-27
Eligible children studied			
Psychological measures	107 (79%)	117 (87%)	224 (83%)
Anthropometric measures	111 (82%)	122 (91%)	233 (87%)

3.1.3 Anthropometric measures of those lost to follow up

An analysis of the effect of attrition was carried out using infant weight data, to ascertain how representative those cases studied were of all cases identified. Mean weight standard deviation scores of cases not psychologically assessed and cases who were assessed were calculated for each age band. These data can be seen in Table 3.1.3a. Cases not tested were relatively smaller than those who were, both at birth when their weight was 0.4 of a standard deviation lower, and throughout infancy. The birth weight of cases lost to follow up was only marginally below the mean of the standardisation population (SDS = - 0.045), but was not significantly below that of the cases followed up ($t = 1.62$, $df = 131$, $p = 0.10$). There were no statistically significant differences in weight SDS between cases lost and those followed up in any subsequent age band.

Given the heavier birth weight of cases followed up, the Thrive Indices of the cases were compared in order to establish whether the rate of weight gain in infancy of children not psychologically assessed was more severely affected than those that were. When the mean Thrive Index was calculated for each age band they were very similar with no statistically significant differences (Table 3.1.3b).

Controls were selected on the basis that their weight was never lower than one standard deviation below the population mean, so, their weight gain in infancy was above average. It can be seen in Table 3.1.3c that, in contrast to the case group, the controls lost to follow up were heavier than those controls followed up. However, none of the differences were statistically significant for any age band, and the trajectory of weight gain was similar.

Table 3.1.3a Weight SDS of cases lost to follow up and cases followed up.

	Cases lost to follow up			*Cases studied*				
Age	*Mean SD*	*sd*	*n*	*Mean SD*	*sd*	*n*	*t*	*p*
Birth	-0.04	1.20	29	0.37	1.20	104	1.62	0.11
1 month	-0.64	1.03	26	-0.3	1.07	98	1.38	0.17
2 months	-0.90	0.91	25	-0.65	1.10	91	1.04	0.30
3 months	-1.23	0.76	28	-1.18	1.05	101	0.26	0.79
6 months	-1.63	0.67	29	-1.56	0.92	100	0.40	0.70
9 months	-1.88	0.67	28	-1.71	0.88	89	0.96	0.34
12 months	-1.87	0.91	23	-1.66	0.81	89	1.07	0.29
18 months	-1.63	0.98	21	-1.62	0.90	72	0.05	0.96

Table 3.1.3b Thrive Indices of cases lost to follow up and cases followed up.

	Cases not studied			*Cases studied*				
Age	*Thrive Index*	*sd*	*n*	*Thrive Index*	*sd*	*n*	*t*	*p*
3 months	-0.84	0.63	28	-1.00	0.57	101	1.21	0.23
6 months	-1.31	0.59	29	-1.42	0.51	100	1.01	0.31
9 months	-1.59	0.46	28	-1.59	0.51	89	0.09	0.93
12 months	-1.51	0.54	23	-1.54	0.51	89	0.22	0.83
18 months	-1.44	0.85	21	-1.43	0.57	72	0.10	0.92

Table 3.1.3c Weight SDS of controls lost to follow up and controls followed up

	Controls lost to follow up			*Controls studied*				
Age	*Mean SD*	*sd*	*n*	*Mean SD*	*sd*	*n*	*t*	*p*
Birth	0.04	0.99	17	-0.16	1.16	114	0.65	0.50
1 month	0.16	0.68	16	-0.25	0.99	104	1.61	0.11
2 months	0.17	0.74	15	-0.18	1.00	98	1.30	0.19
3 months	0.35	0.83	17	-0.07	0.94	113	1.76	0.08
6 months	0.45	1.13	16	0.08	0.96	109	1.39	0.17
9 months	0.70	0.95	12	0.22	0.99	84	1.58	0.12
12 months	0.52	1.04	12	0.19	0.92	91	1.16	0.25
18 months	0.44	1.59	8	0.17	0.93	68	0.71	0.48

There were some missing anthropometric data at follow up at eight years old. However, more parents consented to have their children measured at school than to participate in the psychological testing. Of those children administered psychological tests or visited at home, two cases and two controls were lost or not available by the time the anthropometric measurements were made. Of the remainder, all cases, there were missing data for one boy with cerebral palsy who only agreed to have head circumference measured; a girl with a major organic condition who became distressed, so anthropometric assessment was discontinued without measuring head circumference; and a girl whose weight was not recorded. On three further occasions, the children had moved out of Newcastle and were measured by the author rather than the research nurse. Head circumference measures made by the author however, were not satisfactory when compared with measurements made by the research nurse, so these measurements were

omitted. Four of the children with incomplete anthropometric measures had a major organic condition. Of the non-organic cases only one had missing data for height, two for weight and four for head circumference, so the impact of these missing data on the analysis will be small.

The anthropometric measurements of cases not psychologically assessed were compared with that of cases who were. Although the cases not tested are on average shorter and lighter at follow up than the cases tested, using a t-test for two samples or a median test (Seigel, 1956) as appropriate, no statistically significant differences were found in the height, weight and head circumference between the cases tested and those cases who were not. However, the number of children measured but not given psychological tests is small and so, the power of a significance test to detect a true difference is low. These data are shown in Table 3.1.3d. Those controls lost to follow up were on average shorter and lighter and with a smaller head circumference than those controls followed up. The differences were smaller than those for cases and were also not statistically significant. These data are shown in Table 3.1.3e.

Table 3.1.3d Anthropometric measures at age eight of cases not given psychological tests and those that were tested

	Cases not tested			*Cases tested*				
	mean	*SD*	*n*	*mean*	*SD*	*n*	*t*	*p*
Height (cm)	124.59	4.19	7	126.05	5.69	103	0.67	0.50
Head circumference (cm)	51.00	1.33	7	51.95	1.76	101	1.40	0.16
	median	*IR*	*n*	*median*	*IR*	*n*	χ^2	*p*
Weight (kg)	23.4	20.6 to 27.6	7	24.25	21.7 to 7.1	102	0.86	0.39

Table 3.1.3e Anthropometric measures at age eight of controls not given psychological tests and those that were tested

	Controls not tested			*Controls tested*				
	mean	*SD*	*n*	*mean*	*SD*	*n*	*t*	*p*
Height (cm)	129.36	6.06	7	130.82	5.89	115	0.64	0.52
Head circumference (cm)	52.60	1.95	7	52.80	1.74	115	0.29	0.77
	median	*IR*	*n*	*median*	*IR*	*n*	χ^2	*p*
Weight (kg)	27.6	23.7 to 29.3	7	28.4	25.4 to 32.0	115	0.76	0.44

3.1.4 Missing psychological test data for those children who were studied

Of the children studied, WISC-III^{UK} and WORD data were not obtained for two cases who were too impaired to be tested at all and one control child who was too impaired to complete the reading test. In two of these instances the mothers were interviewed at home but were not asked to complete the IQ test (WAIS-R^{UK}). As one mother withdrew consent to be interviewed after her child had been tested and five mothers repeatedly failed to keep appointments, it was not possible to collect data from six mothers of children who were studied. Of those missing, three mothers were cases and three were controls. A further 13 WAIS-R^{UK} scores were not obtained from mothers from the Indian sub-continent (6 cases, 7 controls) whose first language was not English. Missing data are summarised in Table 3.1.4.

Table 3.1.4 Missing psychological test data of those children studied

	Cases	*Controls*
Total number of children studied	107	117
WISC-III^{UK} and WAIS-R^{UK} data missing	2	
WORD data missing	2	1
WAIS-R^{UK} and demographic data missing	3	3
WAIS-R^{UK} for mothers from the Indian sub-continent	6	7

3.1.5 Summary of study attrition

In any longitudinal study it is regrettable that some subjects will be lost to follow up. Loss of subjects occurs for a number of reasons, such as moving away from the area. The children in this study and their families had never been contacted by any member of the research group, although the children's infant weight data had been collected from child health clinics six to seven years before. Of the 136 cases originally identified, 79% were traced and given psychological assessment, and 87% of the 134 controls. Height and weight were obtained for 82% of cases and 91% of controls. The higher follow up for controls can largely be explained by the replacement of those who had moved out of the area. Cases could not be replaced as all cases are already selected.

It was important to establish if loss to follow up produced a bias in the cases studied, that is to say that the severest cases were more likely to be lost. The cases lost to follow up were slightly lighter in infancy than those cases tested, but they were not more severely affected than cases followed up. All these differences were small as no statistically significant differences were found. In contrast to this, the controls lost had slightly higher

standard deviation scores, although these did not differ significantly in any of the age bands from those controls followed up.

At age eight years, those cases and controls who were not given psychological tests but were measured, were shorter and lighter on average than those followed up. None of the anthropometric differences were statistically significant, although numbers are small.

The possible effect of attrition can only be speculated upon. However, it is clear that although the differences in infant weight standard deviation scores and anthropometric measures at age eight between those cases studied and those lost are marginal, the cases lost to follow up were slightly smaller. The controls lost were on average heavier in infancy and slightly shorter and lighter at age eight.

3.2.1 Demographic variables and medical history.

A comparison of the demographic and medical history data was carried out to ascertain if any of the variables measured were significantly different between the case and control groups. In most analyses the data were categorical and are presented in a two by two contingency table. Chi-square tests of association are used to test for a statistically significant difference between the groups. For continuous data an independent samples t-test or a Mann-Whitney U test is used as appropriate.

3.2.2 Age and sex of children followed up

The mean age of the children when tested with the WISC-III^{UK} was 8.12 years (SD 0.62) for cases and 8.10 years (SD 0.57) for controls. Mean ages of cases and controls when given the WORD reading test was 8.3 years (SD 0.54) and 8.22 years (SD 0.54)

respectively. At follow up 43 boys and 64 girls in the case group were studied, and 44 boys and 73 girls in the control group.

3.2.3 Child care and the characteristics of the main caregiver

All but two of the carers interviewed were mothers. The two were both fathers of control children, one a widower and one who had gained custody of the child and had subsequently lost contact with the child's mother. All the parents were asked if they had always been the main carer; three control parents, including the two fathers previously mentioned, and one case parent said they had not. If the carer interviewed had not always been the main carer, they were asked who had cared for the study child. Three children (two controls and one case) were cared for by their grandmother, and one control child was cared for by an aunt. The mean (SD) age of case and control carers at the time of interview was 35.9 (5.73) years and 36.4 (5.32) years respectively ($t = 0.66$, $df = 200$, $p = 0.5$).

There were no statistically significant differences in the age of the main carers, or the care arrangements of the children whose carers were interviewed. Almost all were cared for mainly by their mothers. Of the four who were not, all were cared for within the family and none had been in foster care or had been adopted. The parents of five children who were repeatedly unavailable for interview had all been present for a first home visit at which a signed consent for their child's participation in the study was obtained. This suggests that none of these children was in care. However, one case was living with foster parents, and the parent was not available for interview. For this child consent for participation was sought from the parent by Newcastle upon Tyne City Council Social Service Department.

3.2.4 Medical background reported by mothers

Mothers were asked if they had any medical problems during the pregnancy or the birth of their child (Tables 3.2.4a and 3.2.4b). Twenty six per cent of mothers of cases reported medical problems during pregnancy compared with 19% of mothers of controls. This was not a statistically significant difference. But 38% of mothers of cases reported problems during parturition compared with 24% of controls, and this was a significant difference ($\chi^2 = 4.90$, df = 1, p = 0.03).

Gestational age of the children ranged from 37 weeks to 43 weeks. The mean gestational age reported for controls (39.7 weeks) was slightly lower than that for cases (40.1 weeks). The difference was statistically significant (Mann Whitney U = 2.07, p = 0.038). These data are shown in Figure 3.2.4.

Mothers were asked if their child had ever been admitted to hospital or attended a hospital outpatient clinic with a medical problem. Disproportionately more children in the case group had attended hospital as an inpatient or outpatient (63% of cases and 49% of controls) and this difference was statistically significant ($\chi^2 = 4.54$, df = 1, p = 0.03). Fewer mothers of case children said they smoked (33% of cases and 45% of controls), but the difference was not quite statistically significant. These data are shown in Tables 3.2.4c and 3.2.4d.

Table 3.2.4a Medical problems during pregnancy

Medical problems during pregnancy	*Cases*	*Controls*	*Total*
Yes	27	22	49
No	77	92	169
Total	104	114	218

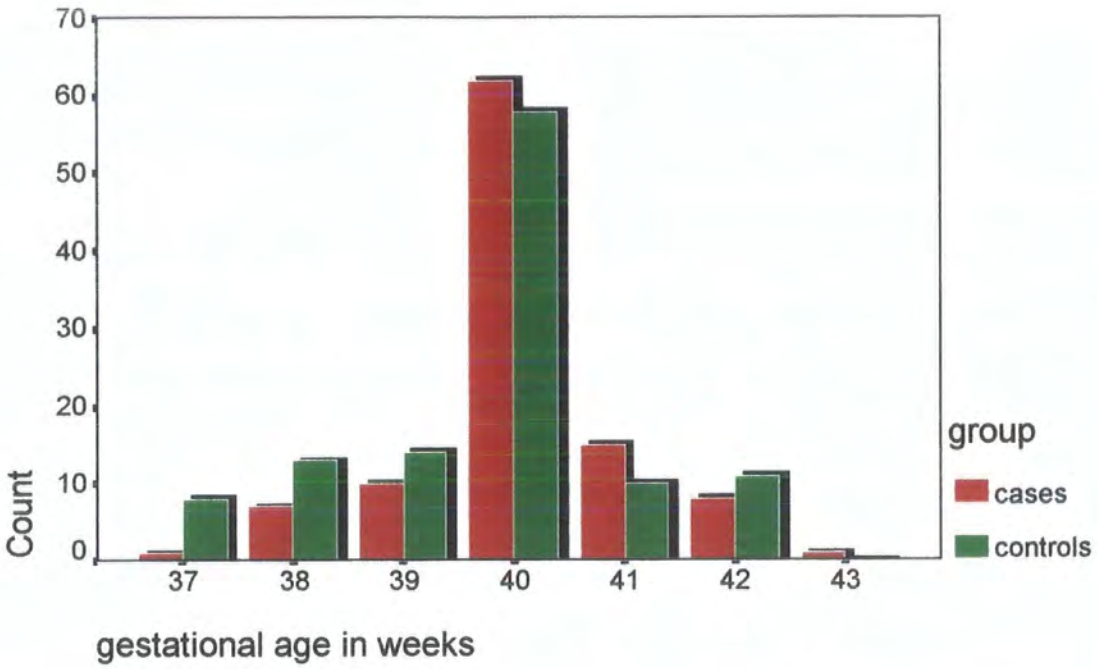
$\chi^2 = 1.39$, df = 1, p = 0.24

Table 3.2.4b Medical problems during parturition

Difficult birth	*Cases*	*Controls*	*Total*
Yes	40	28	68
No	64	86	150
Total	104	114	218

$\chi^2 = 4.90$, df = 1, p = 0.03

Figure 3.2.4 Gestational age by case and control group



Mann Whitney U = 2.07, p = 0.038

Table 3.2.4c Children who had received treatment in hospital as an inpatient or outpatient

Hospital Admission	*Cases*	*Controls*	*Total*
Yes	66	56	122
No	38	58	96
Total	104	114	218

$\chi^2 = 4.54$, df = 1, p = 0.03

Table 3.2.4d Mother's who smoked

Mother smokes	*Cases*	*Controls*	*Total*
Yes	34	51	85
No	70	63	133
Total	104	114	218

$\chi^2 = 3.32$, df = 1, p = 0.068

3.2.5 Feeding History

The number of mothers who had ever breast fed their child was almost identical between the groups (58% of cases and 55% of controls). Mothers who had breastfed their infants were asked approximately how long they had continued breastfeeding. Answers were categorised as follows: less than a week, 1-6 weeks, 6 weeks - 4 months, more than 4 months. There was no statistically significant difference between groups ($\chi^2 = 1.29$, df = 3, p = 0.73). These data are shown in Tables 3.2.5a and 3.2.5b.

When asked if they had problems feeding their child 47% of case mothers said they had, compared with 25% of controls and this was a significant difference ($\chi^2 = 12.11$, df = 1, p = 0.0005). These data are shown in Table 3.2.5c. All feeding problems described by mothers have been listed in Appendix XI. When these were broadly categorised, the problems reported for cases were similar but more numerous than those of controls. Feeding problems also appear to be more persistent in the case group as the number of reported problems decline more slowly after weaning (shown in Table 3.2.5d).

Table 3.2.5a Breastfeeding by case and control group

Ever breast fed	*Cases*	*Controls*	*Total*
Yes	60	63	123
No	44	51	95
Total	104	114	218

$\chi^2 = 0.13$, df = 1, p = 0.7

Table 3.2.5b Length of time breastfeeding was continued

Breast fed	*Cases*	*Controls*	*Total*
Less than a week	5	8	13
1-6 weeks	11	15	26
6 weeks to 4 months	17	16	33
More than 4 months	26	24	50
Total	59	63	122

$\chi^2 = 1.29$, df = 3, p = 0.73

Table 3.2.5c Feeding problems by case and control group

Feeding problems	*Cases*	*Controls*	*Total*
Yes	49	28	77
No	55	86	141
Total	104	114	218

$\chi^2 = 12.11$, df = 1, p = 0.0005

Table 3.2.5d Type of feeding problem reported for cases and controls

Type of problem	*Cases*	*Controls*
Bottle or breast feeding		
Intubated	111	1
Not feeding	111111	
Problems with breast feeding	11111111111	1111111
Aversion to milk	11111	11
Small quantity consumed	11111	11
Colic	111	11
Distress when feeding	11	11
Vomiting	111111111111	1111
Slow feeder	11111	11
Poor suck	11111	1
Diahorrea		1
Weaning		
Didn't like solids	1111111111111	111
Problems swallowing	111	1
Problems chewing	111111	
Slow eater	1	
Poor appetite (not interested in food)	1111	111
Faddy (eats limited range of foods)	11111111111111	1111
Small portions	11	

3.2.6 Physical impairments likely to affect test performance

Twenty nine children out of 218 were reported as having impaired vision, with a slightly lower proportion of cases. This was not a statistically significant difference. Thirty eight mothers out of 218 reported that their children had hearing problems. These were almost evenly distributed between the groups (17% of cases and 18% of controls). These data are shown in Tables 3.2.6a and 3.2.6b.

Table 3.2.6a Impaired vision by case and control group

Visual problems	*Cases*	*Controls*	*Total*
Yes	13	16	29
No	91	98	189
Total	104	114	218

$\chi^2 = 0.11$, df = 1, p = 0.74

Table 3.2.6b Impaired hearing by case and control group

Hearing problems	*Cases*	*Controls*	*Total*
Yes	18	20	38
No	86	94	180
Total	104	114	218

$\chi^2 = 0.00$, df = 1, p = 0.96

In order to ascertain the prevalence and severity of speech problems by group, mothers were asked not only if their child had a speech problem, but if they had been referred to a speech therapist and if so for how long. These data are shown in Tables 3.2.6c and 3.2.6d.

Forty three out of 218 mothers reported that their children had speech problems. There was no significant difference between groups for reported speech problems (20% for cases and 19.3% for controls). Twenty eight of the 43 children with impaired speech had been referred (16 cases and 12 controls) and of those referred, the cases were twice as likely to

receive speech therapy for more than one year, although this difference did not achieve statistical significance as numbers were small.

Mothers were asked if anyone else in the family had speech difficulties, reading difficulties or speech and reading difficulties (data shown in Table 3.2.6e). In spite of the trend towards case children experiencing more severe speech problems, there was no significant difference in familial speech and reading difficulties between groups (χ^2 = 0.03, df = 3, p = 0.87). When asked which member of the family was affected (sibling, parent, grandparent, or other) proportionately more relatives of cases (34 out of 104 cases and 33 out of 114 controls), in particular siblings, were reported to have speech or reading problems. However, more parents of controls were also reported as having speech and reading problems and the differences between the groups failed to reach conventional levels of statistical significance (data shown in Table 3.2.6f).

Table 3.2.6c Impaired speech by case and control group

Speech problems	*Cases*	*Controls*	*Total*
Yes	21	22	43
No	83	92	175
Total	104	114	218

χ^2 = 0.03, df = 1, p = 0.87

Table 3.2.6d Length of time children with impaired speech received therapy

Speech therapy	*Cases*	*Controls*	*Total*
More than a year	8	3	11
Less than a year	8	9	17
Total	16	12	28

$\chi^2 = 1.80$, df = 1, p = 0.18

Table 3.2.6e Familial speech and reading difficulties

Familial speech and reading difficulties	*Cases*	*Controls*	*Total*
No problems	71	82	153
Speech difficulties	15	18	33
Reading difficulties	10	10	20
Speech and reading difficulties	8	4	12
Total	104	114	218

$\chi^2 = 1.94$, df = 3, p value = 0.58

Table 3.2.6f Relationship of learning impaired relative

Relationship of member of the family with speech or reading difficulties	*Cases*	*Controls*	*Total*
Sibling	23	16	39
Parent	2	9	11
Grandparent	0	1	1
Other	9	7	16
Total	34	33	67

The Parent, Grandparent and Other categories were collapsed for analysis as expected cell size was less than 5.
$\chi^2 = 2.53$, df = 2, p = 0.28

3.2.7 Children with special educational needs

More cases (14.4% of cases and 9.6% of controls) were registered by the Education Authority as having special educational needs (Table 3.2.7a), but this difference was not statistically significant ($\chi^2 = 1.18$, df = 1, p = 0.3).

In order to ascertain the severity of learning difficulties, mothers were asked which stage of special needs assessment their child had reached (Table 3.2.7b). There are four stages; the last two (stages 3 and 4) require assessment by an educational psychologist. Although, according to the protocol followed in the Newcastle area, parents should be informed in writing of the stage a child has reached in their special needs assessment, many find that it is difficult to distinguish between stage one and two, or three and four. For this reason the stages were placed into two categories which are easily distinguished by asking whether the child has been assessed by an educational psychologist or not. Significantly more cases were at assessment stage three or four ($\chi^2 = 4.55$, df = 1, p = 0.03).

Table 3.2.7a Children on the register of special educational need

On the register of special educational needs	*Cases*	*Controls*	*Total*
Yes	15	11	26
No	89	103	192
Total	104	114	218

$\chi^2 = 1.18$, df = 1, p = 0.3

Table 3.2.7b Assessment stage for special educational need

Assessment stage	*Cases*	*Controls*	*Total*
Stages 3&4	9	2	11
Stages 1&2	6	9	15
Total	15	11	26

$\chi^2 = 4.55$, df = 1, p = 0.03

3.2.8 Preschool education

Preschool education was assessed by recording if the child had attended playgroup (Table 3.2.8a) or nursery school (Table 3.2.8b), for how many months (Figure 3.2.8) and for how many days per week (less than four days per week or four days or more: Table 3.2.8c). The same percentage of cases and controls attended preschool playgroup (58%) and there was only a small difference for those attending nursery school (88% for cases and 84% for

controls). Only two controls and no cases did not attend any form of preschool group. There was almost no difference between the groups in the number of months they had attended preschool education. Twenty nine per cent of cases went to nursery for more than four days per week compared with 24% of controls, but the difference was not significant.

Overall, there were no significant differences in the amount of preschool education received by the case or control groups when measured by the type of preschool provision, the length of attendance or the number of days attended per week.

Table 3.2.8a Playgroup attendance

Attended playgroup	*Cases*	*Controls*	*Total*
Yes	60	66	126
No	44	48	92
Total	104	114	218

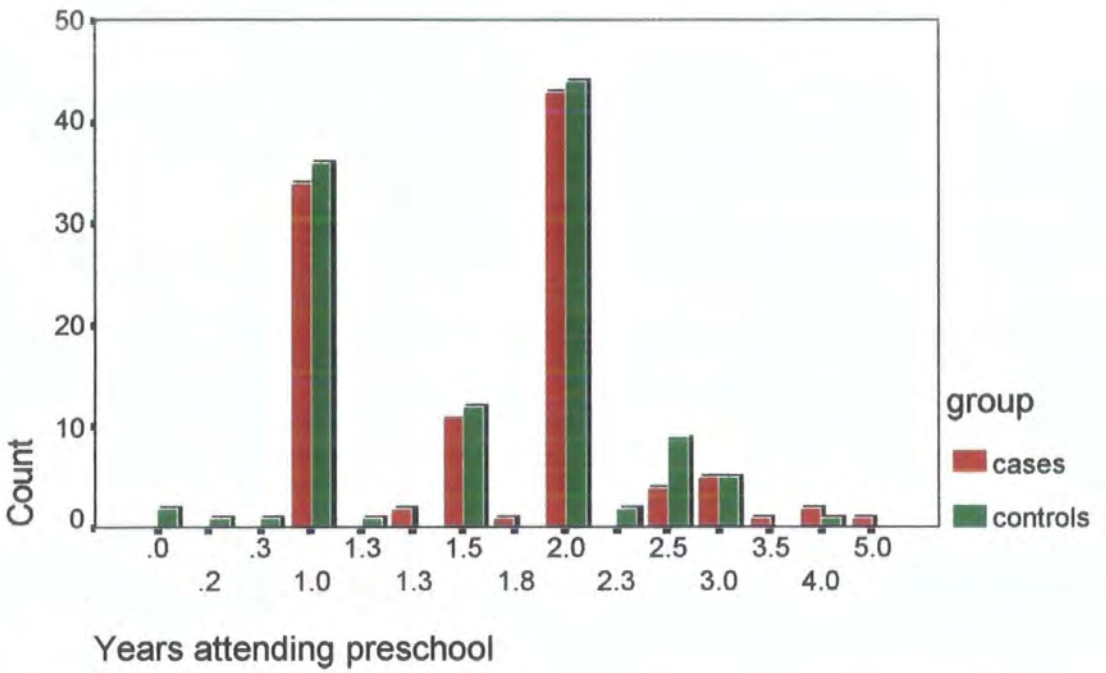
$\chi^2 = 0.00$, df = 1, p = 0.98

Table 3.2.8b Nursery school attendance

Attended nursery	*Cases*	*Controls*	*Total*
Yes	91	96	187
No	13	18	31
Total	104	114	218

$\chi^2 = 0.48$, df = 1, p = 0.49

Figure 3.2.8 The number of years children had attended preschool by case and control group



Mann Whitney U = 0.3, p = 0.75

Table 3.2.8c The number of days nursery school attended per week

Number of days attended per week	*Cases*	*Controls*	*Total*
Less than four	30	27	57
Four days or more	74	85	159
Total	104	112	216

$\chi^2 = 0.62$, df = 1, p = 0.43

3.2.9 Ethnicity

As ethnic origin was not one of the original selection criteria for controls and is known to be associated with verbal IQ (Sameroff et al, 1987), it was important to analyse between group differences in ethnic origin. There were 14 mothers from the Indian sub-continent (6 cases and 7 controls) who were all either bilingual or who spoke only their language of origin. There was one Afro-Caribbean mother whose first language was English. As numbers were small, the Afro-Caribbean and Indian sub-continent categories for ethnic origin were collapsed for analysis. There were no significant differences in ethnic composition of the groups ($\chi^2 = 0.03$, df = 1, p = 0.86). Table 3.2.9.

Table 3.2.9 Ethnic composition of the case and control groups

Ethnic group	*Cases*	*Controls*	*Total*
Caucasian	97	107	204
Afro Caribbean	1	0	1
Indian sub-continent	6	7	13
Total	104	114	218

The Afro-Caribbean and Indian Subcontinent categories were collapsed for analysis
$\chi^2 = 0.03$, df = 1, p = 0.86

3.2.10 Maternal Education

Mother's education was assessed by the age at which they finished full time education, the type of institution attended after leaving school and the qualifications received. These data are shown in Tables 3.2.10a, 3.2.10b and 3.2.10c. Seventy four per cent of mothers of cases had left school at 16, and 67% of mothers of controls. Those mothers who stayed at school were asked where they completed their full time education. Although fewer case

mothers stayed on, the type of institution attended did not differ significantly (χ^2 = 3.53, df = 3, p = 0.32). Case mothers had fewer qualifications, but the difference was not statistically significant (χ^2 = 1.72, df = 5, p = 0.89).

There was a trend for mothers of case children to have left school earlier, and to have slightly lower levels of educational attainment, but these differences are very small and none were statistically significant.

Table 3.2.10a Mothers who left school by the age of 16 years

Mother left school at 16 years old	*Cases*	*Controls*	*Total*
Yes	77	77	154
No	27	37	64
Total	104	114	218

χ^2 = 1.11, df = 1, p = 0.29

Table 3.2.10b Institution where mothers completed their full time education

Institution where mothers completed their full time education	*Cases*	*Controls*	*Total*
Sixth form	5	4	9
Further education college	6	13	19
Institute of higher education	13	12	25
University	3	8	11
Total	27	37	64

χ^2 = 3.53, df = 3, p = 0.32

Table 3.2.10c. Qualifications reported by mothers

Qualifications	*Cases*	*Controls*	*Total*
None	45	42	87
'O' Level	23	30	53
'A' Level	3	3	6
Diploma	3	4	7
Degree	10	15	25
Other	20	20	40
Total	104	114	218

$\chi^2 = 1.72$, df = 5, p = 0.88

3.2.11 Family structure

The study child's siblings' dates of birth, sex and ages were recorded. Family size and birth order were separately noted in order to cross check the data.

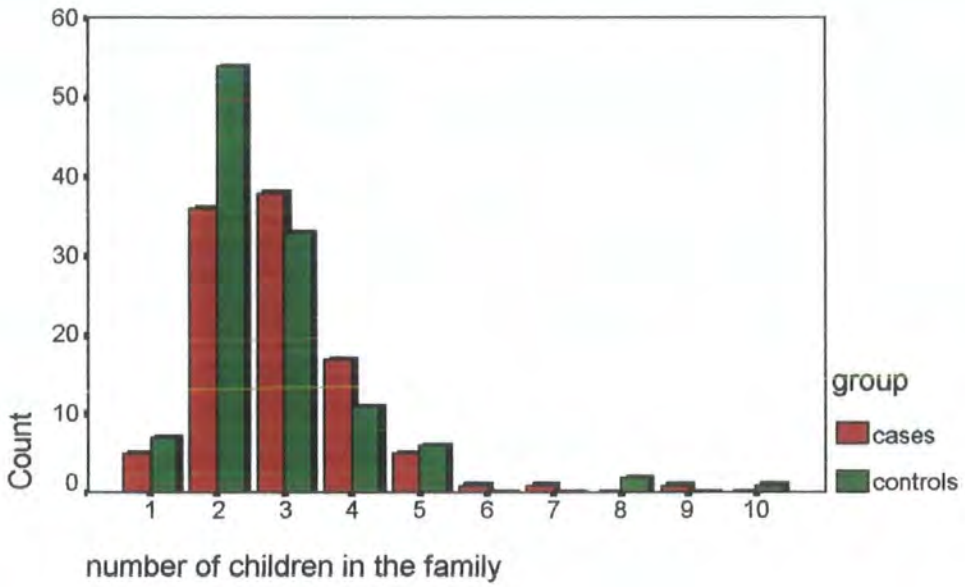
A between group comparison of family size was carried out. There were more siblings in the families of cases than controls, with a median of three for cases and two for controls. The data were skewed and so a non-parametric test was used to test for a difference. Family size did not significantly differ between groups (Mann Whitney U = 1.88, $p < 0.06$). A histogram of these data is shown in Figure 3.2.11a.

As the median number of children in the case families was larger, it was expected that the cases would also be later born. The median for birth order was second for both groups. When the numbers of children who were first born were compared with those later born by group there was no statistically significant difference (Mann Whitney U = 1.53, p = 0.12). These data are shown as a histogram in Figure 3.2.11b.

The number of days between births of all siblings was calculated. The difference between groups in mean number of days between births of siblings is shown in Table 3.2.11. None of the differences were found to be statistically significant. Three children with twins were studied, all of whom belonged in the control group.

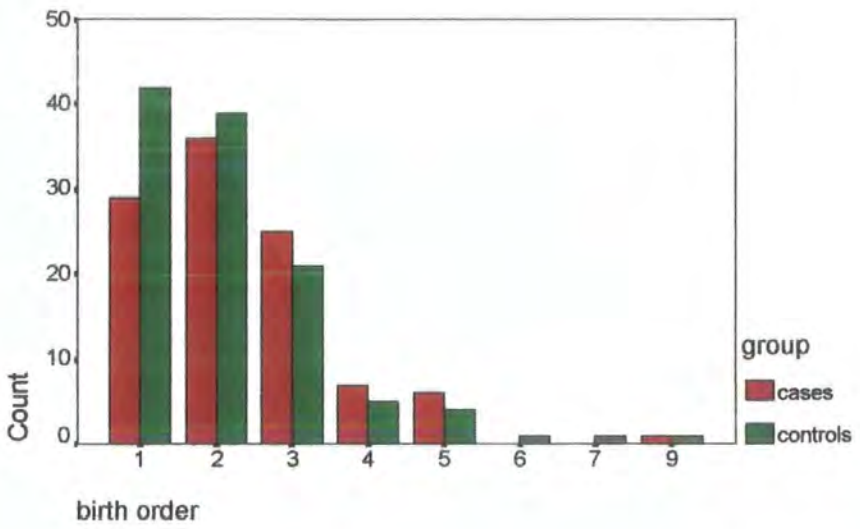
There was a trend for case families to be larger than controls, but cases were not significantly later born, nor was the spacing between births in the case group significantly shorter than controls.

Figure 3.2.11a Number of children in the families of case and control groups



Mann Whitney U = 1.88, $p < 0.06$

Figure 3.2.11b Order of birth of cases and controls



Mann Whitney U = 1.53, p = 0.12

Table 3.2.11 Number of days between birth of siblings

	Mean number of days between the birth of siblings by group				*Difference between groups in the mean number of days between the birth of siblings*		
	Cases	*n*	*Controls*	*n*	*Difference*	*t*	*p*
1st-2nd	1146	99	1077	105	69	0.72	0.524
2nd-3rd	1317	65	1342	52	25	0.15	0.877
3rd-4th	1096	26	1057	18	39	0.18	0.848
4th-5th	1071	8	1418	7	347	0.57	0.582
5th-6th	908	3	856	1	52	na*	na*
6th-7th	530	2	477	1	53	na*	na*

na* is small therefore no statistical analysis is reported

3.2.12 Measures of socio-economic deprivation

Three measures of socio-economic deprivation were used, levels of employment and car and home ownership.

In order to assess levels of employment in each household three questions were asked: whether the main carer (mostly the mother) was employed (Table 3.2.12a), whether this was part time or full time work (Table 3.2.12b) and whether there was another wage earner in the home (Table 3.2.12c). Mothers of cases were slightly less likely to be employed (52% of mothers of cases worked and 59% of controls). Of the mothers who worked, 22% of mothers of cases had a full time job compared with 25% of mothers of controls. Lastly cases were slightly more likely to have another wage earner in the house (64% vs 61%). None of these differences was statistically significant.

The families of cases were slightly more likely to rent, not own, their home (34.6% of cases and 31% of controls), and less likely to own a car (63% of cases and 66% of controls), but again the differences were small and not statistically significant. These data are shown in Tables 3.2.12d and 3.2.12e.

Table 3.2.12a Maternal employment by case and control group

Maternal employment	*Cases*	*Controls*	*Total*
Employed	54	67	121
Not employed	50	47	97
Total	104	114	218

$\chi^2 = 1.03$, df = 1, p = 0.3

Table 3.2.12b Full time or part time work of mothers employed

Full time or part time	*Cases*	*Controls*	*Total*
Full time	12	17	29
Part time	42	51	93
Total	54	68	122

$\chi^2 = 0.13$, df = 1, p = 0.72

Table 3.2.12c Presence of another wage earner in the home

Another wage earner in the home	*Cases*	*Controls*	*Total*
Yes	67	69	136
No	37	45	82
Total	104	114	218

$\chi^2 = 0.35$, df = 1, p = 0.55

Table 3.2.12d Rented and privately owned accommodation

Home ownership	*Cases*	*Controls*	*Total*
Council House	36	35	71
Privately rented	10	11	21
Owner occupier	58	68	126
Total	104	114	218

$\chi^2 = 0.40$, df = 2, p = 0.82

Table 3.2.12e Car ownership

Car ownership	*Cases*	*Controls*	*Total*
Yes	66	75	141
No	38	39	77
Total	104	114	218

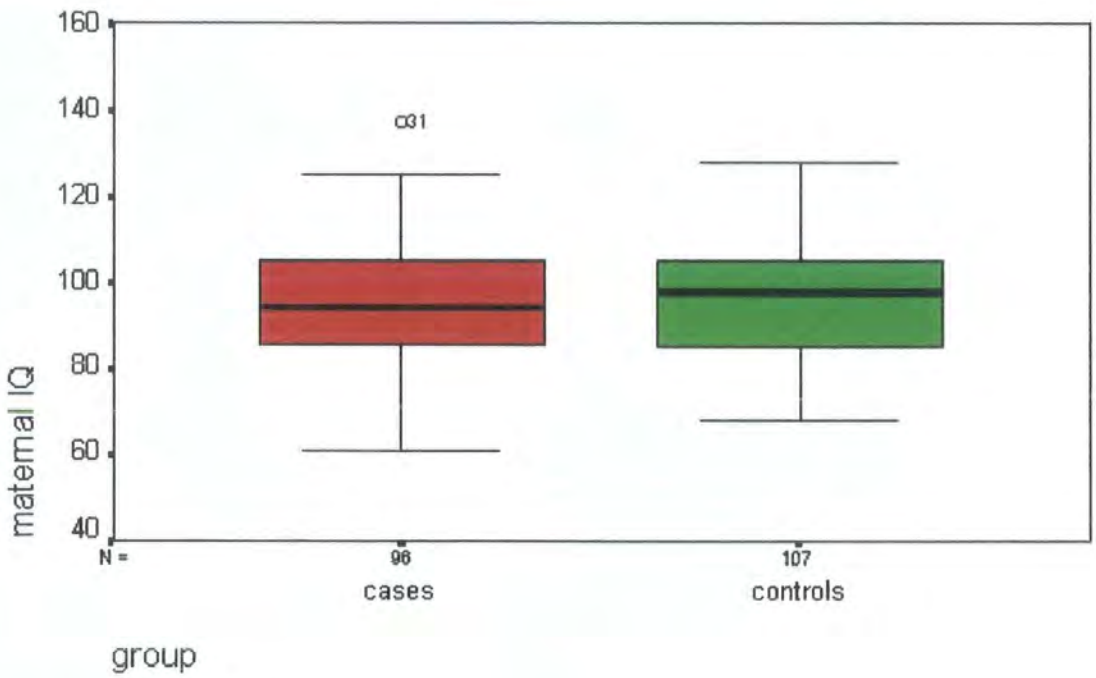
$\chi^2 = 0.13$, df = 1, p = 0.72

3.2.13 Maternal IQ

Mothers whose first language was not English, and two mothers whose children were too impaired to be tested with the WISC-III^{UK}, were not tested with the WAIS-R^{UK}. A further six mothers were not interviewed or tested after their child had participated in the study. Therefore, an IQ test was carried out on 96 mothers of cases and 107 mothers of controls.

connect the largest and smallest values that are not categorised as outliers (i.e. more than 1.5 box lengths from the box). Outliers (O) and their case numbers are also shown. The mean WAIS-R^{UK} score of mothers of cases was less than one IQ point lower than that of controls (96.3 vs 97.0). Since the data are normally distributed a t-test was used to test for a difference. The difference in mother's IQ scores by group was small and not statistically significant ($t = 0.37$, $df = _{1, 202}$, $p = 0.58$).

Figure 3.2.13 Maternal IQ for case and control groups



$t = 0.37$, $df = _{1, 202}$, $p = 0.58$

3.2.14 Summary of demographic variables and reported medical problems

Selection criteria for controls, age, sex and deprivation score of the area of residence at eighteen months of age, were based on the information which was available before the families had been contacted. At follow up we were able to obtain information about other factors which may influence psychological or anthropometric outcomes which could not have been known beforehand. For this reason additional information was gathered during the interview with the mother about demographic and educational variables and medical history, in order that adjustment could be made to the outcome measures where potentially confounding variables differed between the groups.

There were no statistically significant differences between the case and control groups for the age and sex of children followed up, their care arrangements, method of feeding, visual, hearing and speech problems, referral for special educational needs assessment, preschool education, ethnic origin, family size, birthorder, birth spacing, socio-economic indices, their mother's education, and IQ. Statistically significant differences found between the case and control groups were mainly related to the medical history of each child. These differences were: having a difficult birth, reported feeding problems, and having a medical problem requiring hospital treatment. All children had a gestational age of 37 weeks or more (this was an inclusion criterion). Even so, there was still a significant between group difference in gestational age reported by the mother, control mothers reporting shorter gestational age for their child.

3.3.1 Medical history derived from medical records

The differences reported by mothers in the health related variables in the questionnaire required further investigation for two reasons. Firstly, the medical histories taken from the mothers were imprecise. Secondly, the difference in hospital admissions between groups,

although statistically significant, should be treated with caution as nine cases had been referred exclusively for growth problems or failure to thrive, and two exclusively for feeding problems. The objective in considering hospital admissions is to identify children who may have an organic condition which has contributed to poor growth rather than to confirm the condition which defines them as a case in the first place. For this reason information recorded on the interview schedule about a child's medical history from birth was abstracted by the author (Appendix. VIII) and coded blind by Dr Charlotte Wright according to the relevance of the reported condition to growth or cognitive outcomes. Initially, the reported medical history for each child was coded in one of three ways: no admission, records not consulted as the hospital admission was unlikely to be significant, the admission was for a condition likely to be relevant to the outcomes. The medical records of those children with a condition likely to be relevant were then examined, and recoded as either not relevant to the outcomes, possibly relevant, or definitely relevant as described in the methods section (pp 72, Tables 2.5.2a and 2.5.2b). In four instances the medical records (all for cases) were not available.

The codes given for each child's medical history according to its likely relevance to growth are shown in Table 3.3.1a. More case children had been referred to hospital and more had conditions likely to affect growth. Three children (two cases and one control) had conditions which would definitely affect their growth. As some expected cell sizes are less than 5, categories 1 and 2, and 4 and 5 were collapsed for statistical analysis. The difference between the groups was statistically significant ($\chi^2 = 12.36$, df = 3, p = 0.006).

Similarly medical conditions reported by the mother were coded according to the likelihood that a condition would lead to cognitive impairment; the criteria are listed in the methods section (pp 72, Table 2.5.2b). Eight children were coded as having such a condition, including the children who were too impaired to be tested with the WISC-III^{UK} and WORD.

The results by group of the coding for medical conditions likely to impair cognitive development are shown on Table 3.3.1b. All the children coded as having a major condition which would definitely cause cognitive impairment were in the case group. For the purposes of statistical analysis the first two categories were collapsed as expected cell values were less than five. Disproportionately more cases had a medical condition likely to impair cognition (odds ratio = 3.7). As an expected cell value is less than five, Fisher exact results are used to test for a difference. The difference failed to reach conventional levels of statistical significance, (Fisher exact test, 1-tailed p = 0.11) but numbers are small.

Description of the conditions found in medical records rated as likely to have an effect on growth or cognitive outcomes are given in Table 3.3.1c and 3.3.1d by case then control group.

Of the 66 cases who had attended hospital or outpatients departments, 12 were rated blind as having an organic condition which affected growth or cognitive outcomes, and of the 56 controls 4 were. Six children were rated both as having an organic condition likely to result in poor growth, and to have a condition likely to result in cognitive impairment. All but one of these were cases. Therefore, not only were the cases with an organic condition more numerous than controls, but the conditions described for the cases were more serious and more likely to be associated with growth and cognitive outcomes. Only one case and one control were coded as having an organic condition which was likely to affect cognitive outcomes but not growth.

Table 3.3.1a Classification of medical condition related to growth

Classification	*Cases*	*Controls*	*Total*
1) Definite	2	1	3
2) Possible	9	2	11
Total	***11***	***3***	***14***
3) Irrelevant	16	19	35
4) No records	4	0	4
5) Not consulted	57	56	113
6) No admission	16	36	52
Total	***93***	***111***	***204***

Categories 1 & 2 and 4 & 5 were collapsed for analysis as expected cell size is < 5
$\chi^2 = 12.36$, df = 3, p = 0.006

Table 3.3.1b Classification of medical condition related to cognitive outcomes

Classification	*Cases*	*Controls*	*Total*
Definite	4	0	4
Possible	2	2	4
Total	***6***	***2***	***8***
No history	98	112	210
Total	***104***	***114***	***218***

Fisher exact test, 1-tailed p = 0.11

Table 3.3.1c Cases rated as having a medical condition which would affect growth or cognition

Cases	*Medical condition*	*Code for medical history likely to affect growth*	*Code for medical history likely to affect cognition*
Case 1	A boy who had septicaemia at birth and has mild cerebral palsy	2 (possible)	1 (definite)
Case 2	A girl who had recurrent urinary tract infections leading to renal scarring by three years old and reflux on one kidney	2 (possible)	3 (not relevant)
Case 3	A boy admitted at 5 months for bronchiolitis and subsequently for a persistent wheeze	2 (possible)	3 (not relevant)
Case 4	A boy admitted for loss of consciousness following a head injury at one year old	3 (not relevant)	2 (possible)
Case 5	A boy with Di George Syndrome/Catch 22, a cleft palate and chromosomal anomaly.	2 (possible)	1 (definite)
Case 6	A boy who had been admitted for surgical correction of a hiatus hernia at 15 months of age, also suffered from recurrent chest infections (croup, pneumonia, pleurisy,) and jaundice	2 (possible)	3 (not relevant)
Case 7	A girl with perinatal hypoxia was subsequently found to have microcephaly and an attention disorder. She had also been found to have mild pulmonary stenosis.	2 (possible)	1 (definite)
Case 8	This girl was found to have a major brain malformation and hydrocephalus.	1 (definite)	1 (definite)
Case 9	This girl was treated for a major scald at 13 months. At 15 months grommets were inserted following recurrent ear infections	2 (possible)	3 (not relevant)
Case 10	A boy with Pierre Robins Syndrome, a pharyngeal anomaly and a cleft palate who has had surgery several times including a tracheostomy and gastrostomy.	1 (definite)	2 (possible)
Case 11	A boy admitted for recurrent urinary tract infections at one year. Found to have renal scarring	2 (possible)	3 (not relevant)
Case 12	A boy admitted with osteomyelitis at the age of nine months	2 (possible)	3 (not relevant)

Table 3.3.1d Controls rated as having a medical condition which could affect growth or cognition

Controls	*Medical condition*	*Code for medical history likely to affect growth*	*Code for medical history likely to affect cognition*
Control 1	A boy with Tetralogy of Fallot and progressive cyanosis who underwent open heart surgery at two years of age	1 (definite)	2 (possible)
Control 2	A girl admitted at one year with frequent ear infections. Grommets were inserted when aged two and a half years. Admission for moderately severe asthma when nearly four years old.	2 (possible)	3 (not relevant)
Control 3	A girl admitted for treatment of epilepsy at 8 months following at least five fits	3 (not relevant)	2 (possible)
Control 4	A boy who had grommets inserted at 14 months of age and was still receiving treatment at 8 years after multiple middle ear infections	2 (possible)	3 (not relevant)

3.4.1 Stratified analysis of cases and controls with and without an organic condition and variables likely to affect IQ test performance

An analysis was carried out to investigate the extent to which ratings of relevant medical problems were associated with other problems reported by the mothers. For the purpose of this analysis, only the children coded as having a medical condition likely to affect growth were classified as 'organic'. This category in fact also includes all those children who had a condition which definitely affected cognitive outcomes, since these children all also had a condition which affected growth. It excludes one case who had a head injury and a control who was treated for mild epilepsy. In both instances they did not have a medical condition likely to affect growth or with more than a possible effect on cognition. These and all other cases are referred to as 'non-organic'.

3.4.2 Feeding problems and medical history

Of the eleven cases and three controls classified as 'organic', nine cases and one control were reported to have feeding problems. Of the non-organic cases, 43% were reported as having feeding problems in comparison with only 24% of controls. This difference in the non-organic cases was statistically significant ($\chi^2 = 8.01$, df = 1, p = 0.005). These data are shown in Table 3.4.2.

Table 3.4.2 Reported feeding problems in cases and controls classified as 'organic' or 'non-organic'

	Organic		*Non-organic*		
Feeding problems	*Cases*	*Controls*	*Cases*	*Controls*	*Total*
Yes	9	1	40	27	77
No	2	2	53	84	141
Total	11	3	93	111	218

For non-organic cases and controls $\chi^2 = 8.01$, df = 1, p = 0.005

3.4.3 Speech problems and medical history

Only 15% of the non-organic cases were reported as having speech problems; this was slightly fewer than the 18% of the control group ($\chi^2 = 0.32$, df = 1, p =0.57). These data are shown in Table 3.4.3a.

Of the children rated as 'organic' who received speech therapy for more than a year, all six were cases. There were no significant differences between cases and controls for children receiving speech therapy after the organic cases were omitted ($\chi^2 = 0.15$, df = 1, p = 0.7). These data are shown in Table 3.4.3b.

Table 3.4.3a Reported speech problems in cases and controls classified as 'organic' and 'non-organic'

	Organic		*Non- organic*		
Speech Problems	*Cases*	*Controls*	*Cases*	*Controls*	*Total*
Yes	7	2	14	20	43
No	4	1	79	91	175
Total	11	3	93	111	218

For non-organic cases and controls $\chi^2 = 0.32$, df = 1, p = 0.57

Table 3.4.3b Case and control children receiving speech therapy classified as 'organic' and 'non-organic'

	Organic		*Non- organic*		
Speech Therapy	*Cases*	*Controls*	*Cases*	*Controls*	*Total*
Less than a year	1	2	7	7	17
More than a year	6	0	2	3	11
Total	7	2	9	10	28

For non-organic cases and controls $\chi^2 = 0.15$, df = 1, p = 0.7

3.4.4 Special educational needs and medical history

Seven cases and one control rated as having an organic condition had special educational needs. After excluding the children with an organic condition 8.6% of cases and 9% of controls had special educational needs ($\chi^2 = 0.01$, df = 1, p = 0.92). These data are shown in Table 3.4.4a

Nine Cases and two controls had reached stage 3 & 4 and had been assessed by an educational Psychologist for special educational needs. Of these, seven were cases and none were controls with an organic condition. There were no statistically significant differences in the assessment stages reached for non-organic cases and controls ($\chi^2 = 0.15$, df = 1, p = 0.7). These data are shown in Table 3.4.4b

Table 3.4.4a Case and control children assessed for special educational needs classified as 'organic' and 'non-organic'

	Organic		*Non- organic*		
Special educational needs	*Cases*	*Controls*	*Cases*	*Controls*	*Total*
Yes	7	1	8	10	26
No	4	2	85	101	192
Total	11	3	93	111	218

For non-organic cases and controls $\chi^2 = 0.01$, df = 1, p = 0.92

Table 3.4.4b Case and control children on the register of special educational needs by assessment stage and classified as 'organic' and 'non-organic'

	Organic		*Non-organic*		
Assessment stage	*Cases*	*Controls*	*Cases*	*Controls*	*Total*
Stage 3 & 4	7	0	2	2	11
Stage 1 & 2		1	6	8	15
Total	7	1	8	10	26

For non-organic cases and controls $\chi^2 = 0.15$, df = 1, p = 0.7

3.4.5 Summary of medical history and the variables likely to affect IQ test performance

On assessment of their medical history, more cases were likely to have a medical problem which would be likely to affect growth and cognition. Once those children with an organic condition had been identified, the differences between case and control groups in the severity of speech and educational problems were better explained in terms of organic disease rather than failure to thrive. Only feeding problems remained disproportionately high in non-organic cases. Since feeding problems are widely reported in children who fail to thrive (i.e. Pollitt and Eichler, 1976, Dahl and Kristiansson, 1987, Mathisen et al, 1989, Ramsay et al, 1993), they may be properly regarded as symptomatic of the condition, and would be expected in a sample of children correctly identifed as failing to thrive.

Further differences between the groups were birth weight and gestational age, the control group having a slightly lower birth weight and shorter gestational age than cases. The effect of an organic condition on gestational age was unlikely to be large, as only four

controls were classified as having an organic condition. It is also unclear whether variation in gestational age after 37 weeks gestation has any clinical significance since all births were at term. However, since the difference is nevertheless statistically significant, adjustment for the potential effects of gestational age on outcome is taken into account in subsequent analyses, along with birth weight and organic condition, the only other potentially confounding differences between the groups.

Chapter Four

Anthropometric measures

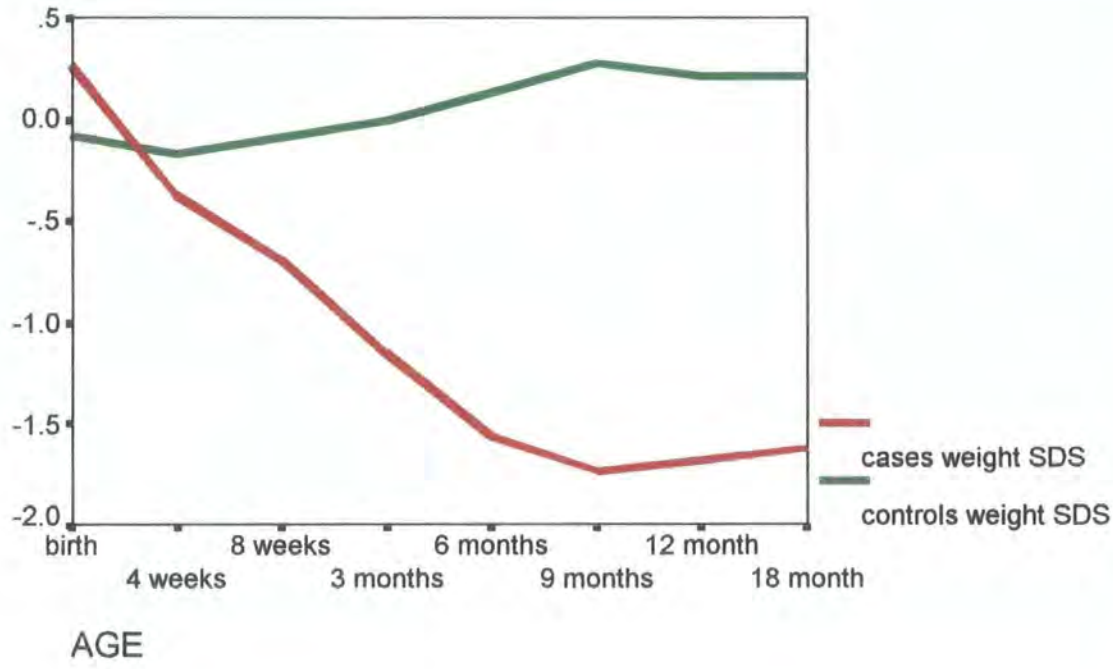
4.1.1 Weight in infancy

Birth weights were available from child health records for 97.2% of the children selected. Average birth weight for the cases was slightly higher than that of the controls. Mean birth weight for the cases was 0.27 of a standard deviation above the reference population mean, whilst for controls it was -0.14 of a standard deviation below. This difference was statistically significant at the p = 0.004 level. This early weight advantage for cases was rapidly lost, as, by four weeks old, cases weighed less than controls, and thereafter the mean weight for the cases was significantly less than for controls at every age, as shown on Table 4.1.1a and Figure 4.1.1a.

Table 4.1.1a Weight standard deviation scores (SDS) for case and control groups in each age band.

	Cases			*Controls*				
Age	*mean SDS*	*sd*	*n*	*mean SDS*	*sd*	*n*	*t*	*p*
Birth	0.28	1.20	133	-0.14	1.14	131	2.87	0.004
4 weeks	-0.37	1.06	124	-0.20	0.96	120	1.38	0.170
8 weeks	-0.69	1.06	116	-0.13	0.98	113	4.17	0.000
3 months	-1.18	0.98	129	-0.02	0.94	130	9.76	0.000
6 months	-1.56	0.86	129	0.13	0.99	125	14.60	0.000
9 months	-1.75	0.83	117	0.28	0.99	96	16.20	0.000
12 months	-1.69	0.83	112	0.22	0.93	103	16.00	0.000
18 months	-1.62	0.90	93	0.20	1.01	76	12.35	0.000

Figure 4.1.1a Mean weight standard deviation scores for case and control groups in each age band.



It is not immediately clear from Figure 4.1.1a at what age the largest average fall in relative weight occurred. In order to show this better, the change in standard deviation scores between two time points for cases and controls were compared by calculating the difference in SDS scores between a later age band and the preceding one (i.e. $SDS_{age2} - SDS_{age1}$). These data are shown in Table 4.1.1b and Figure 4.1.1b.

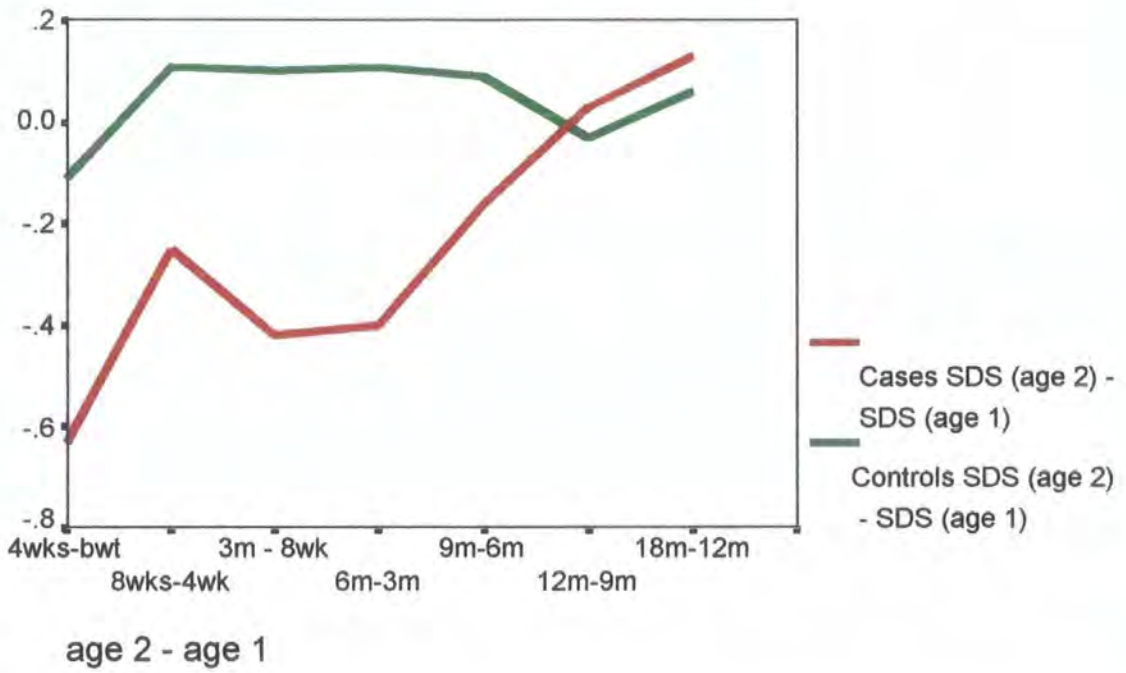
The mean difference in relative weight of the control group from one age band to the next was very small and in fact the control goup was gaining relative to the reference population. The case group, on the other hand, showed a substantial fall in relative weight

compared with the control group. Although this method of analysis does not allow for regression to the mean, the differences show that on average the fall in relative weight occurred early, there being no statistically significant differences between the groups in the change in relative weight after nine to twelve months, when it appears that the case infants have stabilised at a new lower weight.

Table 4.1.1b Difference in mean weight standard deviation scores between one age band and the previous age band by case and control group

	Cases			*Controls*				
Difference	*mean*	*sd*	*n*	*mean*	*sd*	*n*	*t*	*p*
Birth -4wks	-0.64	0.69	121	-0.11	0.62	117	6.18	0.000
4wks - 8wks	-0.26	0.44	111	0.13	0.34	104	7.23	0.000
8wks - 3mths	-0.44	0.39	112	0.11	0.36	111	10.98	0.000
3mths - 6mths	-0.37	0.49	124	0.13	0.50	123	8.07	0.000
6mths - 9mths	-0.17	0.58	111	0.09	0.42	91	3.60	0.0004
9mths - 12mths	0.02	0.53	97	-0.03	0.37	81	0.77	0.440
12mths - 18mths	0.14	0.49	81	0.04	0.56	64	1.12	0.260

Figure 4.1.1b Difference in mean weight standard deviation scores between one age band and the previous age band by case and control group



Thrive Index values were calculated for each age band from 3 months old (3, 6, 9, 12, 18 months), but not all infants had a weight recorded within every age band. The cases, who had been weighed more often, had a mean of 4.2 Thrive Index values and the controls had 3.9. Sixty five cases had fallen below the 5% threshold in two age bands, 46 in three age bands, 18 in four age bands and 7 in five age bands; thus no case had failed to thrive for less than three months in infancy, the majority failing to thrive for six months or more. One of the selection criteria for controls was that they had maintained a relative weight within one standard deviation of their predicted weight SDS, so none of the controls had fallen below the 5% threshold for weight gain in any of the age bands.

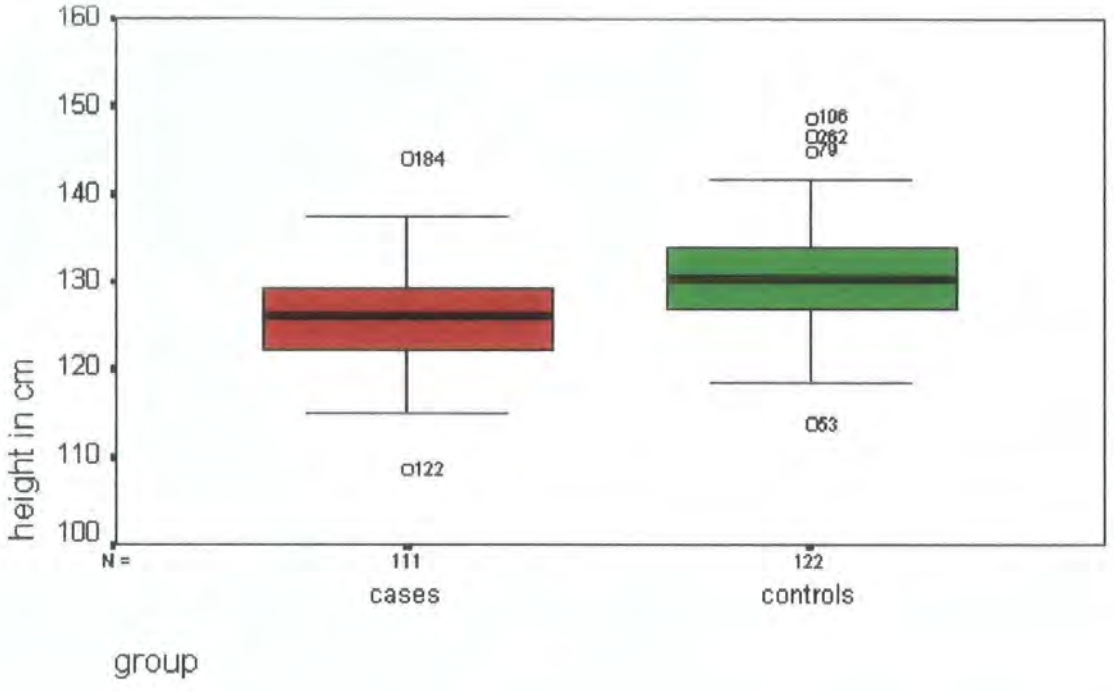
The median age at which cases first crossed the 5% screening threshold was 5.0 months (the interquartile range was 3.11 months to 8.5 months), although, as shown, the mean relative weight of the cases was clearly falling from birth. The fall in relative weight of cases was sustained and at times severe. For cases the mean deficit in weight, calculated as the mean of Thrive Index values, was -1.38 (SD 0.38). The controls, on the other hand had gained in relation to their predicted weight standard deviation score as their mean Thrive Index value was 0.24 (SD 0.58). The comparative severity of failure to thrive in the case group is evident when the mean of the lowest Thrive Index values are compared with those of the controls. The mean for the lowest Thrive Index values recorded for cases was -1.87 (SD 0.51), and for controls was -0.15 (SD 0.53).

4.2.1 Anthropometric measures at eight years old

The anthropometric measures for cases and controls were compared using an independent samples t-test. They are represented by boxplots showing the median, interquartile range, and the range of scores.

When they were eight years old, the height of cases ranged from 108.5 cm to 144 cm (median 126 cm) and that of controls from 113.8 to 148.6 (median 130 cm). The data for height are shown on Figure 4.2.1a. Since height is normally distributed, the mean values of 125.96 cm (SD 5.6) and 130.74 cm (SD 5.9) for cases and controls are close to the median values. On average cases were 4.78 cm shorter than that of controls and this was a statistically significant difference ($t = 6.33$, $df = {}_{1, 232}$, $p < 0.001$). When compared with a reference population the average height of cases was -0.8 standard deviations below the mean and that of controls was 0.02 standard deviations above, close to the reference population mean.

4.2.1a Height at age eight of cases and controls



A possible explanation for the difference in height between the groups is that the height of the parents differed. Tanner (1970) constructed standards for children's heights correcting for the effect of the height of parents using correlation coefficients derived from a number of studies of children's and parent's height coordinated by the International Children's Centre. Each parent was found to contribute equally to the height of their offspring. At age eight the correlation coefficient for the height of children and the mean of their parent's height was 0.53 for boys and 0.49 for girls.

Since a child's height is related to the height of both parents, it was important to establish if there were between group differences in the height of the parents. Parental height was compared in the case and control groups, mother's height having been measured and father's height reported by the mothers who were interviewed. The mean height (161.6 cm, SD 6.3 cm) for the mothers of cases was shorter than for the mothers of controls

(162.8 cm, sd 6.4 cm). This difference was not statistically significant (two tailed test, t = 1.43, df = $_{1, 214}$, p = 0.10). The reported mean height of fathers of cases was less than the fathers of controls (mean height 174.2ms, SD 7.8cms, and 177.0, SD 7.2 respectively). This was a statistically significant difference (two tailed test t = 2.69, df = $_{1, 211}$, p = 0.01).

In order to control for the height of their fathers and mothers, an analysis using regression methods was carried out. As the time between collecting height data for parents and children differed, mid parental height was calculated and converted to standard deviation scores and entered into the regression of children's height standard deviation scores at follow up. A dummy variable denoting case or control group was also created and entered into the regression. After adjusting for parental height the difference between the groups was 0.67 of a standard deviation and remained highly significant (t = 5.63, p < 0.0001). The regression of children's height and parental heights is shown in Table 4.2.1.

Table 4.2.1 Case and control children's height SDS aged 8 years, adjusted for mid-parental height SDS

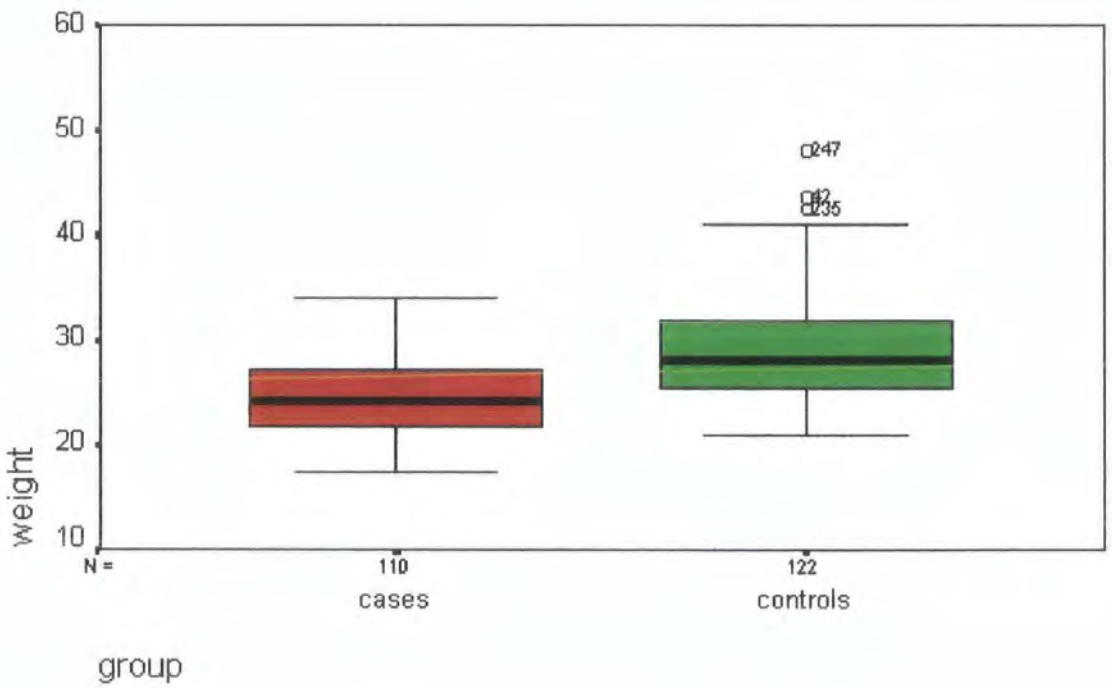
Variable	*Coefficient*	*Standard error*	*t*	*p*
Mid-parent SDS	0.48	0.07	7.2	<0.000
Group	-0.67	0.12	5.6	<0.000
Constant	0.11	0.08		

R = 0.57, df = $_{2, 210}$, F = 50.48, p < 0.0000
Group (0 = control; 1 = case)

The median is the preferred measure of central tendency for weight, as weight is not normally distributed (Cole, 1990). The median weight of cases (23.8 kg) was 4.1 kg lighter than that of controls (27.9 kg), with the standard interquartile range 21.5 kg to 26.9 kg and 25.2 kg to 31.6 kg respectively. The data are shown in Figure 4.2.1b.

Weights, adjusting for skewed distribution using the LMS method (Cole, 1990), were converted into SDS scores derived using The British Standards (Freeman et al, 1995). The mean of SDS score for cases was on -0.93 (SD 0.97) and that of controls was 0.16 (SD 0.96). This was a statistically significant difference ($t = 8.63$, $df = {}_{1, 231}$, $p < 0.001$).

Figure 4.2.1b Weight at age eight of cases and controls



As the cases were both shorter and lighter it was important to establish whether the difference in weight was proportionate to that in height, in other words to test if cases were thinner than controls. Body Mass Index (BMI) is a measure of weight adjusted for height (weight (kg) / height (m^2)) for which newly updated age-related population standards exist (Cole et al, 1995). As with weight, the distribution of BMI is not normal and so adjustment was made for skewed distribution using the LMS method (Cole, 1990), and BMI was converted into SDS scores derived using Body Mass Index reference curves for the UK (Cole et al, 1995). Cases had lower body mass indices than controls. Mean BMI SDS for cases was -0.65 (SD 0.87) and for controls was 0.2 (SD 1.00). This difference was statistically significant ($t = 6.87$, $df = {}_{1, 231}$, $p < 0.001$). Median BMI for cases and controls (14.9 and 16.3 respectively), the standard interquartile range of 14.1 to 16.0 for cases and 15.3 to 17.8 for controls and the range are shown in Figure 4.2.1c.

The data for head circumference were normally distributed, with mean head circumference of 51.9 cm (SD 1.8) in cases and 52.8 cm (SD 1.7) in controls. Using an independent samples t-test, this difference was statistically significant ($t = 3.86$, $df = {}_{1, 229}$, $p < 0.001$). The median (51.9 cm for cases and 53.0 cm for controls), the interquartile range (50.9 cm to 53.0 cm for cases and 51.9 cm to 54.0 cm for controls) and range (47.3 cm to 57.3 cm for cases and 47.0 cm to 56.5 cm for controls) of head circumference measures are shown in Figure 4.2.1d.

Figure 4.2.1c Body Mass Index at age eight of cases and controls

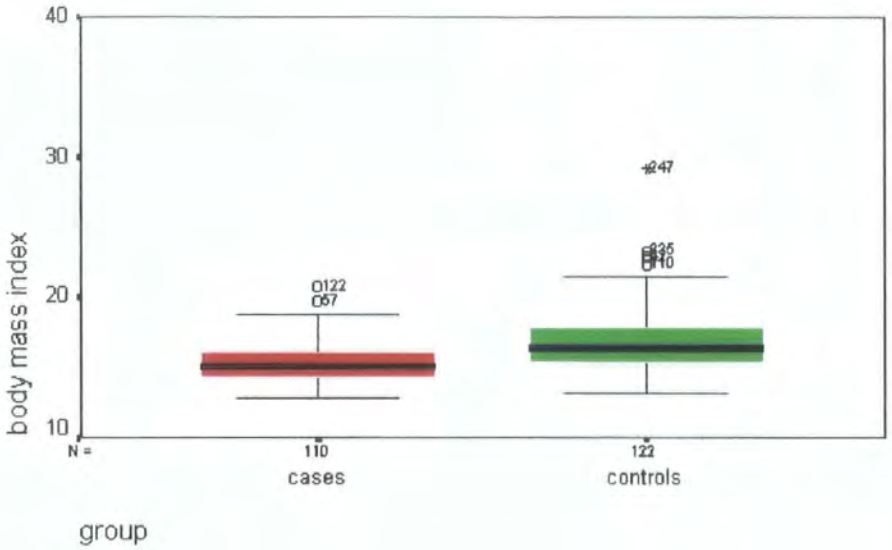
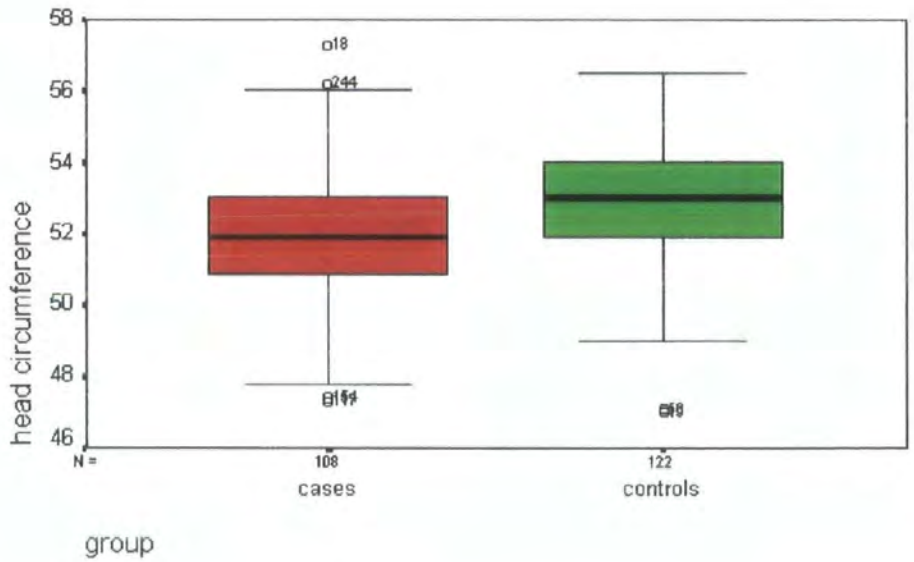


Figure 4.2.1d Head circumference at age eight for cases and controls



4.3.1 Summary of infant weight data and anthropometric measurements at age eight years

The screening criterion, the Thrive Index, identified a group of children whose early weight gain substantially differed from controls. Despite a slight birth weight advantage, when compared with controls, the cases showed a statistically significant deficit in weight standard deviation scores from 8 weeks of age to 18 months. By calculating the difference between a standard deviation score in one age band and that in the preceding age band, it could be seen that on average, the relative weight of cases was falling from birth, although they did not cross the 5% screening threshold until aged 5.0 months. After 6 months the decline in relative weight slowed down, although deficits in relative weight persisted up to 12 months of age. After 12 months of age there was no further fall. By contrast, despite a slightly lower birth weight, the controls' rate of weight gain in infancy was marginally greater than in the reference population.

At follow up at eight years of age, cases were substantially shorter than controls even after adjusting for their parent's shorter stature. The cases were also lighter, which would be expected as weight summarises a number of anthropometric measures including height. But weight was also less after adjusting for height using Body Mass Index, so cases were thinner than controls. Their mean head circumference was also significantly smaller than that of controls. The cases were smaller than the controls, but they were also well below the reference population mean for three indices, height, weight and BMI. At follow up, the controls were slightly above the 50th percentile for height, weight and for BMI. So controls gained weight in infancy at an above average rate, and were slightly heavier than average at follow up.

Chapter Five

Psychological Outcomes

5.1.1 Intelligence in case and control groups

This analysis sets out to test the central hypothesis that failure to thrive is associated with cognitive deficits. The total number of children tested with the WISC-III^{UK} was 222 as it was not possible to test two cases who were attending special school. Of those tested 105 were cases and 117 were controls.

The WISC-III^{UK} Full Scale scores for the case and control groups were normally distributed and the variances (304.3 for cases and 292.8 for controls) were similar. The range of scores for both groups was large (45-130 IQ points for cases, 45-136 IQ points for controls) with median values of 85 and 88 IQ points for cases and controls respectively (shown in Figure 5.1.1a). The mean of IQ scores for the case group was 87.6, 3 points lower than the mean for controls (90.6 IQ points). This difference was not statistically significant ($t = 1.32$, $df = _{1, 221}$, $p = 0.19$).

Although no significant difference was found in Full Scale IQ, a separate analysis of verbal and performance sub scale scores was carried out, as previous studies found a preponderance of verbal IQ deficits in case groups (Hufton and Oates, 1977, Oates et al, 1984, 1985).

The WISC-III^{UK} Verbal scores for the case and control groups were normally distributed and the variances (298.5 for cases and 288.9 for controls) were similar. The range of scores for both groups was 50-130 points for cases and 50-139 points for controls, with median values of 87 and 89 points for cases and controls respectively (shown in Figure 5.1.1b). The mean of verbal IQ scores for the case group was 87.9 and the mean for controls was 91.2. This difference was not statistically significant ($t = 1.43$, $df = _{1, 221}$, $p = 0.15$).

The WISC-III^{UK} Performance scores for the case and control groups were also normally distributed and the variances similar (318.1 for cases and 282.8 for controls). The range of scores was 47-130 IQ points for cases and 47-133 IQ points for controls, with median values of 90 and 91 points for cases and controls respectively (shown in Figure 5.1.1c). The mean Performance score for the case group was 89.9 and the mean for controls was 92.2. This difference was not statistically significant ($t = 0.98$, $df = {}_{1, 221}$, $p = 0.67$).

Only one statistically significant difference was found in an analysis of the ten subtests. This was for the arithmetic subtest where the mean standardised scores for cases and controls were 8.3 and 9.4 respectively (SD 3.1 for both). This difference was found to be significant at the $p = 0.008$ level ($t = 2.67$, $df = {}_{1, 221}$). The results of an independent samples analysis for all subtests are in Appendix XII.

Figure 5.1.1a WISC-III^{UK} Full Scale Scores for cases and controls

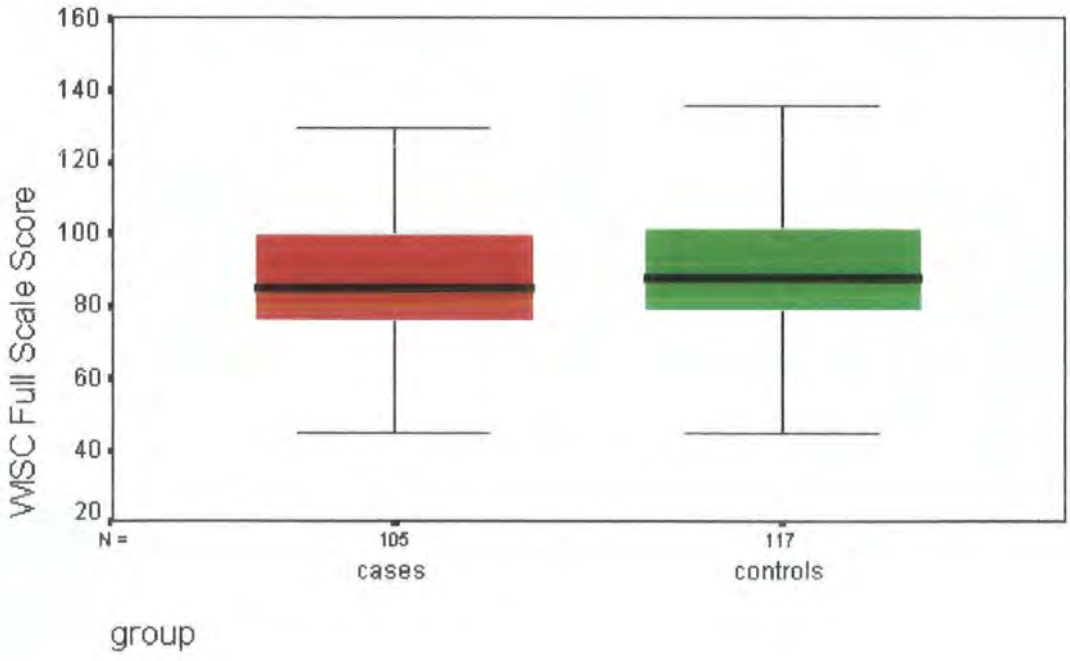


Figure 5.1.1b WISC-III^{UK} Verbal Scale Scores for cases and controls

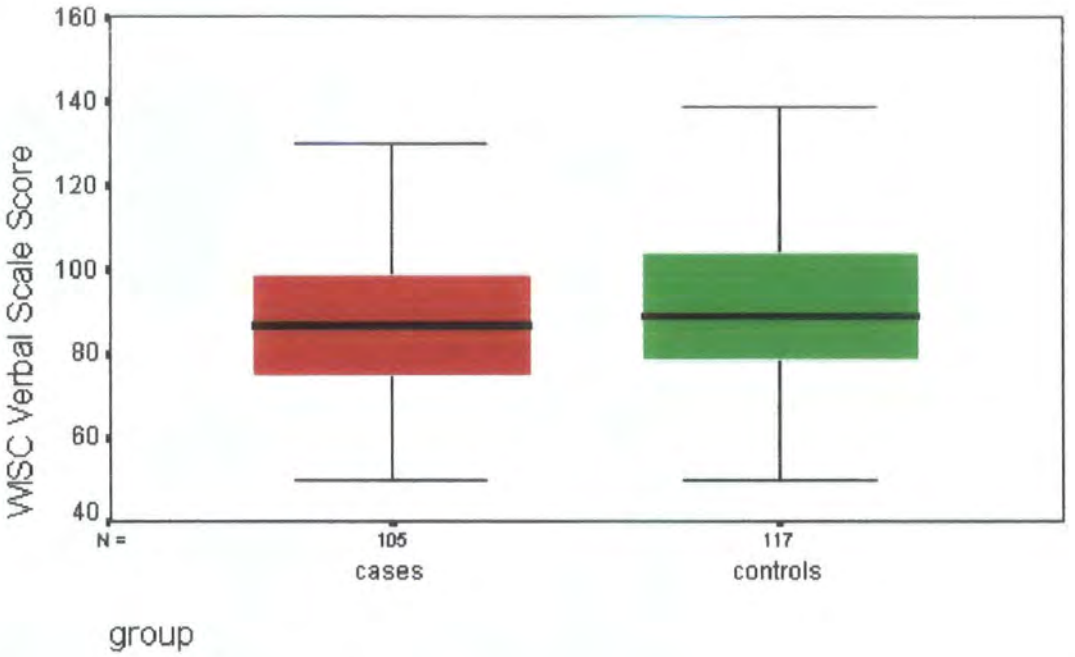
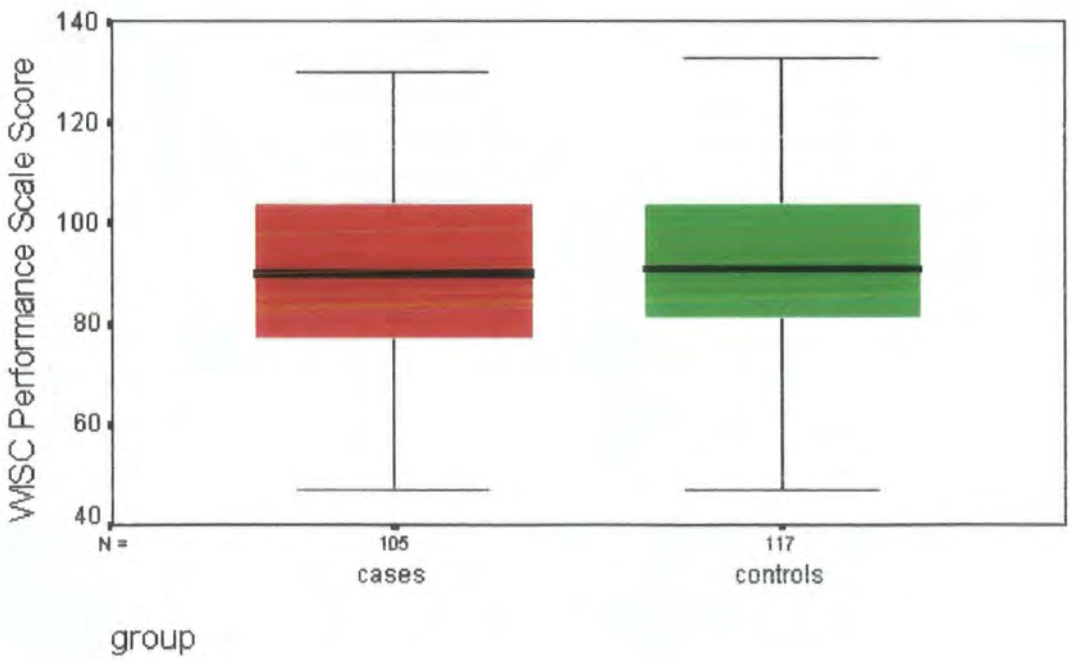


Figure 5.1.1c WISC-III^{UK} Performance Scale Scores for cases and controls

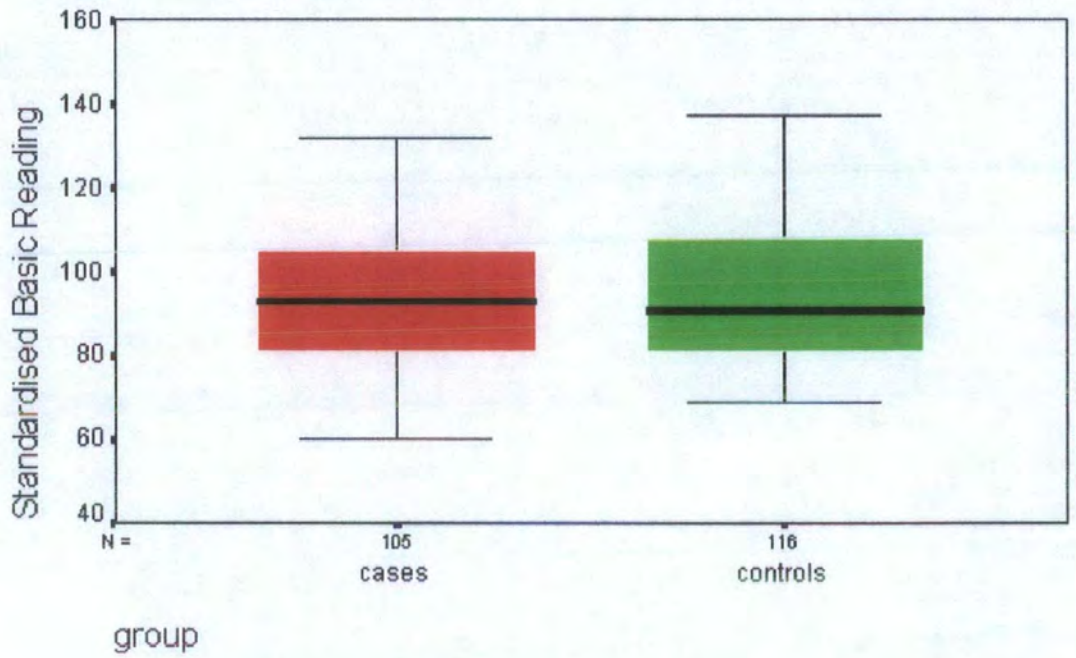


5.1.2 Reading in case and control groups

The WORD reading test was administered to 221 children, as it was not possible to test three children (one control child and two cases) attending special school.

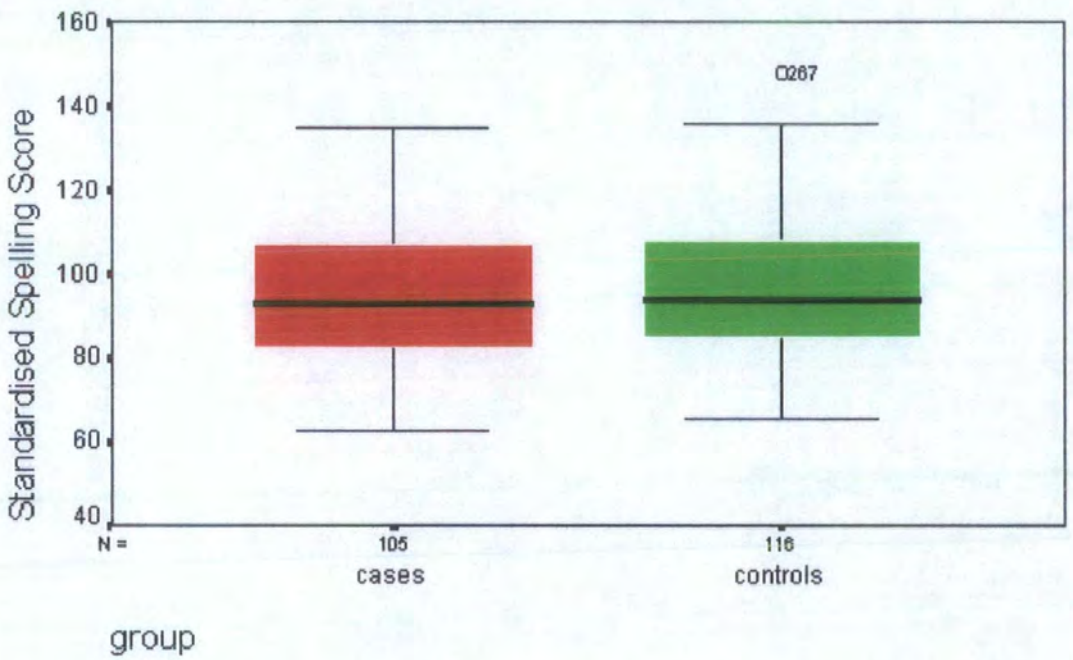
The Basic Reading scores for the case and control groups were normally distributed and the variances (263.1 for cases and 243.4 for controls) were similar. The range of standardised scores was 60 to 132 for cases and 69 to 137 for controls, with a median of 93 for cases and 91 for controls (shown in Figure 5.1.2a). The mean for the case group was 93.5 and the mean for controls was 94.5. The difference was less than one point and was not statistically significant ($t = 0.2$, $df = {}_{1, 220}$, $p = 0.65$).

Figure 5.1.2a Standardised basic reading scores for cases and controls



The Spelling scores for the case and control groups were also normally distributed and the variances similar (273.6 for cases and 260.1 for controls). The range of standardised spelling scores was 63 to 135 for cases and 66 to 148 for controls, with a median value of 93 for cases and 94 for controls (shown in Figure 5.1.2b). The mean value for standardised spelling scores for the case group was 94.7 and the mean for controls was 96.9. The difference was two points and was not statistically significant ($t = 0.99$, $df = {}_{1, 221}$, $p = 0.67$).

Figure 5.1.2b Standardised spelling scores for cases and controls

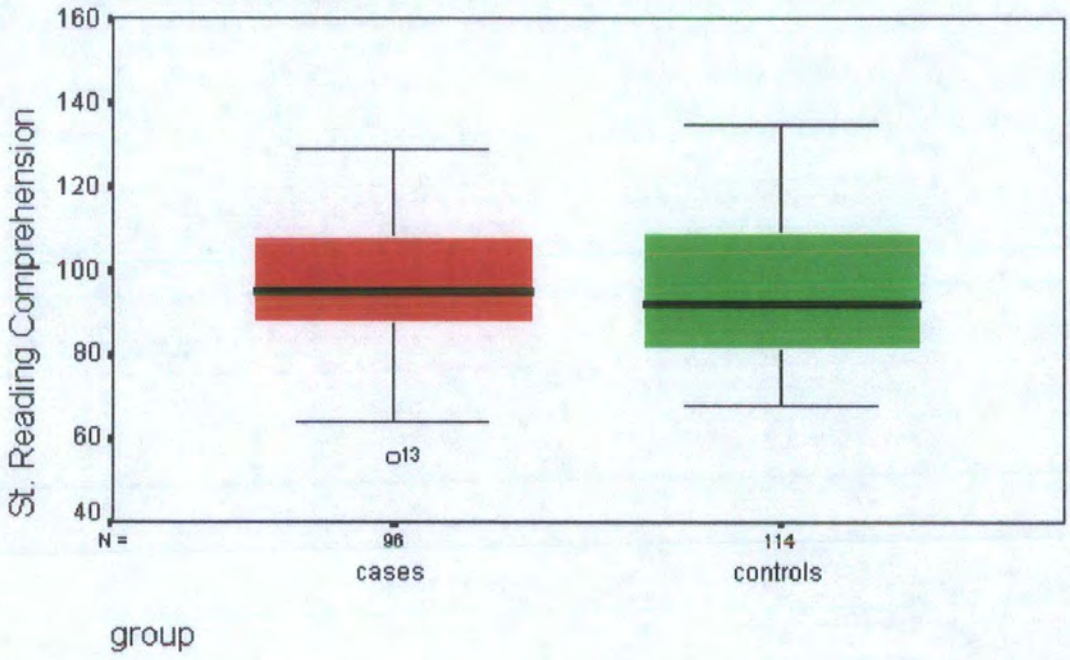


The Reading Comprehension test was only administered if a child received a raw score greater than eight on the Basic Reading test. Two controls and nine cases scored too low on Basic Reading and were not given the full test.

The Reading Comprehension test was completed by 96 cases and 114 controls. The scores for the case and control groups were normally distributed and the variances for

cases (232.6) and controls (271.4) were similar. The range of standardised reading comprehension scores were 56-129 for cases and 68-135 for controls, with a median value of 95 for cases and 92 for controls (shown in Figure 5.1.2c). The mean for the case group was 95.8 and the mean for controls was 95.3. The difference in the means was less than half a point (in favour of the cases) and not statistically significant ($t = 0.2$, $df = {}_{1, 209}$, $p = 0.8$).

Figure 5.1.2c Standardised reading comprehension scores for cases and controls



5.1.3 The relationship between WISC-III^{UK} Full Scale scores and reading ability

A comparison was made between children's IQ and the standardised reading scores. As the WORD was standardised on the same population as the WISC-III^{UK}, the manual for WORD provides an estimate of average standardised scores for Reading,

Spelling and Reading Comprehension predicted from WISC-III^{UK} Full scale scores. This comparison can test the findings of previous work which suggests that reading ability is poorer in children who have failed to thrive than would be expected given their IQ score (Oates, 1985).

The mean WISC-III^{UK} Full Scale score for the case group was 87.6 IQ points. According to the manual for WORD, average scores for basic reading ability, spelling and reading comprehension, conditional upon a WISC-III^{UK} Full Scale score of 88, are 93, 94 and 92 respectively. The actual mean standardised scores for cases of 93 for Basic Reading, and 95 for Spelling were close to these. The mean Reading Comprehension score of 96 was slightly higher than expected, but the poorest readers (n = 9) were not administered the Reading Comprehension test, as their scores for Basic Reading and Spelling were too low. As the mean WISC-III^{UK} Full Scale score of the nine poor readers was 66.5, this would account for the higher mean standardised Reading Comprehension score conditional upon a mean IQ calculated with those cases included.

The mean WISC-III^{UK} Full Scale score for controls was 90.6, from which Basic Reading, Spelling and Reading Comprehension scores of 95, 95 and 94 respectively, would be expected. Actual scores were 94, 97 and 95 for the three tests.

Therefore, standardised reading scores for both groups were close to those predicted from WISC-III^{UK} Full Scale score in a standardisation population.

5.1.4 Summary of between group comparison of IQ and reading ability

The IQ scores were below those of a standard population in both the case and control groups, as were the reading subscales. The means of standardised test scores for cases and controls were in the expected direction in that cases scored less than controls, with the exception of the Reading Comprehension standardised score. But these differences

were all small and none of them was statistically significant, except in the arithmetic subtest.

The means of the standardised scores for Basic Reading, Spelling and Reading Comprehension, predicted from the WISC-III^{UK} Full Scale scores, are close to the actual means found in the case and control groups. Therefore, reading scores were comparable to IQ scores and no specific reading deficit was identified.

However, of those children whose IQ scores were more than two standard deviations below the mean (i.e. 70 IQ points or less), 20 (19%) were cases and 9 (7.6%) were controls. Furthermore, cases who failed to score high enough on the Basic Reading test to enable them to receive the Reading Comprehension test outnumbered controls (9 cases and 2 controls).

Before concluding that no significant difference exists between the groups, further analysis of outcomes is needed which takes account of other differences between the two groups demographic make up, in particular the IQ of their mothers, their medical history birth weight and gestational age.

5.2.1 Further analysis of IQ

Mother's IQ is associated with that of their child (Skodak and Skeels, 1949, Defries et al, 1987, and Scarr et al, 1993). So, the strong correlation ($r = 0.66$, $p < 0.0005$) found between WISC-III^{UK} Full Scale scores and maternal IQ (WAIS-R scores) in this study was expected. Although there were no statistically significant differences in maternal IQ between groups, since maternal IQ explained 43% of the variance and was such an influential covariate, it was the first term to be entered into the regression.

A dummy variable was created to indicate the presence of an organic condition. Only three children (two cases and one control) had been rated as definitely having a major condition likely to affect growth and only two of these children (one case and one

control) had been administered an IQ and reading test. As there were only two eligible children in this category, they were included with those eleven children rated as possibly having an organic condition likely to affect growth. This 'organic' term was then entered into the regression in order to estimate mean IQ separately for those children classified as having an organic condition.

When adjustment was made for both maternal IQ and the presence of an organic condition, both covariates adjusted for each other, made a significant independent contribution to the explained variance of children's IQ. However, the difference between the mean IQ for the case and control groups after adjustment for maternal IQ and organic condition was just 1.56 IQ points and was not statistically significant. The regression is shown on Table 5.2.1a.

Table 5.2.1a The regression of children's IQ on maternal IQ, organic illness and group

Variable	*Coefficient*	*Standard error*	*t*	*p*
Maternal IQ	0.77	0.06	12.30	<0.0005
Organic	-7.98	3.8	2.09	<0.05
Group	-1.56	1.79	0.86	ns
Constant	17.56			

R = 0.67 df = 3, 199 F = 53.38 p = <0.001
Organic (0 = no; 1 = yes)
Group (0 = control; 1 = case)

Three other variables needed to be taken into account. The first was birth weight, as the difference in birth weight between cases and controls was statistically significant (cases heavier than controls). The second variable was gestational age as the mothers of controls, on average, reported a lower gestational age for their child than mothers of cases. All were term births but this reported difference was also statistically significant. The third variable was the number of siblings in the family. This difference

between groups was approaching statistical significance (p = 0.06) and so, this variable was also included.

Maternal IQ, organic condition, birth weight, gestational age and family size were entered into a regression of WISC-III^{UK} Full Scale scores on group. A stepwise regression was used, which has the advantage that covariates are ranked according to the amount of the variance accounted for by each when analysed in relation to one another; it eliminates those that do not make a significant contribution to the explained sum of squares. The results of this analysis are shown in Table 5.2.1b.

Table 5.2.1b Stepwise multiple regression of children's IQ on maternal IQ, organic illness, birthweight, gestational age, family size and group

Variables in the equation	*Coefficient*	*Standard error*	*t*	*p*
Maternal IQ	0.78	0.06	12.25	<0.0000
Organic	-9.44	3.75	2.5	0.01
Gestational age	1.81	0.78	2.3	0.02
Constant	-56.15			

R = 0.68 df = 3, 193 F = 55.45 p < 0.001
Organic (0 = no; 1 = yes)
Group (0 = control; 1 = case)

Variables not in the equation	*Coefficient*	*Standard error*	*t*	*p*
Group	-0.05	-0.06	0.91	0.37
Birth weight SDS	0.058	0.08	1.07	0.29
Family size	-0.03	-0.04	0.61	0.5

As expected, maternal IQ was the most important covariate of WISC-III^{UK} Full Scale scores. The mean IQ of those children with an organic condition was 9.4 IQ points less than those without. Gestational age as reported by the mother, surprisingly, made

a significant contribution to the explained variance of WISC-III^{UK} Full Scale Scores. Birth weight and family size were not significant covariates. The group variable remained statistically non-significant with less than 1 IQ point difference between the groups. Thus the final regression equation was:

E (WISC-III^{UK}) = -56.15 + 0.78 (maternal IQ) - 9.44 (organic condition) + 1.81 (gestational age)

5.2.2 Regression analysis of Reading Scores

In a similar exercise to the regression of children's IQ on group, each of the standardised reading scores were regressed onto group, maternal IQ and organic condition. For the remaining analyses birth weight, reported gestational age and family size were not included in the regression. Results were comparable to the regression analysis of children's IQ shown in Table 5.2.1a. None of the group differences for standardised scores for Basic Reading (-1.43), Spelling (- 0.19) and Reading Comprehension (-1.58) are statistically significant. They are shown in Tables 5.2.2a, 5.2.2b and 5.2.2c.

Table 5.2.2a The regression of Standardised Basic Reading Scores on maternal IQ, organic illness and group

Variable	*Coefficient*	*Standard error*	*t*	*p*
Maternal IQ	0.55	0.07	8.3	<0.005
Organic	-11.42	4.09	2.8	<0.05
Group	-1.43	1.94	0.74	ns
Constant	43.86			

$R = 0.53$ $df = {}_{3, 199}$ F 25.81 $p < 0.001$
Organic (0 = no; 1 = yes)
Group (0 = control; 1 = case)

Table 5.2.2b The regression of Standardised Spelling Scores on maternal IQ, organic illness and group

Variable	*Coefficient*	*Standard error*	*t*	*p*
Maternal IQ	0.54	0.07	7.67	<0.005
Organic	-12.75	4.29	2.97	<0.05
Group	-0.19	2.03	0.09	ns
Constant	45.47			

$R = 0.51$ $df = {}_{3, 199}$ $F = 23.07$ $P < 0.001$
Organic (0 = no; 1 = yes)
Group (0 = control; 1 = case)

Table 5.2.2c The regression of Standardised Reading Comprehension Scores on maternal IQ, organic illness and group

Variable	*Coefficient*	*Standard error*	*t*	*p*
Maternal IQ	0.64	0.07	9.57	<0.005
Organic	-6.87	4.53	1.5	ns
Group	-1.58	1.91	0.82	ns
Constant	36.51			

$R = 0.57$ $df = {}_{3, 189}$ $F = 31.08$ $p < 0.001$
Organic (0 = no; 1 = yes)
Group (0 = control; 1 = case)

5.2.3 Summary of regression analysis

When the mean IQ of the case and control groups were compared, a small and statistically non-significant difference of 3.0 WISC-III^{UK} Full Scale IQ points was found. Influential covariates, such as maternal IQ, and the presence of an organic condition, were used in a regression analysis in order to provide a better estimate of an individual child's IQ given their case or control status. A strong association was found with maternal IQ. The average IQ of children with an organic condition was nearly eight points lower than those children with no organic condition. However, the group difference in IQ between cases and controls after adjusting for maternal IQ and organic condition was less than 1.56 points. This very small deficit was not statistically significant and was in fact essentially the same as its standard error (1.79).

A stepwise regression was carried out to determine which covariates made a significant contribution to the explained variance of WISC-III^{UK}. Three additional covariates were entered into a stepwise regression. Birth weight and gestational age reported by the mothers were entered because there was a statistically significant difference between the groups for both of these. Family size was also entered into the analysis as the difference between groups in the number of siblings in the family approached statistical significance.

Maternal IQ, organic condition and gestational age were retained as statistically significant predictors of WISC-III^{UK} Full Scale scores, but birth weight and family size were both eliminated. The difference in WISC-III^{UK} Full Scale scores between case and control groups after adjustment for all significant covariates was less than one IQ point and remained statistically non significant.

Mean standardised scores for the reading subscales were also estimated separately for children with an organic condition and adjustment made for maternal IQ using regression techniques. The results were comparable to those for WISC-III^{UK} Full

Scale scores, the between group difference remaining small and statistically non-significant.

5.3.1 Association of WISC-III^{UK} Full Scale scores and anthropometric measures

Previous studies have found associations between psychological measures and the characteristics of infant weight gain, such as the severity of failure to thrive (Corbett, Drewett and Wright, 1996) and age at which the child was diagnosed as failing to thrive (Drotar et al, 1985, Skuse et al, 1993). As no significant between group differences were observed in psychological outcomes, a more probable finding is that any association with infant weight would be small and would be unlikely to explain more of the variance of children's IQ than the association normally found with concurrent weight for height (Pollitt and Mueller, 1982).

In order to determine if there is an association between severity of failure to thrive and IQ at age 7-9 years, a regression analysis of IQ at 7-9 years on infant weight variables reflecting severity was carried out. In addition, a regression of IQ at 7-9 years on height, weight and head circumference at eight years was carried out to determine the relationship between these.

5.3.2 Association of WISC-III^{UK} Full Scale scores at follow up with infant weights

To evaluate the association between infant growth on IQ at follow up, two variables were derived from the infant weight data. The first was a measure of severity using the greatest difference, at any age, between the predicted weight standard deviation score and actual weight standard deviation score, that is the lowest Thrive Index value for each child. In a previous study by the author (Corbett, Drewett, Wright, 1996), it was the lowest Thrive Index which was associated with WIPPSI-R Full scale scores in children aged 6 to 7 years. The second was the age at which the child first crossed the five per cent threshold.

In the regression of WISC-III^{UK} Full scale scores on group and the lowest Thrive Index value, only 1% of the variance of children's IQ was explained. There was no significant independent association between the lowest Thrive Index value and children's IQ ($t = 1.0$, $df = {}_{2, 200}$, $p = 0.3$) and this remained so after maternal IQ was entered into the regression ($t = 0.64$, $df = {}_{3, 199}$, $p = 0.5$).

The second infant growth variable was the age at which the infant first fell below the screening threshold. The median age at which the cases crossed the screening threshold was 5.0 months, with an interquartile range of 3.1 to 8.5 months. The hypothesis being tested was that the earlier the failure to thrive, the greater the deficit in IQ at follow up. When WISC-III^{UK} Full Scale scores were regressed onto the age at which the infant first fell below the screening threshold, the coefficient for WISC-III^{UK} Full Scale scores and age was 0.38, there being almost no relationship ($t = 0.69$, $df = {}_{1, 94}$, $p = 0.49$). After controlling for maternal IQ, the coefficient for age band was even smaller (coefficient for age band = 0.17, $t = 0.4$, $df = {}_{2, 93}$, $p = 0.69$).

It was not possible to derive a measure of chronicity of failure to thrive in infancy which was not highly correlated with measures of severity. Two variables were considered for this purpose; the area below the threshold (mean of thrive indices below the 5% threshold multiplied by the number of age bands below the 5% threshold), and the number of age bands in which an infant was below the threshold. Neither variable was independent of a measure of severity as the most severely affected cases were those who, generally also took longest to climb back above the threshold.

5.3.3 Association of WISC-III^{UK} Full Scale scores and height, weight and head circumference at follow up.

Stunting has been shown to be a significant predictor of developmental deficits in third world studies of undernutrition (Lasky et al, 1981, Grantham McGregor et al 1991) so, the aim of this analysis is to establish whether there is a relationship between

anthropometric measures and IQ at follow up. An association between IQ and weight for height has been reported in a study of well nourished children in the United States (Pollitt and Mueller, 1982), and so it would be important to know if shorter, lighter cases have greater deficits in IQ than controls with a similar stature.

WISC-III^{UK} Full Scale scores were regressed onto height at follow up, and group. Height was significantly associated with IQ ($t = 2.18$, $df = {}_{2, 215}$, $p < 0.05$), and the group effect, controlling for height was reduced to a 0.84 IQ point deficit and was not statistically significant ($t = 0.34$, $df = {}_{2, 215}$, $p > 0.05$). In order to ascertain if shorter stature was more strongly associated with IQ deficits in the case group, an interaction term, height * group, was entered into the regression. The interaction term tests whether the difference between the slope of the regression for each of the groups is significantly different from zero. Height remained a significant predictor of IQ ($t = 2.00$, $df = {}_{3, 214}$, $p < 0.05$), but there was no significant interaction between height and group ($t = 0.55$, $df = {}_{3, 214}$, $p > 0.05$). Therefore, short stature in the case group was not more strongly associated with IQ than short stature in the control group.

IQ and weight at follow up were not significantly associated and lighter case children did not have a larger IQ deficit than light control children. Neither was there a significant association between Body Mass Index (BMI) and children's IQ, and so thinner cases were not more severely affected. None of these variables were statistically significant after maternal IQ and the 'organic' variable were entered.

Infants failing to thrive have been found to have smaller head circumference (e.g. Chase and Martin, 1970, Field, 1984), although head circumference has not been associated with developmental quotient (Field, 1984, Wilensky et al, 1996). In the regression of WISC-III^{UK} Full Scale scores on head circumference and group, a significant association was found between head circumference and IQ ($t = 4.55$, $df = {}_{2, 213}$, $p < 0.0005$). When cases and controls with the same head circumference were compared, the difference between cases and controls was 0.27 of an IQ point which was not significant ($t = 0.23$, $df = {}_{2, 213}$, $p > 0.05$). There was no significant interaction between head circumference and group ($t = 1.2$, $df = {}_{3, 212}$, $p > 0.05$), so cases with a

smaller head circumference did not have a lower IQ than controls with a smaller head circumference.

The problem with assessing the independent contribution each anthropometric variable made, is that they were strongly correlated with each other, as can be seen on the correlation matrix for head circumference, height, weight and Body Mass Index shown in Table 5.3.3a. The only two measures that were not significantly related were Body Mass Index and head circumference. This means that any related variable entered later is less likely to significantly improve the amount of variance explained by the first variable entered.

Table 5.3.3a Correlation matrix for head circumference, height, weight and Body Mass Index

	Height	*Weight*	*Body Mass Index*
Head circumference	.58	.41	.113*
Height		.73	.377
Weight			.895

$p < 0.005$ except $p > 0.10$ for Body Mass Index and head circumference
$n = 230$

To overcome this problem of collinearity of anthropometric variables a stepwise regression method was used. Height, weight, Body Mass Index, head circumference maternal IQ and organic condition were entered into a stepwise regression. The results are shown on Table 5.3.3b.

Only two variables made a significant contribution to the explained variance of WISC-III^{UK} Full Scale scores. These were maternal IQ and head circumference which together explained 43.7% of the variance of WISC-III^{UK} Full Scale scores, maternal IQ alone accounting for 42.4% of the variance. For every 1 point increase in maternal IQ

there was a 0.75 increase in WISC-III^{UK} Full Scale scores and for every 1 cm increase in head circumference there was a 0.28 increase in WISC-III^{UK} Full Scale scores. After these variables were entered, height, weight, organic condition and group were eliminated as they failed to make a statistically significant contribution. The group difference was only 0.03 of an IQ point.

Table 5.3.3b Stepwise regression of WISC-III^{UK} Full Scale Scores on maternal IQ, organic condition, height, weight, and head circumference at age eight years and group

Variables in the equation	*Coefficient*	*Standard Error*	*t*	*p*
Maternal IQ	0.75	0.06	11.97	<0.0000
Head circumference	0.28	0.14	2.06	0.04
Constant	4.04			

R = 0.66, df = 2, 196, F = 75.7

Variables not in the equation	*Coefficient*	*t*	*p*
Height	0.06	1.20	0.23
Weight	0.05	1.02	0.31
Body Mass Index	0.03	0.61	0.54
Organic	-0.10	1.87	0.06
Group	-0.03	0.62	0.54

Organic (0 = no; 1 = yes)
Group (0 = control; 1 = case)

5.3.4 Summary of association of WISC-III^{UK} Full Scale scores and anthropometric measures

If failure to thrive has a direct effect on later intelligence, then there should be an association between the characteristics of infant weight gain and IQ at 7-9 years old. No association was found between the severity or the age at which the child crossed

the 5% threshold and later IQ. It is possible that there is an association between IQ and the length of time the child fails to thrive, but it is difficult to derive from infant weights a measure of chronicity of failure to thrive which is independent of the severity.

In separate analyses of anthropometric measures at age eight, a significant association was found between WISC-III^{UK} Full scale scores and height and head circumference (but not weight or BMI) at follow up. However, there was no significant difference in IQ between cases and controls after separately adjusting for height, weight, BMI or head circumference.

As anthropometric variables are highly correlated, a stepwise regression was used to select the variables which make a significant increment to the explained variance of WISC-III^{UK} Full Scale scores. Only two variables were selected in the regression. These were maternal IQ and head circumference. There was no significant independent contribution by height or weight, the presence of an organic condition or group.

Despite evidence for chronic slow growth in infancy continuing into childhood in the case group, there was no evidence that intellectual deficits in that group were greater than those in the control group. After maternal IQ, the best explanatory variable for WISC-III^{UK} Full Scale scores for all children tested was head circumference. Although the association with head circumference was statistically significant it was not a large effect, representing an average difference of only 2.8 IQ points across the 10 cm range of head circumference measures. No other variable significantly improved the proportion of variance explained by maternal IQ and head circumference together.

Chapter Six

Discussion

6.1 Overview of the study

There are two important results in this study. The first is that children selected for very slow weight gain in infancy were shorter, lighter, thinner and had a smaller head circumference at age eight years than a control group with similar socio-demographic characteristics. The second is that no statistically significant differences were found between the cases and controls in IQ or reading scores. Therefore, low intelligence and poor reading ability do not result from failure to thrive.

These results are contrary to findings and interpretations placed upon a number of the studies reviewed earlier, although the results from previous studies are inconsistent.

The aim of this study has been to fulfil a number of conditions in order to provide a robust design for the investigation of an association between failure to thrive, cognitive deficits and educational problems. These conditions include: using a theoretically powerful anthropometric criterion in order to identify cases by whole population screening, screening a large population in order to provide sufficient cases to detect a statistically significant difference with reasonable confidence, controlling for confounding variables, and testing children blind to their case status at school age. The study will be evaluated in the context of each of these design issues and the additional effect of sample attrition on the representativeness of the sample studied.

6.2 Screening criterion

One problem in evaluating previous research is that the criteria used to define failure to thrive are not the same, and sometimes not even explicitly described (Wilcox et al, 1989). The method of screening weight data and identifying cases used in this study

was devised both for research purposes and for clinical use (Wright et al, 1994a, Wright et al, 1994b, Wright, 1996). Its principal advantage is that an expected weight for age, and its confidence limits, can be calculated for infants of any given early weight. In this way the range of normal weight gain between two time points is defined allowing for regression to the mean. The use of a measure of relative weight conditional upon a previous weight to monitor infant growth, has been endorsed by Hall (1996), who cited the work in Newcastle on which this thesis was based and the more recent development, by Cole (1995), of a conditional reference chart for weight monitoring.

Failure to thrive, defined in this way, is a measure of post natal weight gain only, unlike an attained weight which is a measure of all the growth that has occurred since conception. The effect of using an attained weight alone is to confuse two separate risk factors for psychological outcomes, low birth weight and slow post natal weight gain. Low birth weight is attributable to short gestation, poor intrauterine growth or both. Birth weight, gestation and birth weight conditional upon gestation are associated with long term cognitive outcomes (Sorensen et al, 1997). Preterm infants were excluded from the present study because their post natal growth is more influenced by intrauterine conditions. However, the Thrive Index defines the expected post natal weight for term infants with any given early weight and thus includes those born at term but small for gestational age.

A number of studies have attempted to control for the effect of birth weight by excluding low birth weight infants from their study samples (e.g. Glaser et al, 1968, Drotar et al, 1985, Dowdney et al, 1987, Skuse et al, 1993, Wilensky et al, 1996). Three of these studies (Dowdney et al, 1987, Skuse et al, 1993, Wilensky et al, 1996) also used controls matched for birthweight. Frank and Zeisel (1988) argue that, even when controls have been selected from the same category for birth weight, cases have been found to weigh less at birth, so adverse outcomes of failure to thrive may still be explained in terms of lower birth weight. Whilst this may explain the deficit in Bayley scale scores in the recent study of Wilensky et al (1996), in the study by Dowdney et al (1987), it was the controls who were, on average, lighter at birth, and in the study by

Skuse et al (1993) the cases with the smallest deficit in Bayley scale scores were lighter at birth and showed the smallest post natal fall in relative weight.

Although earlier definitions of failure to thrive had tended to identify a disproportionate number of low birth weight infants as failing to thrive (Frank and Zeisel, 1988), in this study the birth weight of the control group was significantly lower than that for cases. Six controls with birth weights below the third centile were fully studied, compared with only two cases. Thus poor prenatal growth was a more prevalent risk factor in the control group and poor post natal growth the principal risk factor in the case group. However, neither birth weight nor post natal weight gain were found to be related to WISC-III^{UK} after allowing for maternal IQ and the presence of an organic condition. Neither were greater deficits found in WISC-III^{UK} Full Scale scores where both low birth weight and poor post natal weight gain were present.

Modifications were made to the published Thrive Index methodology in recognition of the problems of relying upon single data points for both the baseline standard deviation score, from which an expected later weight was calculated, and the later weight, used to compare with an expected weight. Instead, the baseline value used to calculate later expected weight was the mean of birth weight and weight at one and two months, and a case was only selected if there was a fall below the five per cent threshold on more than one occasion. These modifications, whilst reducing the effect of measurement error, thus improving discrimination between cases and controls, had the disadvantage that insufficient data points were available for 606 out of the 3418 full term infants in the cohort, and they were omitted from further analysis. The effect of missing data on the selection of cases was small, as in both this study and previous work carried out in Newcastle (Edwards et al, 1990, Corbett, 1994) infants whose growth was slow tended to be taken to clinic and weighed more often.

A greater problem with the inclusion of mean weight SDS up to two months of age in the calculation of the baseline value was that since the relative weight of cases was falling from birth, weights recorded after many of the infants had started to fail to

thrive were included in the calculation for the baseline value. The effect of this would be to exclude some of the milder cases, since the baseline value would be below that expected from birth weight alone, and the later expected weight was calculated conditional upon this lower value. It is all the more remarkable, therefore, that no major differences were found in IQ and reading ability that were attributable to failure to thrive.

However, as the baseline value is the mean of early weight SDS, it is not only likely to exclude mild cases but also those with a large early fall in relative weight. Four published papers (Drotar et al, 1985, Drotar and Sturm, 1988, Skuse et al, 1993, 1994) have reported that the earlier the fall in relative weight, the larger the developmental and cognitive deficits. No association was found in this study with IQ and the age at which the child first crossed the 5% Thrive Index threshold, but a problem was identified with using the age when the threshold is crossed to determine age at onset. As the relative weight for the case group was falling from birth, only crossing the fifth centile at an average age of five months, by the time infants crossed the screening threshold they had already been failing to thrive for some time. This discrepancy between the age at onset of failure to thrive, and age at which a child fell below the 5% threshold for the Thrive Index, has identified a limitation with using the age at which a child crosses the screening threshold (Drotar et al, 1985, Drotar and Sturm, 1988 and Wilensky et al, 1996) as a criterion for the detection of early onset failure to thrive.

A measure of severity, the lowest Thrive Index value, was derived from the dataset. In earlier work in Newcastle (Corbett, Drewett and Wright, 1996) a significant association was found between IQ at age 6 to 7 years and the lowest Thrive Index in infancy. However, no such association was found using this larger dataset, so even infants with the longest fall away from their expected weight did not have a lower IQ.

The Thrive Index method has enormous practical value for the screening of populations for failure to thrive, using routinely collected data, where data are often missing or very scant. Provided the possibility of measurement error is taken into account, in practise the method has the advantage that an individual case can be

detected with only one early and one late weight. However, whilst every effort has been made to reduce the amount of missing data, the limitations of the dataset for more detailed analysis of the characteristics of infant weight gain must be recognised. For example, it was not possible to determine the chronicity of failure to thrive as the number of age bands an infant fell below the screening threshold correlated with the number of weights available for each age band, suggesting a strong effect of missing data. Another way to interpret this correlation is that mothers who were most concerned about their child's weight came to the clinic more often to have their child weighed, or they were followed up more diligently by the health visitor. However, in order to explore the finer details of patterns of weight gain in infancy a more systematic approach would be required for collection of data on early weight to reduce the effect of missing data. Even with a complete dataset no measure of chronicity that was considered would have been independent of severity or the number of age bands below the screening threshold as those whose relative weight falls farthest could be expected to take longest to climb above the threshold.

6.3 Sample selection

A criticism of many studies of failure to thrive is that the sample of cases studied are biased by the use of non-anthropometric criteria during the diagnostic process (Batchelor and Kerslake, 1990), or by the use of deprived populations (Mitchell et al, 1980, Edwards, 1990, 1994). Thus, characteristics already selected for, such as family dysfunction or economic deprivation, become considered part of the syndrome (Hufton and Oates, 1977), or are used to explain the poor nutritional intake which, it is argued, results in poor growth (Skuse, 1985). In order to allow for the effect of sample selection on outcome measures, the recommendation made by Drotar (1990) is that the population from which the sample is drawn should be fully described.

This is a difficult criterion to fulfill as it requires a major epidemiological survey. Prior to this study, just such a survey was undertaken of the cohort from which the cases for this study were selected (Wright, 1993, 1994). The cohort included children from all

social classes. All the children born within the Newcastle Health Authority area whose routinely collected weight data met the selection criteria comprised the case group. There was no selection by referral, or after physical examination or interview, as none had been seen by any member of the research group prior to follow up at age eight. Thus the selection process was not influenced by attributions made by any member of the research team based on non-anthropometric criteria.

The prevalence and distribution of failure to thrive across the city by level of deprivation was described (Wright et al, 1994b). Proportionately more children failed to thrive in the most deprived areas of the city, accounting for the low mean IQ in both case and control groups (87.6 and 90.6 respectively), which is slightly above the mean of 85 found in a deprived Newcastle population previously studied (Corbett, Drewett and Wright, 1996). However, proportionately more children in the most affluent areas also failed to thrive, although numbers were smaller. The slightly larger standard deviation of IQ scores (17.3 IQ points) than expected in a standardisation population (15 IQ points) may be attributed to the attenuated socio-economic status of the case sample and selection of controls from the same deprivation strata.

Most cases (56%) were found to be living in homes given an intermediate or affluent designation for economic status (Wright et al, 1994b). It is these cases who Batchelor and Kerslake (1990) argued were least likely to be referred to hospital, and constituted 24% of a referred sample classified by Sills (1978) as having failure to thrive of undetermined aetiology, since they were neither sick nor deprived. As failure to thrive is often attributed to psychosocial factors (Hall, 1996), in any study of referred cases of failure to thrive, a number of these more affluent children with unattributed failure to thrive are likely to be missed. The strategy of selecting cases by screening city wide has resulted in screening in a case group with wide socio-economic diversity, but as with many indicators of risk to health, economically deprived cases were overrepresented.

6.4 Controlling for covariates of IQ and reading ability

A matched pair design has the advantage of controlling for differences between the case and control pair that may independently affect outcome. However, the variables likely to affect growth and cognitive outcomes are numerous, and perfect pairwise matching was unlikely to be achievable for all variables. This leads to unsatisfactory matching, or to having to exclude cases for which no match can be found. Others are excluded if, for any reason, data collection cannot be completed with their match. Instead, the groups were selected after stratification for age and sex, and level of deprivation. However, the level of deprivation measure only reflected an averaging of census data, such as level of employment, car and home ownership for that area, not for each family and so was inexact at the individual level.

It was only during the home visit that more detailed information was gathered about other relevant variables, such as: care arrangements, medical and feeding history, physical impairments, maternal education, number of siblings in the family and birth order of the study child. Each mother was also asked about employment, car and home ownership as an indicator of economic deprivation. This omitted the fourth measure on the Townsend Scale (Townsend et al, 1989), overcrowding, as measured by the percentage of private households with more than one person per room, as it was considered that this would require intrusive questioning.

Whilst the Townsend scales provide a useful indication of deprivation in a defined population, for example a percentage of unemployed, which can be compared with other populations, it does not discriminate well between moderate and very affluent individuals within a population, where not only employment, but levels of income are needed to make such a distinction. Nevertheless, in the context of other information gathered, such as maternal education, a more complete picture of each child's social and economic environment was established.

The case and control groups were found to be highly comparable, with the exception of four statistically significant differences, in feeding problems, birth weight and reported gestational age, and the presence of an organic condition likely to affect growth. There was also a trend towards larger families in the case group, but this did not achieve statistical significance. The data on feeding problems from the present study should be treated cautiously as they are dependent on mother's recall. Many different problems were recalled by mothers, so it is not possible to shed light on the origin of feeding problems in failure to thrive, merely to endorse the findings of many previous studies of more feeding problems in the case group (Pollitt and Eichler, 1976, Kotelchuck and Newberger, 1983, Hepinstall et al, 1987, Mathisen et al, 1989, Ramsay et al, 1993, Wilensky et al, 1996). It is of interest that there is a greater reduction in reported feeding problems in the control group after weaning suggesting greater persistence of problems in the cases. However, as feeding problems have been identified in previous studies of failure to thrive they would be expected where a population of failure to thrive infants had been correctly identified. This variable could not, therefore, be regarded as an accidental confounder.

Using regression methods, outcome measures were adjusted for the four potentially confounding variables found to differ between the groups, birth weight, gestational age, family size and the presence of an organic condition. Maternal IQ was also entered as a covariate. Whilst no statistically significant differences were found between maternal IQ in the case and control groups, it was found to be the single most important explanatory variable for the IQ of the child. For this reason it was retained as a covariate in all subsequent regression analyses.

All births were regarded as term births (> 37 weeks), but mothers of controls reported shorter gestational age and the controls were lighter at birth. Differences in gestational age after 37 weeks are not regarded as clinically significant and so it was unclear what these differences signified. The majority of births, were reported as 40 weeks gestation (54% for cases and 53% for controls), so it was surprising that, in a limited range of scores, these small differences for the remainder achieved statistical significance. In the regression of WISC Full Scale scores on maternal IQ, organic condition,

gestational age, birth weight, family size and group, gestational age but not birth weight was a statistically significant covariate for children's IQ. However, the gestational age variable is of limited interest for this study, as the between group difference for children's IQ remained statistically non-significant after gestation was entered into the regression.

Organic illness was not used as an exclusion criterion as it has been in previous studies (e.g. Drotar 1985) because it has been shown to be difficult to make a distinction between children with failure to thrive with an organic or non-organic aetiology (Homer and Ludwig, 1981). Children with organic failure to thrive may also have a number of psychosocial factors which predispose them to failure to thrive and it is possible to misattribute failure to thrive to organic conditions. However, it has been observed that children who fail to thrive are more likely to be ill in their first year of life (Sherrod et al, 1984) and are more likely to be admitted to hospital (Wilensky et al 1996). Wilensky and colleagues ascribed the differences they found to the increased biological vulnerability of infants who fail to thrive. It is not clear whether the organic conditions they identified were the cause or the effect of poor growth in infancy, so they did not adjust for the potential effect of the organic conditions found on developmental outcomes.

In the present study, as a medical history was not available before case selection, medical records were examined blind at follow up, in order to determine whether the child suffered from a diagnosed organic condition which would definitely or possibly affect growth or cognitive development. The two categories, conditions that would definitely or possibly affect growth, provided a measure of the degree of certainty about the likely effect conditions identified would have on growth. Whilst cases were found to have significantly more organic conditions likely to affect growth, determining the extent to which a condition would affect growth was not straightforward as three controls were also found to have a condition likely to affect growth. It is also probable that some organic conditions were missed as a number of notes did not contain a diagnosis. For example, one child had been admitted repeatedly for suspected gastritis, and at one stage was queried as having microcephaly, but there was no confirmed

diagnosis and so she was not categorised as having an organic condition. As only two children administered psychological tests, one case and one control, were categorised as having a condition which would definitely affect growth, they were included for analysis with the children categorised as having an organic condition which would only possibly affect growth. When entered into a regression of WISC-III^{UK} Full Scale scores, maternal IQ and group, children with an organic condition scored nearly 8 IQ points less than those without, the effect on reading scores being more severe, with average deficits of 11 and 13 points in the reading and spelling subtests.

In an analysis of variables associated with poor test performance, between group differences in frequency and severity of speech and learning difficulties were accounted for by the larger number of cases with an organic condition. In particular, there was no evidence for increased oral motor difficulties in speech in the case group once organic condition had been taken into account. Oral motor problems and feeding skills disorder have been observed in infants who fail to thrive (Mathisen et al, 1989, Ramsay et al, 1993). Early oral motor problems might have been expected to emerge in the form of speech and language problems, which have also previously been observed in children who have failed to thrive (Bithoney, 1986, Chase and Martin, 1970, Dowdney et al, 1987, Elmer et al, 1969, Oates et al, 1985), but there were no differences between cases and controls in reported speech disorders, referral to a speech therapist or the length of time they were treated by a speech therapist, that could not be explained in terms of organic illness, and there were no significant differences in verbal ability as measured by the WISC-III^{UK}. However, reported feeding problems in the case group remained disproportionately high, even after allowing for the effect of an organic condition.

The distribution of many variables between the groups was analysed and with few exceptions, the cases and controls were found to be comparable. After adjustment for the few statistically significant differences found, the IQ of cases and controls remained similar.

It is possible that some of the data gathered poorly discriminated certain categories of participant, for example measures of unemployment do not discriminate between well rewarded or poorly paid employment. However, it is unlikely that, after other variables have been considered, such as car and home ownership, maternal education and IQ, important differences between the groups in the variables measured would be overlooked. Where possible, verification was sought for reported problems, for example an extensive search of medical records was made where hospital or outpatient admission was reported.

Of course a retrospective survey, such as the interview with the mother, cannot take into account factors which may have been relevant at the time the child was found to be failing to thrive. It is possible that a measure taken concurrently with the episode of failure to thrive, such as the HOME inventory of Caldwell and Bradley (1976) would have revealed differences between the groups, but a limitation of this study is that no earlier work with the group was possible. However, lack of earlier contact is also a strength of this study. No differences were found between the groups, despite lack of intervention by the research team which would have been expected to reduce the differences between groups. It is also the case that some of the covariates measured at follow up, such as maternal IQ, would not be expected to vary greatly over time.

6.5 Tester awareness of clinical status of the child

The unintentional effect exerted by researchers in favour of the hypothesis being tested, both during collection and analysis of the data, is well known (Rosenthal, 1966) and may have influenced results in previous studies where cases and controls were tested unblinded (Elmer et al, 1969, Glaser et al, 1968, Hufton and Oates, 1977).

Several steps were taken to prevent those administering psychological tests from knowing the case status of the child prior to testing: cases and controls were identified by a third party, and testing of a child and entry of test results into the database were

always completed before interviewing the mother, as she may have been aware that her child had growth problems.

All testing of children and adults adhered to the protocol for the administration of that test found in the handbook. To ensure a standardised approach to testing, both psychologists (SSC and KP) were given training by a clinical psychologist, with a number of trial tests. Records were maintained of scoring decisions to ensure consistency. In this way the effect of the appearance of the child on the way the test was administered was minimised. In practice the children were between the ages of 7 and 9 years old when tested, and so it was difficult to determine which children were small for their age, particularly to the unpractised eye.

Care was taken when contacting parents to ensure that they were not given any indication whether their child was a case or control. The introductory letter sent to parents explained that the study was concerned with patterns of infant growth and its relationship to outcomes (Appendix V). The term 'failure to thrive' was not mentioned in any correspondence. This was not deception as the intentions of the study were accurately described.

When visited at home most mothers were unaware of their child's case status, few cases having been identified as such by health professionals, and only nine having been hospitalised for failure to thrive. The researcher did not know the case status of the child either. As the groups were found to be demographically similar, and both cases and controls were found to have organic conditions and feeding problems, in general it was not possible to make a distinction between case and control children based on information given at the time of interview.

Data were entered into the EPI5 database firstly by the researcher, on the day of testing, and then independently by an administrative assistant. Errors were checked using the data validation programme. The coding for case and control was entered into a different database by a third party, only being merged with that containing the outcome data when all entries had been made and checked. Programmes to analyse

the data were written into separate files and were not run interactively. In this way both data entry and analysis was carried out without revealing which individuals were cases.

Throughout the process of selection of cases and controls, data collection and data entry, blinding procedures were carefully adhered to, thus ensuring that results could not be influenced by knowledge of the child's case status.

6.6 Exclusions and attrition

One question that needs to be addressed is the likely effect of exclusions and subject attrition on the findings if those lost to follow up were different from those followed up.

Exclusionary criteria were few. Of those infants with sufficient weight data available, 235 infants with a gestational age of less than 37 weeks were excluded from the original population cohort. Whilst this may have excluded a group of infants more prone to fail to thrive, the post natal growth of preterm infants is strongly influenced by gestation and intrauterine conditions and is more difficult to evaluate alongside term infants. The requirement that multiple datapoints be available for the calculation of the baseline and the comparison with later weights lead to the exclusion of 606 individuals from the birth cohort, although the reduction in the number of cases as a result of this was minimal (n = 7). No other exclusions were made when selecting cases.

In previous research carried out in Newcastle by the author (Corbett, 1994), 89% of the study sample had been traced and studied at the age of 6-7 years old. Although many of these families had moved, and some several times, all but a few were still resident in the City. For this reason, an optimistic estimate was made of expected attrition for the present study, despite higher rates of loss to follow up reported by other studies of 37% over three years (Aylward et al, 1985), and a 35% rate of refusal

to participate (Beck et al, 1984). Low subject attrition was also avoided in earlier work in Newcastle, as families were very willing to participate in research, having been contacted by research workers at least three times since their child was 18 months of age. However, the families had never been contacted before for this study, and the children were two years older than those previously studied in Newcastle by the author and so loss to follow up could be expected to be higher.

Every effort was made to achieve the maximum rate of follow up. The task of tracing the children was formidable. Families were traced, and often retraced, through medical records and information held on the child health computer. Many families were traced through the schools, head teachers often sending letters on behalf of the research group, as they were not able to disclose the new address. It was not possible to locate 14 families who were no longer at their last known address. All families who had moved out of the North East were sent a letter about the project. In all but one instance, no reply was received, or the letter was returned by the Post Office. Arrangements to visit the one family that replied and now living in the South of England were attempted, but they were unavailable when it was possible for a researcher to visit. However, families now resident in neighbouring health authorities were more easily located. All those children resident in Northumberland, County Durham, North and South Tyneside were located and studied. Only four cases had moved away from the area and could not be traced. Where controls had moved away from the area (n = 14), they were replaced by controls with the same characteristics (sex, age, deprivation score of area of residence at 18 months).

The ten per cent rate of parental refusal was disappointing, although it was considerably better than would be predicted from some previous long term follow up studies in which up to two thirds of the original case group were lost (Mitchell et al, 1980, Oates et al , 1984, 1985). Of the ten per cent of parents (n = 27) who refused permission for their child to be tested at school, 17 gave permission for their child's height and weight to be measured in school. So not only was it possible to analyse data on weights routinely recorded during infancy of all those lost, but also the

anthropometric data at follow up was analysed for some of those children for whom there was no psychometric data available.

The cases lost to follow up were smaller than those fully studied, both in infancy and at age eight. Birth weights of those lost were lower than for those followed up, but they were not excessively low, being on average only -0.13 of a standard deviation below the standardisation population mean. The differences in weights for each age band between those lost and those followed up were not statistically significant, and the trajectory of weight gain in infancy for cases not studied was the same as cases studied. There were no significant differences, either, in height, weight and head circumference at age eight, although on average, cases lost were smaller than those followed up.

By contrast the infant weights of controls lost to follow up were higher than those followed up, but there were no statistically significant differences in any age band. There was little difference between those controls not fully studied and those given psychological tests at age eight. Despite their higher rate of weight gain in infancy, they were marginally smaller and lighter at follow up than those controls tested.

It is impossible to evaluate the likely effect of subject attrition on psychological outcomes, especially when refusal is associated with more social problems (Aylward et al, 1985, Beck et al, 1984). But the difference in rates of subject loss between groups was small. Only four more cases than controls were untraced and only three more cases than controls refused permission to participate in the study. These small differences suggest that refusal is not significantly motivated by factors which relate specifically to the case group, although the slightly higher rate of parental refusal in the case group may have been attributable to concerns about drawing attention to their child's stature, as some parents of children studied had expressed this concern. However, there is no evidence that those lost differed substantially from those followed up.

Subject attrition lead to some loss of statistical power as it was not possible to test the 120 cases and their controls needed to detect a 5 IQ point difference with 80% power.

However, the number of children followed up (107 cases and 117 controls) represented 82% of the cohort, and provided between 75% and 80% power. Despite a slightly lower rate of follow up than planned, this study nevertheless remains the largest study of its type to date.

6.7 Outcome measures

The association between failure to thrive and deficits using the Bayley Scales has now been replicated in two well conducted population based studies. This association is not surprising as testing was close in time to the episode of failure to thrive. However, the Bayley Scales are not strong predictors of later IQ (McCall, 1979, DiLalla et al, 1990) except in risk samples where neurologically disordered infants living in unstimulating environments are likely to be more prevalent (McCall, 1979, McCall and Carriger, 1993). In general socioeconomic measures correlate more strongly with childhood IQ than standardised infant tests (McCall, 1979). Thus the predictive validity of the Bayley scales is largely dependent on the presence of neurological damage and continuing social disadvantage found in samples of children failing to thrive. There was no evidence for greater neurological impairment in the nonorganic cases in the present study and, the control group were selected for comparable levels of socioeconomic deprivation.

An alternative view is that the Bayley Scales are not as good a measure as childhood IQ tests of 'whatever general intelligence is' (Slater, 1995). In order to establish whether significant IQ deficits endure into childhood or for longer, a prolonged follow up period is required in order to test children at the appropriate age using tests that have greater predictive validity for adult IQ and educational achievement. As there has been no consensus of results for studies with a longer follow up period, the aim was to test the largest sample to date with tests comparable to those previously used. The psychological outcome measures selected for this study were the WISC-III^{UK} and a reading test, WORD. Both of these tests had the advantage of a very low floor effect and no ceiling effect for children of that age and they had been standardised on the

same sample, so direct comparisons could be made with IQ and reading ability The WISC is predictive of academic achievement (Zimmerman and Woo-Sam, 1972) and could therefore be expected to identify children with more general educational problems.

In the present study, only two children could not be tested with the WISC-III^{UK}, and only three were not tested with WORD. Mean IQ scores were below average (87.6 for cases and 90.6 for controls), with a disproportionate number of children scoring more than two standard deviations below a standardisation population mean (n = 29, 10.7%). All but 17 reading tests were administered by a different psychologist, but the results were comparable to those for IQ, with low mean scores for standardised basic reading (93.5 for cases and 94.5 for controls) and 11 children scoring too low to receive the full test. The low mean scores and the high rate of low scores is attributable to the higher number of children living in deprived areas in the sample, and the inclusion of a number of children with an organic illness who scored very low.

In previous studies where reading ability was found to be worse than that expected from a standard IQ test (Hufton and Oates, 1977, Oates et al, 1984, 1985), the difference in results may be explained by the use of unrelated IQ and reading tests. For this study the WORD reading test was selected as it had the same standardisation sample as the WISC-III^{UK} . According to the manual for WORD, the case group which had a mean WISC-III^{UK} Full Scale score of 87.5 would on average have a standardised Basic Reading score of 93. They actually scored 93.5. The control group, whose mean WISC-III^{UK} Full Scale score was 90.6 would be expected to have an average mean standardised Basic Reading score of 95, the actual score being 94.5. It can be seen that reading ability for both cases and controls was close to that predicted by the relationship between the WISC-III^{UK} and WORD in the standardisation sample, and does not suggest any worrying discrepancy.

6.8 Comparison of findings with previous research

As studies of referred cases do not meet minimal design requirements, the results of non-referred population based studies only will be discussed. The criterion for failure to thrive used in most of these studies was attained weight and controls were matched for birth weight (Dowdney et al, 1987, Skuse et al, 1993, 1994, Wilensky et al 1996, Boddy, 1997), but in some later weight was compared with earlier weight (Mitchell et al , 1980, Corbett et al, 1996) thus all providing a measure of post natal weight gain. Whilst the Thrive Index is a more exact measure of this, the criterion used for these studies are comparable with the present study.

The results of this study support the overall findings of two population based studies (Mitchell et al, 1980, Boddy, 1997) and also replicate a comparison of cases and controls in an earlier study in Newcastle carried out by the author (Corbett, Drewett and Wright, 1996). However, they are discrepant with population based studies where the average age of the children tested is younger (Skuse, 1993, 1994 and Wilensky, 1996), and the study of four year olds by Dowdney et al, (1987). These discrepancies can be explained in a number of ways.

An error where a true effect was not detected, might be explained by the appropriateness of the test for the age of the child. For example, where no significant effect of failure to thrive was found on overall GCI scores (Mitchell et al, 1980, Boddy, 1997), differences between case and control groups may have been reduced as a result of a ceiling effect found when testing older children with the McCarthy scales (Kaufman and Kaufman, 1977). However, a ceiling effect would not have been anticipated in either study, as the maximum age of the children was six years old, a year below the age when a ceiling effect is usually found. Furthermore, the children studied by Mitchell et al (1980) had been selected from a poor rural population, the mean GCI scores for cases and controls being 12.5 and 7.5 points below the standardisation population mean. Mitchell and colleagues also found no significant differences in any of the subscales, but the final number tested was very small (n = 28).

In the study of Boddy (1997) the controls had scored just above the standardisation population mean when tested using the Bayley Scales in infancy, which Kaufman and Kaufman (1977) argue is more continuous with the McCarthy Scales than the WIPPSI or Stanford Binet. At follow up GCI scores for cases were 100 (sd 15.8) and for controls 104 (sd 19.7), so although the scores of controls were still above average, there was no evidence of a ceiling effect. In any case, the McCarthy Scales were able to discriminate between the groups as Boddy (1997) found statistically significant differences in the memory and quantitative subscales. In the earlier study in Newcastle (Corbett, Drewett and Wright, 1996) the WIPPSI-R was used to test 6 to 7 year olds. On average, both cases and controls scored well below the standardisation population mean, so the failure to detect a difference could not be attributed to a ceiling effect in this study either.

In the present study of eight year olds, which replicates the overall results of these studies, the WISC-III^{UK} was used. This test is appropriate for children up to the age of 16 years, so the similarity of the scores for each group cannot be explained in terms of a ceiling effect, and as a number of tests in the WISC-III^{UK} have been extended to provide easier items, it is unlikely that differences would be reduced by a floor effect. Again, like the study of Boddy (1997) in which some subtests were found to be significantly different, the cases scored significantly lower in the arithmetic test. However, the probability of finding one statistically significant difference in one of the ten subtests as a result of a type 1 error was high.

If the tests for older children are appropriate, but no clear effect has been found, then what has to be explained is the replication of statistically significant deficits in infants tested with the Bayley Scales (Skuse, 1993, 1994 and Wilensky, 1996). One way to account for this is that the scales used to test younger children test different domains. The Bayley scales, designed to test infants up to 2 years of age, consist of two subscales, the Mental Development Index (MDI) and the Psychomotor Development Index (PDI). The PDI measures gross motor abilities (e.g. sitting, standing, walking, climbing stairs) and the ability to manipulate objects. There is a wide variation in the age range within which infants achieve milestones, such as sitting, but once they

achieve the milestone, further measures of that particular motor skill have no discriminatory power. The two scales (the MDI and PDI) are linked, as manipulation of objects plays an important part in the cognitive development of infants, but as a test of sensori-motor development both scales have a low correlation with later IQ measures (McCall, 1979).

Anastasi (1990) argues that a similar problem with construct validity occurs with the McCarthy Scales, which are designed for slightly older, pre-school children. When used on children in the younger age range some tests are more a test of motor skills than cognitive development. She gives the drawing test as an example of a single task influenced by these developmental trends, as it has a large motor component in young children and a greater conceptual component in older children.

The studies of Mitchell et al (1980) and Boddy (1997) used the McCarthy Scales with children in the older age range for the test, and therefore measured abilities similar to those measured in the IQ tests used in both the Newcastle studies. However, some of the children tested by Mitchell and colleagues were younger (3 years old) and so it is unlikely that the same test used for the four year olds in the study of Dowdney (1987) has not identified a major discrepancy even if a different ability was measured in younger children. This large discrepancy (20.6 points) at age four is not likely to be explained in terms of transient deficits.

An overall explanation for the inconsistency in results of follow up studies may instead lie in their statistical power. The numbers of cases generated in population based studies are very small, despite the very large populations surveyed. The power calculations made to estimate an appropriate sample size for this study showed that to have 80% power to detect a difference in IQ of five points between cases and controls, it was necessary to study 120 children in each group. Mitchell et al (1980) followed up 12 of the 30 cases they identified and 16 controls. Dowdney et al (1987) followed up 23 out of 25 cases and their controls. Skuse et al (1993, 1994) recruited a sample of 49 cases and their controls, 47 of whom were studied by them and 42 studied later by Boddy (1997). Wilensky (1996) followed up 50 cases and controls. None of these

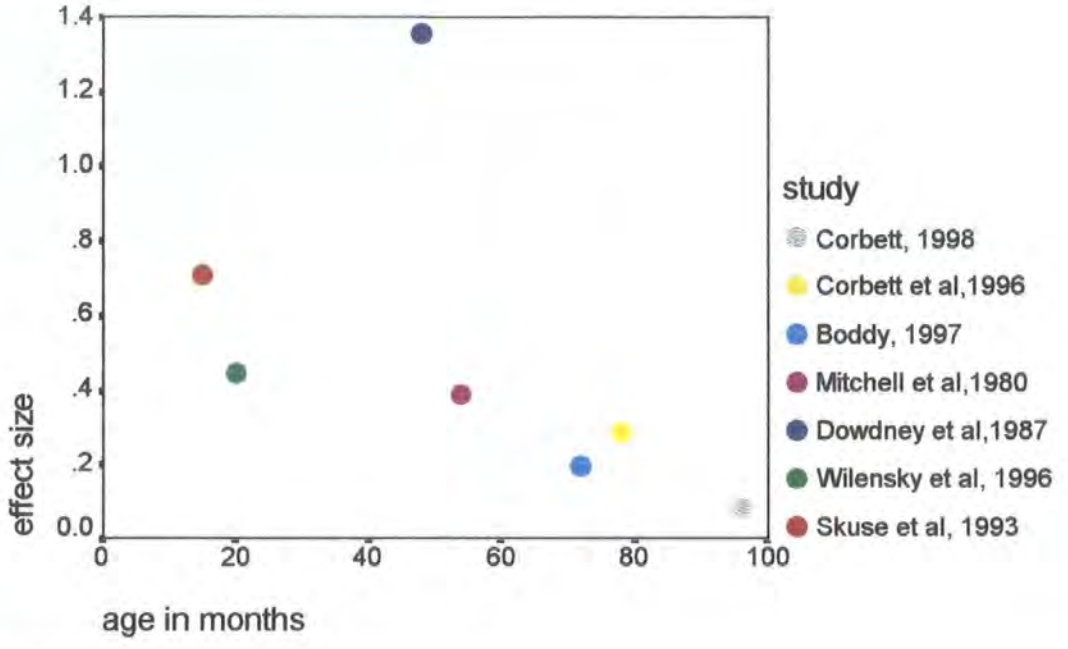
studies alone have the statistical power to reliably detect anything other than very large differences in IQ. However, when combined, Skuse et al (1993, 1994, 1996) and Wilensky et al (1996) studied 97 cases and their controls. In each study more than half a standard deviation difference in Bayley Scale scores was found. This constitutes a sufficient sample size to reliably detect such a difference.

In the present study, it was not possible to recruit the small number of additional children required to achieve 80% power. However, after subject attrition, 107 cases and 117 controls were followed up. The size of the sample followed up provides in excess of 75% power to detect a five IQ point effect, and constitutes a larger sample followed up than in any previous single population based study.

One way to overcome the problem of small sample size and lack of power in a number of studies is to combine their results (Mulrow, 1994). However, it is not possible to combine results in any simple way as studies are not comparable in terms of the tests used and the age at which testing was carried out. Instead, to make a useful comparison between studies, the difference in mean psychometric test scores between case and control groups was divided by the standard deviation of the control group. This gives an estimate, d, of the size of the effect found. When plotted against the age of the children when tested, there is a negative correlation of effect size and age ($r = -.53$), suggesting a different explanation for results. With the exception of the study by Dowdney et al (1987), it appears that as the children get older, the effect of failure to thrive on cognitive development diminishes. The analysis is shown in Figure 6.3.1a.

This is not a unique finding. Pollitt (1993) in his evaluation of a twenty year follow up of children given nutritional supplementation in Guatemala found that the beneficial effect on cognitive measures of nutritional supplementation of the mother during pregnancy and of the child up to the age of two years ceased to be detectable in children after the age of sixty months. This accords with work on children with failure to thrive of organic origin (Lloyd-Still, 1974) where cases tested when 2 to 5 years old were significantly more delayed than siblings, but older cases tested between 5 and 15 years were not.

Figure 6.8 Effect size of psychometric test plotted against age at testing



Effect size = $(\mu_1 - \mu_2) / \delta_{controls}$

The effect of early adversity appears to decline at the same time as the influence of maternal IQ increases. Only 8% of the variance of the Bayley Scales MDI was explained by maternal IQ in the study of Skuse et al (1993). However, in the study presented in this thesis, it was the single most influential covariate of children's IQ, explaining 43% of the variance. This effect was noted in the classic adoption study of Skodak and Skeels (1949) where the correlation of the biological mother's IQ and infant tests was zero, rising to 0.44 by the age of fourteen. Defries et al (1987) calculated the heritability of IQ in adopted infants aged 1, 2, 3 and 4 years old. They observed that heritability of IQ increases as a function of age from 0.10 at one year of age to 0.26 at 4 years. Similarly, in a study of transracial adoption, Scarr et al (1993) found an increase in heritability from 0.34 at age 7 to 0.50 at 17 years, and a concurrent decline in correlations between unrelated children reared as siblings in a shared environment, from 0.31 to 0.19. Taken together, these findings support the

view that the effect of slow growth on cognitive development is likely to diminish as heritability increases and the correlation of IQ with environment decreases. It is clear that the evaluation of long term effects of failure to thrive on intelligence must be conditional upon the IQ of the parent's. Unfortunately, in this study it was not practical in the time available to test the IQ of fathers, but the effect of this is reduced by assortive mating and the slightly higher correlations between the IQ of children and their mothers (Scarr et al, 1993).

The exceptionally large difference in the study of Dowdney et al (1987) remains to be explained. In this study, the cases were selected on the basis that they were below the 10th centile for height at age four, although they also failed to thrive in infancy. Therefore, the selection criterion was based on a measure of chronic slow growth and so a more severe deficit could be expected. Turkheimer and Gottesman (1991) found that extremely poor environments can reduce the heritability of IQ which would result in a slower diminution of cognitive deficits over time, thus the severe and chronic nature of poor growth provides a possible explanation for the persistence of the deficits found.

6.9 Conclusion

The work presented in this thesis represents the most robust investigation of the long term outcomes of failure to thrive to date, both in terms of design and numbers studied. The cases were selected for relatively slow weight gain in infancy, and were shorter, lighter and thinner, with a smaller head circumference than controls at eight years. Yet, there was no evidence that these children had lower IQ or poorer reading ability, or were more likely to require special needs education than the comparison group, either before or after adjustment for maternal IQ and organic illness. Neither was there any evidence to suggest a direct association with characteristics of their early growth, such as the severity of poor growth, or age at which they were screened in as a case. The associations found between IQ and anthropometric measures at age eight were small, and were found in equal measure in the case and control groups.

The most compelling explanation considered is that deficits found in infancy are reduced as children get older. Whilst this is a reassuring finding, it would be wrong to assume that there is no problem which needs to be addressed. The majority of case children in this study showed a dramatic fall in relative weight from birth, with persistent deficits in anthropometric measures at age eight. These anthropometric characteristics have been found to be associated with higher blood pressure in children, and increased incidence of coronary heart disease in adults (Barker and Osmond, 1992).

A number of issues have been raised by the findings of this study. The first observation which requires further investigation is that the relative weight of the majority of cases was falling from birth, with the largest fall occurring within the first six months post natally. As the infants did not cross the 5% Thrive Index threshold until they were on average between five and six months old, most infants will have completed this large early fall before being screened in as a case. This is not a unique problem as Drotar and Sturm (1988) defined age at onset as the age at which a child crossed the 5th centile of the NCHS standards. The average age at onset was three months, suggesting a subtantial fall in relative weight before three months of age. This observation has a number of implications for research, not least the need to develop a method of identifying cases early, enabling them to be studied at the time when their growth faltering is greatest.

The study found evidence consistent with other studies which suggest possible explanations for failure to thrive. There were more problems during parturition in the case group and the cases had more hospital and outpatient admissions in the first years of life. The higher incidence of illness in case groups has been observed in previous studies (Mitchell et al, 1980, Sherrod et al, 1984, Wilensky et al, 1996) and has been attributed to the greater biological vulnerabilty of children who fail to thrive (Wilensky et al, 1996). This study was not designed to, nor has been able to establish the causal direction of this association, so, it is not clear if multiple minor illnesses cause failure to thrive, or if failure to thrive predisposes a child to illness. As in a number of previous

studies (Pollitt and Eichler, 1976, Kotelchuck and Newberger, 1983, Altemeier et al, 1985, Hepinstall et al, 1987, Mathisen et al, 1989, and Wilensky et al, 1996), feeding problems were more prevalent in the children who failed to thrive, even after allowing for organic illness. As the relative weight of cases fell from birth, then failure to thrive cannot simply be regarded as a weaning problem, and investigations are needed of the development of feeding skills and appetite disorders from birth. Feeding problems were reported for both cases and controls. However, proportionately fewer feeding problems were reported after weaning in the control group than in the case group. This persistence of feeding problems at weaning, and the lower BMI for cases at eight years old, an age when food intake is largely controlled by the child, suggests a continuity of feeding or appetite problems that require further study especially as evidence exists that a lack of interest in food at nine years old is associated with eating disorders in adolescence (Marchi and Cohen, 1990).

A limitation of this study was that there are no early data for cognitive development and so conclusions about the continuity of developmental deficits are based on an analysis of separate studies of cognitive outcomes of failure to thrive and age at follow up, not repeated measures. Only one population based study has so far published results for cases studied in infancy and retested in childhood (Skuse et al, 1993, 1994 and Boddy, 1997) and their findings provide particularly strong evidence for a diminution of deficits over time. Replication of a follow up of this kind would be important in establishing whether cognitive deficits are reduced in time and what factors are influential in their rate of reduction.

It was not possible on this occasion to systematically observe other psychological outcomes. With such a large sample the decision was taken to focus on cognition and not emotional or behavioural problems. Early studies of failure to thrive reported high levels of emotional and behaviour disorder (Glaser et al, 1968, Elmer et al, 1969 and Hufton and Oates, 1977, Oates et al, 1984). As discussed, these studies had methodological limitations. In work previously carried out by the author in Newcastle no differences were found between cases and controls in parent or teacher ratings of behaviour problems, however, the criterion for failure to thrive was found to include

children who would not now be regarded as failing to thrive. The children studied for this thesis were selected using a theoretically stronger criterion, they were selected city wide and not just from poorer areas within the city, and thus would provide an excellent sample for a study of long term emotional and behaviour problems.

As the detection and treatment of failure to thrive requires substantial resources it is important to determine in what way it matters in order that resources are used appropriately. Failure to thrive may affect individuals by disrupting their physiological development predisposing them to health problems in later life, or may be indicative of poor appetite with the potential to lead to eating disorders, or may adversely affect emotional and behavioural development. The work presented in this thesis is concerned only with potential cognitive deficits and can conclude that no deficits found in IQ and reading ability in eight year old children were attributable to failure to thrive.

References

Abel-Smith, A.E. and Knight-Jones E.B. (1990) The abilities of very low-birthweight children and their classroom controls. *Developmental Medicine and Child Neurology*, 32, 590-601

Altmeier, W.A., O'Connor, S.M., Sherrod, K.B. and Vietze, P.M. (1985) Prospective study of antecedents for nonorganic failure to thrive. *Journal of Pediatrics*, 106, 360-365.

Anderson, J. E. (1939) The limitations of infant and preschool tests in the measurement of intelligence. *The Journal of Psychology*, 8, 351-379.

Aylward, G.P., Hatcher, R.P., Stripp, B., Gustafson, N.F., and Leavitt, L.A., (1985) Who goes and who stays: subject loss in a multicenter longitudinal follow-up study. *Journal of Developmental Pediatrics*, 6, 3-8.

Baddeley, A. (1990) Chapter 15: Memory, emotion and cognition. Mood and memory, pp390-395. *Human memory.* LEA: London.

Barker, D.J.P. and Osmond, C. (1992) Intrauterine and early post natal origins of cardiovascular disease and chronic bronchitis. In D.J.P.Barker (Eds.) *Fetal and Infant Origins of Adult Disease*. British Medical Journal Publication: Tavistock Square, London. Chapter 4.7

Batchelor, J., and Kerslake, A. (1990) *Failure to find failure to thrive: The case for improving screening, prevention and treatment in primary care*. Whiting and Birch Ltd.: London.

Bayley, N. (1933) Mental growth during the first three years: A developmental studyof 61 children by repeated tests. *Genetic Psychology Monographs*. 14, 1-92

Bayley, N. (1954) Some increasing parent-child similarities during the growth of children. *Journal of Educational Psychology*, 45, 1-21.

Bayley, N. (1970) Development of mental abilites. In *Carmichael's Manual of Child Psychology 3rd Edition*. In Mussen (Eds): Vol. 1 16, pp1163-1209

Beck, S., Collins,L., Overholser, J. and Terry, K. (1984) A comparison of children who receive and who do not receive permission to participate in research. *Journal of Abnormal Child Psychology*, 12, 573-580.

Bellman, M. and Cash, J. (1987) *The Schedule of Growing Skills*. London: NFER. Nelson.

Bennett, B., Fulker, D.W. and Defries, J.C., (1985) Familial resemblance for general cognitive ability in the Hawaii Family Study of Cognition. *Behaviour Genetics*, 15, 401-406

Berkey, C.S., Reed, R.B. and Valadian, I. (1983) Longitudinal growth standards for preschool children. *Annals of Human Biology*. 10, 1, 57-67.

Berwick, D.M. (1980) Non-organic failure to thrive. *Pediatric Reviews*, 1, 265-270.

Birns, B. and Golden, M. (1972) Prediction of intellectual performance at 3 years from infant tests and personality measures. *Merrill-Palmer Quarterly*, 18, 53-58.

Bishop, D.V.M., North, T. and Donlan, C. (1995) Genetic basis of specific language impairment - Evidence from a twin study. *Developmental Medicine and Chiuld Neurology*, 37, 1, 56-71.

Bithoney, W.G. (1986) Elevated Lead Levels in Children with Non-Organic Failure to Thrive. *Pediatrics*, 78, 891-895.

Black, M.M., Dubowitz, H. Hutcheson, J. Berenson-Howard, J. and Starr, R. (1995) A randomized clinical trial of home intervention for children with failure to thrive. *Pediatrics*, 95, 807-814.

Boddy, J.M. and Skuse, D.H. (1994) Annotation: The process of parenting in Failure to thrive. *Journal of Child Psychology and Psychiatry*, 35, 3, 401-424.

Boddy, J.M. (1997*) Maternal characteristics and development of children who failed to thrive*. PhD Thesis. University of London: Institute of Child Health.

Bowlby, J. (1953) *Child care and the growth of love*. Penguin.

Bradley, R. and Caldwell, B. (1976) The relations of Home environment to mental test performance at 54 months: A follow up study. *Child Development*, 47, 1172-1174.

Butler, N.R., Golding, J. (1986*) From birth to five: A study of the health and behaviour of Britain's five year olds.* Pergamon Press.

Caldwell, B.M. and Bradley, R.H. (1976*) HOME observation for the measurement of the environment.* Little Rock, AK: Center for Child Development and Education, University of Arkansas.

Casey, P.H. and Arnold,W.C. (1985) Compensatory growth in infants with severe failure to thrive. *Southern Medical Journal*, 78, 1057-1060.

Cattell R.B. (1940) *The measurement of intelligence in infants and young children.* New York: Science Press.

Cattell, R.B. (1963) Theory of fluid and crystallized intelligence: a critical experiment. *Journal of Educational Psychology*, 54, 1-22

Cavanaugh, M.C., Cohen, I., Dunphy, D., Ringwell, E. A. and Goldberg, I. D. (1957) Prediction from the Cattell Infant Intelligence Scale. *Journal of Consulting Psychology*, 21, 33-37.

Chase, H.P. and Martin, H.P. (1970) Undernutrition and child development. *The New England Journal of Medicine*, 282, 17, 933-939.

Coddington R.D.(1972) The significance of Life Events as etiologic factors in the diseases of children: A study of a normal population. *Journal of Psychosomatic research* 16, 205-213.

Code Of Practice on the identification and Assessment of Special Educational Needs (1994) Department for Education. HMSO.

Cohen, J. (1992) A power primer. *Psychological Bulletin*, 112,155-159.

Cole, T.J. (1990) The LMS method for constructing normalized growth standards. *European. Journal of Clinical Nutrition.* 44:45-60.

Cole, T.J. (1994) Do growth chart centiles need a face lift? *British Medical Journal.* 308: 641-2.

Cole, T.J. (1995) Conditional reference charts to assess weight gain in British infants. *Archives of Disease in Childhood*: 73: 8-16.

Cole, T.J., Freeman, J.V., Preece, M.A. (1995) Body mass index reference curves for the UK, 1990. *Archives of Disease in Childhood*: 73: 25-29

Corbett, S.S. (1994) *Cognitive and behavioural outcomes of non-organic failure to thrive*. MA Thesis: University of Durham.

Corbett, S.S., Drewett, R.F. and Wright, C.M. (1996) Does a fall down a centile chart matter? The growth and developmental sequelae of mild failure to thrive. *Acta Paediatrica.* 85. 1278-1283

Cravioto, J., DeLicardie, E.R., and Birch, H. G. (1966) Nutrition, Growth and Neurointegrative Development: An Experimental and Ecologic Study. *Pediatrics,* 38, 2, 129-372

Dahl, M., and Kristiansson, B. (1987) Early feeding problems in an affluent society. IV: Impact on growth up to two years of age. *Acta Paediatrica,* 76, 881-888

Davies, D.P. (1980) Early growth of overweight and underweight babies at birth. *Proc Nutr Soc* 39, 25-31.

Davies, D.P. and Williams, T. (1983) Is weighing babies in clinics worth while? *British Medical Journal,* 286, 860-863.

Dean, A.G., Dean, A.J., Burton, A.H., Dicker, R.C. (1990) *Epi Info, Version 5: a word processing, database, and statistics program for epidemiology on microcomputers.* USD, Incorporated, Stone Mountain, Georgia.

Defries, J.C., Plomin, R., and LaBuda, M.C. (1987) Genetic stability of cognitive development from childhood to adulthood. *Developmental Psychology,* 23, 4-12.

Dewey, K.G., Heinig, J. Nommsen, L.A., Peerson, J.M. and Lonnerdal, B. (1992) Growth of breast fed and formula fed infants from 0-18 months: The DARLING Study. *Pediatrics,* 89, 1035-1041.

DiLalla, D.L. and Fulker, D. W. (1989) Infant measures as predictors of later IQ: The Twin Infant Project (TIP). *Behaviour Genetics,* 19, 753-754

Doll, E.A.(1965) *Vineland Social Maturity Scale*. Circle Pines, MN: American Guidance Service.

Dowdney, L., Skuse, D., Hepinstall, E., Puckering, C., Zur-Szpiro, S. (1987) Growth retardation and developmental delay amongst inner city children. *Journal of Child Psychology and Psychiatry*, 28, 529-540.

Downie, A.B., Mulligan, J., Stratford, R.J., Betts, P.R., Voss, L.D. (1997) Are short normal children at a disadvantage? The Wessex growth study. *British Medical Journal*, 314, 97-100.

Drotar, D., Malone, C.A., Devost, L., Brickell, C., Mantz-Clumpner, C., Negray, J., Wallace, M., Woychik, J., Wyatt, B., Eckerle, D., Bush, M., Finlon, M.A., El-Amin, D., Nowak, M., Satola, J., Pallotta, J., (1985) Early preventive intervention in failure to thrive: Methods and preliminary outcome. In D.Drotar (Eds.) *New directions in Failure to thrive: implications for research and practice*. New York: Plenum Press.

Drotar, D. and Sturm, L., (1988) Prediction of intellectual development in young children with early histories of non-organic failure to thrive. *Journal of Paediatric Psychology*, 13, 2, 281-296.

Drotar, D. (1990) Sampling issues in research with non-organic failure-to-thrive children. *Journal of Pediatric Psychology*, 15, 2, 255-272.

Edwards, A.G.K. (1987) *Failure to thrive in Newcastle children*. B. Med. Thesis. University of Newcastle Upon Tyne: Medical School.

Edwards, A.G.K., Halse, P.C., Parkin, J.M. and Waterston, A.J.R. (1990) Recognising failure to thrive in early childhood. *Archives of Disease in Childhood*, 65, 1263-1265.

Edwards, A.G.K., Halse, P.C., and Waterston, A.J.R. (1994) Does poor weight gain identify children in need. *Child Abuse Review*, 3, 107-119.

Elardo, R., Bradley, R. and Caldwell, B. M. (1975) The relation of infants' home environments to mental test performance from six to thirty-six months: A longitudinal analysis. *Child Development*, 46, 71-76.

Elmer, E., Gregg, G., and Ellison, P. (1969) Late results of the 'failure to thrive' syndrome. *Clinical Pediatrics*, 8, 10, 584-589.

Escalona, S. K. and Moriarty, A. (1961) Prediction of school-age intelligence from infant tests. *Child Development*, 32, 597-605.

Field, M., (1984) Follow-up developmental status of infants hospitalized for non-organic failure to thrive. *Journal of Pediatric Psychology*, 9, 2, 241-256.

Fillmore, E. A.(1936) Iowa tests for young children. *University of Iowa Studies in Child Welfare*, 11, 1-58

Frank, D.A. and Zeisel, S.H.(1988) Failure to thrive. *The Pediatric Clinics Of North America*, 35, 6, 1187-1206.

Freeman, J.V., Cole, T.J., Chinn, S., Jones, P.R.M., White, E.M., and Preece, M.A. (1995) Cross sectional stature and weight reference curves for the UK, 1990. *Archives of Disease in Childhood*, 73: 17-24.

Freeman, H. E., Klein, R.E., Kagan, J. and Yarborough, C. (1977) Relations between nutrition and Cognition in rural Guatemala. *American Journal of Public Health*, 67, 233-239.

Gairdner, D. and Pearson, J. (1985) Revised Gairdner-Pearson growth charts. *Archives of Disease in Childhood*, 60, 1202

Giani, U., Filosa, A. and Causa, P. (1996) A non-linear model of growth in the first year of life. *Acta Paediatrica*, 85: 7-13.

Glaser, H.H., Heagarty, M.C., Bullard, D.M. and Pivchik, E.C. (1968) Physical and psychological development of children with early failure to thrive. *Journal of Pediatrics*, 73, 5, 690-698.

Goffeney, B., Henderson, N. B. and Butler, B. V. (1971) Negro-white, male-female eight month developmental scores compared with WISC and Bender test scores. *Child Development*, 42, 595-604.

Goldfarb, W. (1943) *Journal of Experimental Education*, 12, 106.

Grantham-McGregor, Powell, C.A. and Fletcher, P. (1989) Stunting, severe malnutrition and mental development in young children, *European Journal of Clinical Nutrition* 43, 403-409.

Grantham-McGregor, Powell, C.A., Walker S.P. and S.M., Himes, (1991) Nutritional Supplementation, psychosocial stimulation, and mental development of stunted children: the Jamaican study. *Lancet*, 338, 1-5.

Griffiths, R. (1967) *The abilities of babies*. London: University of London Press

Groth-Marnat (1990) *Handbook of Psychological Assessment*. Wiley - Interscience Publication

Gunston, G.D., Burkimsher, D., Malan, H. and Sive, A.A. (1992) Reversible shrinkage in kwashiorkor: an MRI study. *Archives of Disease in Childhood*. 67, 1030-1032.

Hall, D.M.B. (1996) *Health for all children*. Third edition. Oxford University Press.

Hammill P.V.V., Drizd, T.A. Johnson, C.L. Reed, R.B. and Roche, A.F. (1976) NCHS Growth Charts. *Monthly Vital Statistics* Report 25, 3 (Supplement) DHEW Pub. No. (H.R.A.) 76.

Healey, M.J.R. (1978) Notes on the statistics of growth standards. *Annals of Human Biology.* 1 (No 1), 41-46.

Hepinstall, E., Puckering, C., Skuse, D., Dowdney, L. and Zur-Szpiro, S. (1987) Nutrition and mealtime behaviourin families of growth retarded children. *Human Nutrition*, 41a, 390-402.

Hermanussen, M., Sippell, W.G. and Valk, I.M. (1985) Knenometric monitoring of early effects of human growth hormone on leg length in children with growth hormone deficiency. *Lancet*, 1, 1069-1071

Herrnstein and Murray (1994) *The Bell Curve: Intelligence and class structure in American life*. New York: Free Press.

Hertzig, M.E., Birch, H.G., Richardson, S.A. and Tizard, J (1972) Intellectual levels of school children severely malnourished during the first two years of life. *Pediatrics*, 49, 6, 814-824.

Hill, R.M., Verniaud, W. M., Deter, R.L., Tennyson, L.M., Rettig, G.M., Zion, T.E., Vorderman, A.L., Helms, P.G., McCulley, L.B., and Hill, L.L. (1984) The effect of interuterine malnutrition on the term infant: A 14-year progressive study. *Acta Paediatrica Scandinavia*, 73, 482-487

Hindley, C. B. (1965) Stability and change in abilities up to five years: Group trends. *Journal of Child Psychology and Psychiatry*, 6, 85-99.

Ho, H.Z. (1987) Interaction of early caregiving environment and infant development status in predicting subsequent cognitive performance. *British Journal of Developmental Psychology*, 5, 183-191.

Homer, C. and Ludwig, S. (1981) Categorization of aetiology of failure to thrive. *American Journal of Disease in Childhood*, 135, 848-851.

Honzik, M. P., Macfarlane, J. W. and Allen, L. (1948) The stability of mental test performance between two and eighteen years. *Journal of Experimental Education*, 18, 309-324.

Hufton, I.W. and Oates, R.K. (1977) Non-organic failure to thrive: A long-term follow-up. *Pediatrics*, 59, 73-77.

Ireton, H., Thwing, E. and Gravem, H. (1970) Infant mental development and neurological status, family, socioeconomic status, and intelligence at age four. *Child Development*, 41, 937-946.

Kangas, J., Butler, B.V. and Goffeney, B. (1966) Relationship between preschool intelligence, maternal intelligence, and infant behaviour. Second Scientific Session, Collaborative Study on Cerebral Palsy Mental Retardation, and Other Neurological and Sensory Disorders of Infancy and Childhood, U.S. Department of Health, Education and Welfare, Public Health Service. 91-102

Kaufman, A.S. and Kaufman, N.L. (1977) *Clinical Evaluation of Young Children with the McCarthy Scales*. Grune and Stratton. New York.

Kellerher, K.J., Casey, P.H. Bradley, R.H., Pope, S.K., Whiteside, L. Barrett, K.W., Swanson, M.E. Kirby, R.S. (1993) Risk factors and outcomes of failure to thrive in low birth weight preterm infants. *Pediatrics*, 91, 5, 941-948.

Klackenburg-Larsson, I. and Stensson, J. (1968) Data on the mental development during the first five years. In 'The development of children in a Swedish urban community: A prospective longitudinal study.' *Acta Paediatrica Scandinavia*, Supplement 187, IV, Stockholm: Almqvist and Wiksell.

Kotelchuck, M. and Newberger, E.H. (1983) Failure to thrive: A controlled study of familial characteristics. *Journal of the American Academy of Child Psychiatry*, 22, 4, 322-328.

Kraemer, H.C. and Theimann, S. (1987) *How many Subjects? Statistical Power Analysis in Research*. Sage. London.

Kristiansson, B., Karlberg, J. and Fallstrom, S.P. (1981) Infants with a low rate of weight gain. 1. A study of organic factors and growth patterns. *Acta Paediatrica Scandinavia* 70: 655-662.

Kuh, D. and Wadsworth, M. (1989) Parental Height: Childhood environment and subsequent adult height in a national birth cohort. *International Journal of Epidemiology*. 18, 3, 663-668.

Lacey, K.A. and Parkin, J.M. (1974) The normal short child: Community study of children in Newcastle Upon Tyne. *Archives of disease in Childhood*, 49, 417-424.

Lasky, R.E., Klein, R.E., Yarborough, C., Engle, P.L., Lechtig, A. and Martorell, R. (1981) The relationship between physical growth and infant behavioural development in rural Guatemala. *Child development*, 52, 219-226.

Largo, R.H., Graf, S., Kundu, S., Hunziker, U., and Molinari, L. (1990) Predicting developmental outcome at school age from infant tests of normal, at-risk and retarded infants. *Developmental Medicine and Child Neurology*, 32, 30-45.

Lea, M. (1986) WAIS-R^{UK}: *A British Supplement to the Manual of the Weschler Adult Intelligence Scale - Revised.* Psychological Corporation. Harcourt Brace & Company. London.

Lloyd-Still, J.D., Hurwitz, I., Wolff, P. and Shwachman, M.D. (1974) Intellectual development after severe malnutrition in infancy. *Pediatrics,* 54, 306-311.

Lucas, A., Morley, R., Cole, T.J., Lister, G. and Leeson-Payne, C. (1992) Breast milk and subsequent intelligence quotient in children born preterm. *Lance*t, 339: 261-4

Lynch, A., Mitchell, L.B., Vincent, E.M., Trueman, M. and MacDonald, L. (1982) The McCarthy Scales of Children's Abilities: A normative study on English 4-year-olds. *British Journal of Educational Psychology*, 52,133-143.

Mathisen, B., Skuse, D., Wolke, D., and Reilly, S. (1989) Oral motor dysfunction and failure to thrive among inner-city infants. *Developmental Medicine and Child Neurology*, 31, 293-302

Mattarazzo, J.D. (1972) *Wechsler's measurement and appraisal of adult intelligence*. Fifth Edition. Baltimore. Williams and WIlkins.

McCall, R.B., Hogarty, P.S. and Hurlburt, N. (1972) Transitions in sensorimotor development and the prediction of childhood IQ. *American Psychologist*, 27, 728-748.

McCall, R.B. (1979) The development of intellectual functioning in infancy and the prediction of later IQ. In J.D. Osofsky (Eds.) *Handbook of Infant Development.* New York: Wiley. Chapter 20, 707-741

McCall, R.B., (1981) Nature-nurture and the two realms of development: A proposed integration with respect to mental development. *Child Development*, 52, 219-226.

McCall, R.B. and Carriger, M.S. (1993) A meta-analysis of infant habituation and recognition memory performance as predicotrs of later IQ. *Child Development*, 64, 57-79.

McCarthy, D. (1972) *The McCarthy Scales of Children's Abilities.* Psychological Corporation. New York.

Mitchell, W.G., Gorrell, R.W. and Greenberg, R.A. (1980) Failure to thrive: A study in the primary care setting. Epidemiology and follow-up. *Pediatrics*, 65, 971-977.

Money, J., Annecillo, C. and Kelley, J.F. (1983) Growth of intelligence: Failure and catchup associated respectively with abuse and rescue in the syndrome of abuse dwarfism. *Psychneuroendocrinology*, 8, 3, 309-319.

Moore, T. (1967) Language and intelligence: A longitudinal study of the first eight years. Part 1. Patterns of development of boys and girls. *Human Development*, 10, 88-106.

Mulrow, C.D. (1994) Rationale for systematic reviews. *British Medical Journal*, 309, 597-599.

Mutch, L., Leyland, A. and McGee, A. (1993) Patterns of neuropsychological function in low-birthweight population. *Developmental Medicine and Child Neurology*, 35, 943-956.

Neale M.D. (1989) *Manual for the Neale Analysis of Reading Ability- Revised British Edition.* NFER-Nelson

Nelson, V. L. and Richards, T. W. (1939) Studies in mental development: III. Performance of twelve-months-old children on the Gesell schedule, and its predictive value for mental status at two and three years. *Journal of Genetic Psychology*, 54, 181-191.

Oates, K., Peacock, A. and Forrest, D. (1984) Development in children following abuse and nonorganic failure to thrive. *American Journal of Diseases in Children*, 138, 764-767.

Oates, K., Peacock, A. and Forrest, D., (1985) Long term effects of nonorganic failure to thrive. *Pediatrics*, 75, 36-40.

Oppenheimer, S., Whitman, P. and Rutman, H. (1965) Prevalence of mental retardation in a pediatric out-patient clinic population. *Pediatrics*, 36, 922

Peterson, K.E., Rathbun, J.M. and Guillermo, M. (1995) Growth data analysis in FTT treatment and research. 157-176. In D.Drotar (Eds.) *New directions in Failure to thrive: implications for research and practice.* New York: Plenum Press.

Pollitt, E. (1969) Ecology, Malnutrition,and Mental Development. *Psychosomatic Medicine*, 31, 193-200

Pollitt, E. and Eichler, A. (1976) Behavioural Disturbances Among Failure to Thrive Children. *American Journal of Disease in Childhood*, 130, 24-29.

Pollitt, E. and Mueller, W. (1982) The relation of growth to cognition in a well nourished preschool population. *Child Development*, 53, 1157-1163.

Pollitt, E., Gorman, K.S., Engle, P.L., Martorell, R., Rivera, J. (1993) Early Supplementary Feeding and Cognition. *Monographs of the Society for Research in Child Development*, 58, 7.

Powell, G.F., Brasel, J.A., Raiti, S. and Blizzard, R.M. (1967) Emotional deprivation and growth retardation simulating idiopathic hypopituitarism. II. Endocrinologic evaluation of the syndrome. *New England Journal of Medicine*, 276, 1279-1283.

Preece, M.A., Freeman, J.V. and Cole, T.J. (1996) Sex differences in weight in infancy: Published centile charts for weight have been updated. Letter to the *British Medical Journal*, 313, 1486.

Psychological Corporation (1993*) WORD: Weschler Objective Reading Dimension Manual*. Psychological Corporation. Harcourt Brace & Company. London.

Ramsay, M. Gisel, E.G. and Boutry, M. (1993) Non-organic failure to thrive: Growth failure secondary to feeding skills disorder. *Developmental Medicine and Child Neurology*, 35, 285-297.

Reed, E.W. and Reed, S.C. (1965) *Mental retardation: A family study*. Philadelphia: W.B.Saunders.

Reed, R.B.and Stuart, H.C. (1959) Patterns of growth in height and weight from birth to eighteen years of age. *Pediatrics*, 24, 904-921.

Ruddy, M.G and Bornstein, M.H. (1982) Cognitive correlates of infant attention and maternal stimulation over the first year of life. *Child Development*, 53, 183-188.

Rutter, D.R. and Quine, L. (1990) Inequalities in pregnancy outcome: A review of psychsocial and behavioural mediators. *Social Science Medicine*, 30, 5, 553-568.

Sameroff, A.J. and Chandler, M.J. (1975) Reproductive risk and the continuum of caretaking casualty. In Horowitz, F.D. (Ed.), *Review of Child Development Research* (Vol.4, pp. 187-244). Chicago: University of Chicago Press.

Sameroff, A.J., Seifer, R., Barocas, R., Zax, M. and Greenspan, S. (1987) Intelligence quotient scores of 4 year-old children: Social environment Risk Factors. *Pediatrics*, 79, 3, 343-350.

Sattler, J.M. (1988) *Assessment of Children's Intelligence*. Revised Reprint. W.B.Saunders Company. London.

Scarr S. (1981) *Race, social class, and individual differences in IQ*. LEA. New Jersey.

Scarr, S., Weinberg, R.A. and Waldman, I.D. (1993) IQ correlations in transracial adoptive families. *Intelligence*, 17, 541-555.

Seigel, S. (1956) *Nonparametric statistics for the Behavioural Sciences.* Tokyo: McGraw Hill Kogakusha Ltd.

Seigel, L.S. (1982) Reproductive, perinatal and environmental factors as predictors of the cognitive and language development of preterm and full term infants. *Child Development*, 53, 963-973.

Sherrod, K.B. O'Connor, S. and Vietze, P.M. (1984) Child Health and maltreatment. *Child Development*, 55, 1174-1183.

Sills, R.H. (1978) Failure to thrive: the role of clinical and laboratory evaluation. *American Journal of Disease in Childhood*, 132, 967-969

Singer, L.T. and Fagan, J.F. (1984) Cognitive development in the failure to thrive infant: A three year longitudinal study. *Journal of Pediatric Psychology*, 9,3, 263-283.

Singer, L.T. (1986) Long-term hospitalization of failure to thrive infants: Developmental outcome at three years. *Child Abuse and Neglect*, 10, 479-486.

Silverstein, A.B. (1982) Two and Four Subtest Short Forms of the Wechsler Adult Intelligence Scale-Revised. *Journal of Consulting and Clinical Psychology*, 50, 3, 415-418.

Silverstein, A.B. (1985) Two and Four Subtest Short Forms of the WAIS-R: a closer look at validity and reliability. *Journal of Clinical Psychology*, 41, 1, 95-97.

Skodak, M. and Skeels, H.M. (1949) A final follow up of one hundred adopted children. *Journal of Genetic Psychology*, 75, 85-125.

Skuse, D. (1985) Non-Organic failure to thrive: a reappraisal. *Archives of Disease in Childhood*. 60, 173-178

Skuse, D., Reilly, S. and Wolke, D. (1993) Psychosocial adversity and growth in infancy. *European Journal of Clinical Nutrition*, 47, suppl 3, 1-18.

Skuse, D., Pickles, A., Wolke, D. and Reilly, S. (1994a) Postnatal growth and mental development: evience for a "sensitive period". *Journal of Child Psychology and Psychiatry*. 35, 3, 521-545.

Skuse, D., Gilmour, J., Tian, C.S. and Hindmarsh, P. (1994b) Psychosocial assessment of children with short stature: a preliminary report. *Acta Paediatrica Supplement* 406, 11-16.

Skuse, D.H. (1996) *Failure to thrive in the first postnatal year: an inner city community survey*. PhD thesis. University of London. Institute of Child Health.

Slater, A. (1995) Individual differences in infancy and later IQ. *Journal of Child Psychology and Psychiatry*, 36, 69-112.

Sorensen, H.T., Sabroe, S. Olsen, J. Rothman, K.J. Gillman, M.W. and Fischer, P. (1997) Birth weight and cognitive function in young adult life:historical cohort study. *British Medical Journal*, 315, 401-403.

Spitz, R.A. (1945) Hospitalism. *Psychanalytic Studies of Children*, 1, 53-74.

Stoch, M.B. and Smythe, P.M. (1963) Does undernutrition during infancy inhibit brain growth and subsequent intellectual development? *Archives of Disease in Childhood*, 38, 546-552

Tanner, J.M., Whitehouse, R.H. and Takaishi, M. (1966a) Standards from Birth to Maturity for Height, Weight, Height Velocity, and Weight Velocity: British Children, 1965. Part I. *Archives of Disease in childhood*, 1966, 41, 454-471.

Tanner, J.M., Whitehouse, R.H. and Takaishi, M. (1966b) Standards from Birth to Maturity for Height, Weight, Height Velocity, and Weight Velocity: British Children, 1965. Part II. *Archives of Disease in Childhood*, 1966, 41, 613-627.

Tanner, J.M. and Thompson, A.M. (1970) Standards of birthweight at gestation periods from 32 to 42 weeks, allowing for maternal height and weight. *Archives of Disease in Childhood*, 45, 566-569.

Tanner, J.M., Goldstein, H., and Whitehouse, R.H. (1970) Standards for children's height at ages 2-9 years allowing for height of parents. *Archives for Disease in Childhood*, 45, 755.

Tanner, J.M. (1989) *Foetus into Man: Physical growth from conception to maturity.* Second Edition. Castlemead Publications. Ware.

Townsend, P., Phillimore, P. and Beattie, A. (1988) *Health and deprivation: Inequaliity in the North.* Routledge.

Turkheimer, E. and Gottesman, I.I. (1991) Individual differences and the canalisation of human behaviour. *Developmental Psychology*, 27, 18-22.

Waterlow, J.C, Buzina, R., Keller, W., Lane, J.M., Nichaman,M.Z. and Tanner, J.M. (1977)The presentation and use of height and weight data for comparing the nutritional status of groups of children under the age of 10 years. *Bulletin of the World Health Organisation*, 55 (4): 489-498

Wechsler, D. (1981) *The Manual of the Wechsler Adult Intelligence Scale - Revised (WAIS-R) Manual*. Psychological Corporation. Harcourt Brace & Company. London.

Wechsler, D. (1992) *The Manual of the Wechsler Intelligence Scale for Children - Third Edition UK (WISC-R^{UK}) Manual*. Psychological Corporation. Harcourt Brace & Company. London.

Werner, E. E., Honzik, M. P., and Smith, R. S. (1968) Prediction of intelligence and achievement at 10 years from 20 months pediatric and psychologic examinations. *Child Development*, 39, 1063-1075

Whitehead, R.G., Paul, A.A. (1984) Growth charts and the assessment of infant feeding practices in the western world and developing countries. *Early Human Development*, 9, 187-207.

Whitehead, R.G., Paul, A.A. and Cole, T.J. (1989) Diet and the growth of healthy infants. *Journal of Human Nutrition and Dietetics*. 2, 73-84.

WHO (1981) Development of indicators for monitoring progress towards Health for All by the year 2000. *'Health for all' series No4*. Geneva.

WHO Working Group (1986) Use and Interpretation of anthropometric indicators of nutritional status. *Bulletin of the World Health Organisation*, 64 (6): 929-941

Wilcox, W.D., Neiburg, P., and Miller, D.S. (1989) Failure to thrive: A continuing problem of definition. *Clinical Paediatrics*, 28, 9, 391-394.

Wilensky, D.S., Ginsberg, G., Altman, M., Tulchinsky, T.H., Ben Yishay, F. and Auerbach, J. (1996) A community based study of failure to thrive. *Archives of Disease in Childhood*. 75, 145-148.

Winick, M., Rosso, P. (1969) Head Circumference and cellular growth of the brain in normal and marasmic children. *The Journal Of Paediatrics*, 74, 5, 774-778

Winick, M. (1969) Malnutrition and brain development. *The Journal Of Pediatrics*, 74, 667-679

Wolke, D., Skuse, D. and Mathisen, B. (1990) Behavioural style in failure to thrive infants: A preliminary communication. *Journal of Paediatric Psychology*, 15, 2, 237-253

Wright, C.M., Aynsley-Green, A., Tomlinson, P., Ahmedn, L. and MacFarlane, J.A. (1992) A comparison of height, weight and head circumference of primary school children living in deprived and non-deprived circumstances. *Early Human Development*, 31, 157-162.

Wright, C.M., Waterston, A. and Aynsley-Green, A. (1993) Comparison of the use of Tanner and Whitehouse, NCHS, and Cambridge standards in infancy. *Archives of Disease in Childhood*, 69, 420-422

Wright, C.M., Matthews, J.N.S., Waterston, A. and Aynsley-Green, A. (1994a) What is a normal rate of weight gain in infancy? *Acta Paediatrica* 83,351-6

Wright, C.M., Waterston, A. and Aynsley-Green, A. (1994b) Effect of deprivation on weight gain in infancy. *Acta Paediatrica* 83; 357-9.

Wright, C.M. (1995) A population approach to weight monitoring and failure to thrive. In T.J.David (Eds.) *Recent Advances in Paediatrics 13.* Edinburgh: Churchill Livinstone. (pp.73-87)

Wright, C.M., S.S. Corbett, and R.F.Drewett (1996) Sex differences in weight in infancy and the British 1990 national growth standards. *British Medical Journal* 313, 513-514.

Wright, C.M. (1996*) The Parkin Project: A study of screening and intervention in Failure to Thrive*. PhD thesis, University of Newcastle upon Tyne.

Wright, P. and Deary, I.J. (1992) Breastfeeding and intelligence. Letter to the *Lancet*, 339: 612.

Zeskind, P.S. and Ramey, C.T. (1978) Fetal malnutrition: An experimental study of its consequences on infant development in two caregiving environments. *Child Development*, 49, 1155-1162.

Zeskind, P.S. and Ramey. C.T. (1981) Preventing intellectual and interactional sequelae of fetal malnutrition: A longitudinal, transactional and synergistic approach to development. *Child Development*, 52, 213-218.

Zimmerman, I.L. and Woo-Sam, J. (1972) Research with the Wechsler Intelligence Scale for Children: 1960-1970 [Special Monograph Supplement] *Psychology in the Schools*, 9, 232-271.

Review Of Outcome Data

Author	Sample	Criteria for selection	N/Total	Controls	Age at FU	Test	Blind	Cognitive Outcome	Comments
Bithoney (1986)	Referred from primary care	Weight and height <5% NCHS	45	45 Matched for age, sex, SES	Age 18.5m at FU	Denver Developmental Screening Test	No	Denver Failure Rates sub tests ftt ctrl p Personal/social 2/45 1/45 ns Fine Motor 0/45 1/45 ns Gross Motor 5/45 0/45 <.02 Language 9/45 0/45 <.002	Ftt group more anaemic with elevated lead levels in ftt group ($p<0.003$), but no sig. relatonship exists between lead, anaemia and Denver Dev. Scr. test.
Boddy (1997)	Population based sample. Annual birth cohort (n=2510) in inner city London health district. Preterm, severe IUGR infants excluded. After exclusions and attrition 1554 subjects remaining	Weight SDS below -1.88 SD by 12 months of age and sustained for 3 months or more. No organic disease	42/49	Matched for age, sex, ethnicity, birth weight (to within 300grm) and ordinal position. SES was matched on the basis of geogrphical proximity.	6 years	McCarthy Scales	blind	Cases Controls Wilcoxon p GCI 99.8 103.7 -1.59 <0.10 Memory 47.6 50.7 -1.73 <0.05 Verbal 48.8 49.2 -0.6 ns Quantitative 47.5 52.2 -2.34 <0.01 Perceptual 52.3 53.9 -1.23 ns	No sig. difference between early and late falterers or between small or large cases at follow up. Sig. covariates of McCarthy scales in cases WAIS SES Eating disorders Marital disharmony
Chase and Martin (1970)	Hospitalised in first year of life	not specified	19	19 Matched for age, birth weight, sex, race, SES retrospectively	3 - 4 years after hospitalisation	Denver Developmental Screening Test WAIS HOME (Caldwell 1967) San Mateo County Family Social Functioning Scale (1963)	Blind	Denver Developmental Quotients sub tests ftt<4m ftt>4m ctrl p Personal/social 95.7 70.6 100 <.01 Fine Motor 97.1 78.7 105.3 <.01 Gross Motor 97.4 69.6 99.4 <.01 Language 83.7 60.2 88.0 <.01 WAIS no sig. difference. Maternal / Child Interaction. Control mothers scored sig. higher 48 vs 42 ($p<0.01$). Social Functioning - Only Health Status of the family differed significantly ($p<.025$ In 10 children with FTT >4M low indices for weight, height, and head circumference were more frequent. Significant effect of treatment for ftt after 4m. Their mean DQ was 70.	Paternal separation during pregnancy, alcohol related problems, inadequate money, many young sibs., birth spacing, and more unwanted pregnancies. No relationship with maternal age, parental education, families with deviant behaviour, parental separation except during pregnancy.

Author	Sample	Criteria for selection	N/Total	Controls	Age at FU	Test	Blind	Cognitive Outcome	Comments
Dowdney, Skuse, Hepinstall, Puckering, Zur-Szpiro (1987)	Population based. 2145 live births registered. 1980 traced. 650 moved away. 26 died. 564 moved into area. Study pop. 1868. Caucasians only studied (n=1,233)	Last recorded weight <10% (n=138). Exclusions, prem. birth, congenital defects, current weight and/or height >10% (n=61 remain) Paediatric examination and interview with the mother, weight and height <10% and below 10% after allowing for height of both parents. Final study group (n=25).	23/25	23 matched for ordinal position, sex, gestational age, ethnic origin, and birth weight. To allow for SES controls chosen from the same clinic or nursery.	4 years of age	McCarthy Scales of Children's Abilities (1972). The Lowe and Costello Symbolic Play Test. (1976)	Bind	McCarthy Quotients ftt ctrl p GCI Scores 77.1 97.7 <0.001 Verbal Score 36.1 46.7 <0.001 Perceptual 42.7 51.9 <0.001 Performance Scores Scales ftt ctrl p Quantitative 39.1 48.4 <0.002 Memory 37.4 47.6 <0.002 Motor 37.9 49.6 <0.001 Symbolic Play ftt ctrl p play age 31.5m 33m NS	Ordinal position - GCI Score for first born cases were on average 16 point s better than late born cases. GCI Scores were not significant when correlated with weight for age and height for age.
Drotar et al (1985)	Hospitalised during first year of life.	1. Weight<5% NCHS 2. Deceleration of weight gain. 3. No organic cause. 4. catch up growth in hospital. 5. BW >1500grm 6. Weight appropriate for gestational age. 7. No physical or neurological disability. 8. Head circumference above 5% and above weight centile. 9. No Abuse. 10. 1-9m of age. 11. Not in foster care. 12. Living within one hours drive of the hospital.	69/80	Within case intervention study. 1. Advocacy. 2. Family Centred. 3. Parent centred.	18m & 24m of age at follow up.	Bayley Scales Infant Behaviour Record (Bayley, 1969) Language Ability (White et al, 1978) Symbolic Play Test (1978) Strange Situation (Ainsworth, 1978) Home Environment Scale (Caldwell & Bradley, 1980)	Blind	Age the child reached the 5th centile predicted Bayley MDI @ 18m and 24m (p<.003 and p<.0004). The younger the child the greater the deficit at FU. Language development was predicted by ratio of adults to children and fewer family stresses No attachment differences except those with more insecure attachment had more chronic ftt. No effects of treatment group	Age at onset and chronicity were both important predictors for Bayley MDI at 18m and 24m. Growth centile at testing bw intake 18m 24m height 32 20 27 23 weight 37 5 27 27 Head C 83 25 38 35 .

Author	Sample	Criteria for selection	N/Total	Controls	Age at FU	Test	Blind	Cognitive Outcome	Comments
Drotar & Sturm (1988)	Hospitalised during first year of life.	As above	59/80	As above	36m of age at follow up	The Stanford Binet Intelligence Scale (Terman and Merrill 1972)		Bayley Scales at intake was 99.6 and mean IQ at follow up 85.4. No sig. effect of intervention. Sig. correlation of age at onset with IQ r=.38 p<.005. Environmental factors accounted for 22% of variance (p<.05) Severity of fit and previous Bayley score were not significant pedictors.	Onset and duration result in sig. increment to variance (p<.05)
Elmer et al (1969)	Hospitalised	Weight <3%. Height <3% in 13 cases and <10% in 2 remaining cases.	15	none	Aged 3years 3m to 11years 7m at follow up.	Oppenheimer Scale for mental dev. Willoughby and Haggerty for behavrl problems. Lapouse method for assessing behaviour of school children.	Not blind	5 children were normal mentally, 6 mildly retarded, 4 moderately retarded. 6/7 school age children were in special education classes. Slow speech development and problems with conceptual thinking.	3/8 preschool children and 4/7 school age children with behavioural disturbances. Only 2 out of these 7 mothers perceived them as difficult.
Field (1984)	Hospitalised	Physician's diagnosis When charted 15 were <5th centile for weight.	17	none	1m, 3m, & 14 infants were seen at 6-13m after discharge	Bayley Scales (1969)		Bayley Scales MDI adm discharge 1m 3m 6-13m 83.3 87.5 88.6 96.9 95 Motor Quotient 80.5 87.7 89.1 84.9 85 Physical and developmental measures had reached normal range at FU.	There was a relationship between improved weight and MDI, but no effect of age at onset was found. 9 families rated as having financial, personal, and caretaking stress. 8 were not. There were routine intervention.
Glaser et al (1968)	Hospitalised	BW>5lbs Weight<3% Stuart growth Charts	40/50	none	8m-8yrs average age 4.5yrs	Not specified	No	Normally distributed results with mean value between 90-100. 6 cases classified as mentally retarded. But also 6 with IQ above 120.	Some emotional disorders reported but 1/3 of children had no emotional, physical or psychological disorders. Most families intact, stable with steady incomes.

Author	Sample	Criteria for selection	N/Total	Controls	Age at FU	Test	Blind	Cognitive Outcome	Comments
Hufton and Oates (1977)	Hospitalised	Weight<10%	21/30	none	6yrs 4m after discharge	Wisc (only given to 14 children). Graded Reading Vocabulary Test. MMPI for mothers	No	One child's Wisc Full Scale score was >110, three had scores less than 90 points. The remaining 10 had scores between 90-110. Three of these had performance scores 20 points higher than verbal scores. 2/3 had delayed reading age. Teacher's evaluation of emotional status classified 10 children as having abnormal personalities MMPI showed 13 abnormal profiles, 3 had attempted suicide.	.5 children remained below the 10th centile for weight and 1 below the 10th centile for height. 12 fathers were still married and with the family, 1 was schizophrenic. 9 families were on welfare, 10 had financial difficulties, 8 homes had 3 or more than people per bedroom.
Mitchell, Gorrell & Greenburg (1980)	Population based. N=312 in three rural primary care centres.	80% of normal (i.e.normal is 50th percentile of Stuart growth curves)	12/30 had complete examination	Matched controls (16 had complete examination)	3-6yrs of age	McCarthy Scales Behaviour Problems Questionnaire Coddington Life Events Scale for Children.	Blind	McCarthy Scales Subtest ftt ctrl P Verbal 43.5 46.4 NS Percept/perf 42.7 46.5 NS Quantitative 43.7 43.5 NS Memory 44.4 47.5 NS Motor 44.6 49.0 NS GCI 87.5 92.5 NS Behaviour problems were not noted more frequently. No difference in Life Events scores	Power calculations, do not appear to allow for the effect of a high rate of attrition. Social turmoil was a strong predictor of developmental deficits and the entire clinic population was deprived. OM more frequent in the case group but no other illness was more common. Mean height 5% below controls and mean weight <15% of controls. BW and prematurity did not differ but neonatal problems were more frequent in cases.

Author	Sample	Criteria for selection	N/Total	Controls	Age at FU	Test	Blind	Cognitive Outcome	Comments
Oates, Peacock, Forrest (1984, 1985)	Hospitalised. (See Hufton and Oates, 1977)	Weight<10%	14/30	Comparison with 39 abused children and controls taken from the same school matched for age, sex, social class and ethnic group..	13 years after admission to hospital	- WISC-R (1974) - Piers-Harris Children Self Concept Scale. (Piers, 1976) - Diagnostic and attainment testing (Schonell, 1952) - Vineland maturity scale (Doll, 1953) - Cattell's High School Personality Questionnaire. - Verbal Language Development Scale (Mecham, 1971) - Behaviour questionnaire Rutter (1967) - Personality test for mothers	Blind	WISC Subscales fit abused ctrl for fit p Verbal 90 95 102 <.02 Performance 98 95 104 .27 Full scale IQ 93 95 103 .06 No sig. effect of the Piers Harris Scale Reading age fit ctrl ±12m of chron age 3 7 13-24m behind 1 4 25-36m behind 2 2 >36m behind 8 1 Verbal Language fit ctrl p mean quotient 80 91 0.005 Social Maturity was sig. lower in the fit group (fit social maturity quotient 98 vs. controls 107, p<0.04) Case mothers scored lower on the personality test. Personality scores have been found to correlate with IQ.	Scores for abused children differed on IQ and social maturity. Weights of the children were all >3rd centile and only one child had height <3rd centile. Their anthropometric measures were not sig. different from the controls. Control mothers were more informed about their child's education.
Singer & Fagan (1984)	Hospitalised at least once before ages 5-9m	Weighed under 3rd centile for their conceptional age at testing.	13	13 children with Organic Failure to Thrive and 13 not failing to thrive Groups did not differ sig. on gestational age, post natal or conceptional age, birth weight, parental education.	20m & 3yrs	Visual Recognition Memory Bayley Scales Both administered after Noft and OFT groups condition had stabilised.and Bayley Scales again at 20m. Stanford-Binet Intelligence Scale at 36m.		VRM and Bayley MDI - initial assessment. Group % to novelty MDI NOFT 68 77.6 OFT 52.6 67.7 CTRL 62.5 120.2 Group MDI at 20m. IQ at 3yrs NOFT 80.5 78.6 OFT 72.2 67.7 CTRLS 109 97.4 MDI HOME REARED PLACED NOFT 90 70.2 OFT 83.3 67.4	At initial assessment Visual Recognition Memory of non-organic failure to thrive infants was found to be equivalent to controls. NOFT and OFT did equally poorly on MDI lagging behind controls (p<.001). This suggests that in NOFT infants delay is attributable to environmental factors rather than impaired CNS functioning. Outcome was most associated with parental educational level for all groups.

Author	Sample	Criteria for selection	N/Total	Controls	Age at FU	Test	Blind	Cognitive Outcome	Comments
Singer (1986)	Hospitalised	Weight <3rd centile of NCHS standards	29	No controls Three groups: 1. Sick in care 2. In care 3. Not in care	3 yrs of age	Stanford -Binet Intelligence Scales (n=11) McCarthy Scales (n=12) Recentlly assessed using unspecified standardised IQ tests (n=6).	No	Group IQ p 1. 78.3 NS 2. 84 NS 3. 79.2 NS	All low SES. Cognitive deficits may be explained in terms of low SES or possible continuing health problems but later health status is not documented.
Skuse, Reilly, Wolke (1993)	Population based study. Annual birth cohort (n=2510) in an inner city London Health District.Preterm, severe IUGR, infants were excluded. 1554 subjects remaining after exclusions and attrition.	Weight Z score below -1.88 (3rd centile) by 12m of age and sustained for 3m or more. No organic disease.	49 Subsets early and late growth faltering.	Comparison of infants with late (n=25) and early (n=22) growth faltering.	15m of age	Bayley Scales (1969) Mental Development Index and Psychomotor Development Index	Blind	Bayley MDI Grand mean 98.2 Growth trajectory explains 21% of variance of MDI after maternal IQ. PDI Grand mean 96.7 Growth trajectory explains 15% of variance of PDI after maternal IQ (Maternal IQ is not sig. associated with PDI) Early FTT is sig. associated with lower MDI scores. Late FTT did not differ from a normal comparison group. 4 subset WAIS given to the mothers. Mothers of late falterers had IQ .75 of a standard deviation lower than early falterers.	Late falterers had lower standardised BW. GHQ scores of late faltering mothers were higher than those of early falterers (5.8 vs 2.9, p=0.039). Social support scores were higher for mothers of early fit babies (early 31, late 25.7, p=0.001) In the early fit group only 35% of siblings had definitely grown normally in their first year. In the late fit group 57% grew normally in their first year..
Skuse, Pickles, Wolke and Reilly (1994)	As Above	As above	47/49	Matched for age, sex, ethnicity, birth weight (to within 300grm) and ordinal position. SES was matched on the basis of geographical proximity	15m	Bayley scales WAIS Tester's Rating of Infant Behaviour (Wolke, 1987) Nursing Child Assessment Teaching Scales HOME Scale Schedule for Oral Motor Assessment (Reilly et al, 1994)	Blind	Bayley Scales FTT CTRL P MDI 98.2 108.5 PDI 96.7 103.6 Sig. correlation for MDI @ 15m and standardised weight at 6wks, 3m, 6m, 9m, (r=.31, .45,.42, .33 respectively). TRIB no sig. differences HOME did not explain more variance once NCAST scores had been entered. SOMA scores higher for cases and correlated with MDI r = -.38. Negative correlation between SOMA and MDI found in the control group.	No sig. correlation was found between duration and mental outcome. Neither BW nor weight at 15m was correlated with MDI or PDI. Head circumference did not correlate with outcome. Predictive model devised taking into account severity, timing and duration of FTT A fall in weight of -2SD after birth = -10 MDI/PDI points. No deficit for a fall after 8m

Author	Sample	Criteria for selection	N/Total	Controls	Age at FU	Test	Blind	Cognitive Outcome	Comments
Wilensky (1996)	Community based non-referred sample. Cohort (n=1452) born in 1991 in 3 neighbourhoods in Jerusalem with mixed SES. Exclusions: BW<2500grm or before 37 wks and infants whose weight/height >10th centile	<3rd centile for weight (NCHS) for at least 3 months.	50/55	Matched from the same clinic, born in the same month, same maternal education, ethnic origin, maternal age, parity, infants BW.	Records reviewed @ 15m, tested @ 20m. 3/5 still below 3rd centile for weight. At 25m the Bates, Home, and maternal psychiatric interview was completed.	Bayley Scales Bayley Infant Behaviour Record HOME Bates Temperament Questiopnnaire.	blind	Bayley Scales FTT CTRL P MDI 99.7 107.2 <.05 More FTT children scored below 80 (11.5% vs 4.?%) Maternal education and Home Score Positively associated with MDI in FTT group, but only the HOME in the control group. Age at onset and weight at 20 months were not predictive of MDI, but only11 children fell below 6 months. In behavioural development, FTT children were less sociable <.05 and more fearful of the examiner p<.05. HOME scales showed that FTT homes were lower in verbal and emotional reactivity, active stimulation, and fam,ily participation in encouraging development. No difference was found in the Bates Temperament Questionnaire. The mother's psychiatric status did not differ.	Prevalence of FTT 3.9% Mean age at onset 8.4m BW of cases was marginally smaller. No differences in maternal age, birth order, maternal education, obstetric or neonatal problems, or ethnicity. Cases had marginally more medical problems (p<.052) and were twice as likely to be admitted to hospital. No difference in OM Similar proportion were breast fed but FTT children were breast fed longer (15.3 vs 11.4 wks, not sig.) and longer per feed (19.5 vs 12.9 mins. p<.05). FTT fell asleep more often during a feed and had more feeding difficulties. Paternal age was sig. associated with FTT (p<.01).

ID ☐☐☐☐

DATE of interview ☐☐ day ☐☐☐ month ☐☐ year

Interview schedule for STUDY OF CHILDREN'S GROWTH OUTCOMES

I would like to ask you a few questions about (child's name). The questions are in two sections. Section A is about your child's health from before birth up to the present day. Section B is about you and your family.

SECTION A

1) General information about your child

Your child's surname.....

Other names.....

Address

Street.....

Town.....

County.....

Postcode.....

Child's date of birth ☐☐ day ☐☐☐ month ☐☐ year

Child's age ☐☐ years

Male/female ☐ enter M or F

Before beginning please could you specify your relationship to the child?

☐

Natural mother (enter 1)
Grandmother (enter 2)
Foster mother (enter 3)
Adoptive mother (enter 4)
Other (enter 5) Please specify

Have you always been the only carer at home or has your child been regularly cared for by people other than yourself?

enter Y (yes, only carer) or N (no, others) ☐

If no, who was the main carer?

☐

Mother was the main carer (enter 1)
Grandmother (enter 2)
Relatives (enter 3) (Please specify)
Other (enter 4) (Please specify)

And for how long?

☐ years ☐☐ months

2) About your child's birth

Thinking back to your pregnancy can you recall if there were any medical problems?

enter Y (yes) or N (no) ☐

If so, can you describe these to me?

Was it a difficult birth?

enter Y (yes) or N (no) ☐

If so, can you tell me about it? (length of 2nd stage of labour, forceps, caesarian or SUD)

Did the baby arrive early, or on time?

enter 1 (early), or 2 (on time) ☐

If early how many weeks early?

enter number of weeks pregnant ☐☐
(Number of weeks=40 - number of weeks early)

Did you smoke whilst you were pregnant?

enter Y (yes) or N (no) ☐

Do you smoke now?

enter Y (yes) or N (no) ☐

3) About your child's health
Firstly, I'd just like to ask you a few questions about how you fed your child in infancy.

Did you ever breastfeed your child?

enter (Y) even if only once or (N), never ☐

If yes can you tell me how long you breastfed? ☐

<1 week	(enter 1)
1-6 weeks	(enter 2)
6 weeks-4 months	(enter 3)
more	(enter 4)

Were there any feeding problems with your child?

enter Y (yes) or N (no) ☐

If so, can you describe these? (Problems sucking, swallowing or chewing are particularly relevant)

Could I just ask about (child's name) general health.

Has your child had any serious problems that lead to him being treated in hospital or as an outpatient?

enter Y (yes) or N (no) ☐

Please describe these.

What age was your child when seen? ☐☐ ☐☐
years months

What was the name of the hospital or hospitals where your child was seen?

What was the name of the consultant or consultants who saw your child?

Has your child had any problems with his sight?

enter Y (yes) or N (no) ☐

Has your child had any hearing difficulties?

enter Y (yes) or N (no) ☐

Has your child any speech problems?

enter Y (yes) or N (no) ☐

Has your child ever been referred to a speech therapist?

enter Y (yes) or N (no) ☐

If so how long was your child seen by the speech therapist?

less than a year (enter 1)
more than a year (enter 2) ☐

Does anyone in your family have speech or reading difficulties?

enter 1 (speech) or 2 (reading) or 3 (both) ☐

If so who in the family is affected?

4) About child care

Did your child attend any of the following before going to school?

Playgroup

enter Y (yes) or N (no) ☐

Nursery school

enter Y (yes) or N (no) ☐

How long did your child attend playgroup or nursery school?

enter years months

How many days per week?

Attended 4 or more days per week (enter 1)
Attended less than 4 days per week (enter 2)

SECTION B

1) About you

To which of these ethnic groups do you belong? Tick the appropriate box.

Caucasian (enter 1)
Afro-Caribbean (enter 2)
Indian subcontinent (enter 3)
Ethnic Chinese (enter 4)
Other (enter 5) (please specify)

Did you leave school at 16 or did you stay on at all?

enter Y (yes) or N (no)

If you stayed on, where did you complete full time education?

Sixth form (enter 1)
Further Education College (enter 2)
Institute of Higher Education (enter 3)
(ie. University, Polytechnic)
Other (enter 4) Please specify

Did you get any qualifications? Tick the appropriate box.

None (enter 1)	☐
'O' level (enter 2)	
'A' level (enter 3)	
Diploma (enter 4)	
Degree (enter 5)	
Other (enter 6) Please specify	

2) About your family

Can I ask for details of your children?
What is their date of birth, age and sex?

	DOB	AGE	SEX
1.			
2.			
3.			
4.			
5.			
6.			
7.			
8.			
9.			
10.			

So there are (Number) of children in your family?

enter number of children ☐

And (target child) is your (birth order) child?

enter birth order ☐

Is (target child) a twin?

enter Y (yes) or N (no) ☐

Are you employed outside the home?

enter Y (yes) or N (no) ☐

If so, is your work full time or part time?

Enter F (full time), enter P (part time). ☐

Is anyone else in the house employed?

enter Y (yes) or N (no) ☐

Do you and your household own or rent the house or flat where you live?

☐

Own (or are buying) enter 1
Rent privately enter 2
Rent a council house enter 3

Does your family own a car?

enter Y (yes) or N (no) ☐

Please estimate the height of your child's natural father.

☐ ☐
Ft ins.

Can I just measure your height?

☐☐☐
enter cms

PARENT'S CONSENT FORM FOR CHECKING MEDICAL RECORDS OF THEIR CHILD IN THE GROWTH STUDY

Name of parent or main carer.....

Address.....

.....

.....

Name of child.....

I am willing for a member of the Growth and Development Study to check the medical records of my child.

I understand that the information will be treated confidentially, and will not be shown to anybody outside the project team.

SIGNED.....

Head Teacher
Address

Dear

I am writing to ask if you could allow us to visit your school to see some children who are enrolled in a research project. The project is a collaboration between the University of Durham and the Department of Child Health at the University of Newcastle Upon Tyne. It has been approved by the Newcastle Health Authority Ethics Committee, and by the City of Newcastle Director of Education.

We are planning to study children, now about eight years old, some of whom were identified as having poor growth in infancy. Some schools helped us with an earlier study of this kind which we carried out in the West End of Newcastle two years ago. During this study we found a relationship between poor growth and subsequent cognitive development.

A small number of the children we would like to follow up are now attending your school, and we would like to see them each individually for about two hours altogether, firstly, to assess their cognitive development using the Weschler Intelligence Scales for Children, and then at a later date to assess reading ability, articulation and fine motor skills. All testing will be carried out by psychologists. Although this will not involve any extra work for your teaching staff we would like to ask to use the schools as a location for testing, as there is more consistency in the testing environment than if the tests were carried out at home.

I am sorry to have to trouble you with this, but it is important that we see every child in the group, and we will do our best to cause you as little inconvenience as possible. I would be glad to discuss the project with you personally, and will telephone you in a few days time.

Yours sincerely,

Mrs. Sally Corbett

For project research group

Mrs. Sally Corbett
Dr. Robert Drewett
Dr. Charlotte Wright

The Parkin Project
3rd Floor,
Shieldfield Health Centre,
Stoddart St.,
Newcastle Upon Tyne.
NE2 1AN
Tel. 091 261 0789

Dear Mrs.

In a study carried out by The Parkin Project the weights of all children in Newcastle born between 1987 and 1988 were collected to find out how children normally grow. We would like to see a number of these children now that they are older, to see how they are getting on. The children selected represent different patterns of growth in infancy. We would be most grateful if you would agree to help us by participating in this study.

I would like to see you for about an hour and a half to collect some information about you and your child's health. I would also like to visit your child at school to assess their progress. The tests used in school are carefully designed to be suitable for children this age whether they do well at school or not. All information will be treated confidentially and will not be shown to anyone outside the project team.

This is a city wide project and many children in each school are likely to have been selected. Each school will be contacted personally before your child is seen.

I would be grateful if you would fill in the enclosed slip and return it in the reply paid envelope. I look forward to hearing from you and if you are able to help us, will call to see you personally to explain the project further.

Yours Sincerely,

For project team
Mrs. Sally Corbett
Dr. R.F. Drewett
Dr. C.M. Wright

PLEASE FILL IN AND RETURN THIS SLIP IN THE ENCLOSED ENVELOPE
To Sally Corbett, The Growth and Development Study, 3rd Floor, Shieldfield Health Centre, Stoddart St. Newcastle Upon Tyne, NE2 1AN.

NAME OF CHILD.....

ADDRESS.....

.....

.....

Please tick the appropriate box below

I am willing to be contacted and to participate in the project ☐

I am not willing to be contacted and to participate in the project ☐

SIGNED.....

The Growth and Development Study
3rd Floor,
Shieldfield Health Centre,
Stoddart St.,
Newcastle Upon Tyne.
NE2 1AN
Tel. 091 261 0789

Dear

Recently you kindly agreed to participate in a study of children's growth and to allow me to visit (child's name) in school. You will be pleased to hear that we have seen (child's name) and now I would very much like to call on you to collect some more information. As I mentioned before, none of the information you give me will be shown to anyone outside the project without your permission.

I shall need to see you for about an hour and a half and I wondered if

(day, date, at time)

would be convenient. Please fill in the enclosed form and return it in the reply paid envelope to let me know if this is a suitable time to call.

Yours sincerely,

Sally Corbett

For project team
Mrs. Sally Corbett
Dr. R.F. Drewett
Dr. C.M. Wright

Please fill in the form below and return in to me in the reply paid envelope.

Name of child.....

Address.....

.....

.....

School.....

Please tick the appropriate statement below and add the information to enable me to contact you at a convenient time.

- The time suggested is convenient ☐

- This time is not convenient please contact me to arrange an alternative time. ☐

  My telephone number is:.....

- This time is not convenient. ☐
  However I am usually available on the following day and/or times:.....

- This time is not convenient. ☐
  Please send an alternative appointment time.

QUESTIONNAIRE INFORMATION ABOUT HEALTH

Record number	Code MEDHIST	Code COGMHIST	Problems during pregnancy	Problems at birth	Medical History	Age of admission
002	1	2			Tetralogy of Fallot Recurrent chest infections Progressive cyanosis Open heart surgery Hole in the heart RTA/ Broke his legsX 2	 2 yrs 5 yrs
003	3	3		Caesarian	Viral infections Lumo removed from navel Grommets	2 month 16 month 6 yrs
005	5	3	Threatened miscarriage		Stitches in head Behaviour problems	18 month 6 years
007	3	3			Gastroenteritis Displaced hip	10 month 1yr
009	5	3		Forceps Long second stage	Vomiting/Dehydration	3yrs
010	2	3			Multiple ear infections Grommets Tonsils/Adenoids	 1months 7 yrs
011	5	3	Toxaemia	21 hour labour caesarian	Referred for growth problems Nothing found	2yrs
012	5	3		21hour labour		
013	5	3	High blood pressure	Induced Forceps 17 hour labour		
014	5	3		Forceps In hospital 1 month afterwards		

Record number	Code MEDHIST	Code COGMHIST	Problems during pregnancy	Problems at birth	Medical History	Age of admission
016	5	3		Forceps Cord round neck		
020	3	3			Vomiting Convulsions	6 months 1yr
021	3	3			Colic Chest infections Grommets	3 months 6 months
022	5	3	High blood pressure			
024	5	3			Colic	3 month
025	5	3			Pneumonia	6yrs
026	3	3			Gastroenteritis Failure to thrive	15 months 1yr
028	5	3		Caesarian	Head injury/minor	6yrs``
029	5	3		Induction Forceps		
030	2	3	Placenta praevia	Caesarian Baby distressed	Recurrent ear infections Grommets	 1yr
031	5	3	High blood pressure / nausea			
033	5	3		Forceps	Constipation, frightened to go to the toilet	5yrs
035	5	3		Caesarian/Breech		
036	4	3		Planned caesarian	Chicken pox / severe skin problem	2yrs
038	5	3			Tested for grommets / not treated	3yrs
039	5	3	High blood pressure	Caesarian	Grommets	6yrs
041	5	3			Check for possible growth in throat /Nothing found	1wk

Record number	Code MEDHIST	Code COGMHIST	Problems during pregnancy	Problems at birth	Medical History	Age of admission
044	5	3			Undescended testicle Grommets Broken collar bone Broken wrist	2yrs 4yrs 6yrs
046	5	3			Grommets	5yrs
047	3	3	Blood in urine	Breech Forceps	Feeding problems Spina bifida/bleeding from bowel	6wks 5yrs
048	3	3		Planned caesarian	Feeding problems Gastroenteritis	9months 9months
050	5	3		Forceps		
053	5	3		Forceps	Broken les	1yr
054	5	3		Ventouse		
056	3	3		Forceps	Grommets	2yrs
057	3	3	Placenta praevia , Haemorrhage, Labour started but prevented.		Virus that affected balance	18 months
058	5	3	Threatened miscarriage		Convulsions/ high temperature Accidental injury x 2	2yrs
059	5	3		Forceps		
060	5	3			Several accidental injuries has been kept in hospital	
061	5	3			Large head circumference, not treated	1yr
063	5	3	Bleeding x 3		Fingers fused together Broke collar bone	1, 2,and 3 months 2yrs 21months
064	5	3			Accidental injury	4yrs
065	3	3		Long labour	Accidental injury excema	7yrs

Record number	Code MEDHIST	Code COGMHIST	Problems during pregnancy	Problems at birth	Medical History	Age of admission
066	2	1		Long second stage /55mins	Septcaemia at birth Mild cerebral palsy	Birth
067	5	3			Iinfantile puerpera Platelet count 10 - not treated	18months
068	5	3	Chicken pox@6wks gestation			
069	3	3			Vomiting	5months
071	3	3	Ulcerative colitis		Meningitis	1yr
072	5	3		Caesarian Breech		
073	5	3			Tonsilitis	7yrs
074	3	3			Urinary infection Hearing problems - build up of wax Cyst removed on arm Cracked elbow	
077	5	3			Bell's palsy Broken collar bone	4yrs 6yrs
078	5	3	Lost weight			
079	5	3	Swollen legs Back ache	Passed out Post antal depression		
080	5	3			Grommets	4yrs
081	5	3	Placenta praevia			
082	3	3		Caesarian	Asthma/Bronchiolitis Heart murmur Broken collar bone	6months
083	3	3	Nervy Sickness	Induced Long labour Forceps	Asthma Excema	3 months 3 yrs
084	5	3	High blood pressure	Induced	Dyspraxia	4yrs

Record number	Code MEDHIST	Code COGMHIST	Problems during pregnancy	Problems at birth	Medical History	Age of admission
085	5	3			Grommets	5yrs
087	5	3			Burn on the hand Cut on the forehead Uunconscious once	6yrs
088	5	3			Displace hip / wore special shoes	5yrs
089	5	3			Allergic reaction to penicillin x 2	2 months 3 years
090	5	3		Caesarian	Undescended testicle	2yrs
093	5	3		Mother lost blood		
094	2	3	Stomach ulcers	Caesarian	Severe hiatus hernia Chest infection (croup, pneumonia, pleurisy) Jaundice	15 months
097	5	3			Grommets	2yrs
099	5	3	Anaemia		Warts on buttocks	18months
100	2	3		Long labour Baby's heart slowed	Bronchiolitis / Persistent wheeze Head injury	5 months 3years
101	3	3			Bronchiolitis	10months
102	5	3			Bump on the head	5yrs
103	5	3			Head injury Burn on the hand Dog bite	3yrs 4yrs
105	5	3		Caesarian		
106	5	3			Operation to correct squint	2 1/2 yrs
107	5	3			Split head open	
108	5	3	Migraine	Long second stage of labour	Failure to thrive	18months
109	5	3		Caesarian/ Twins		

Record number	Code MEDHIST	Code COGMHIST	Problems during pregnancy	Problems at birth	Medical History	Age of admission
110	2	1		Mother discharged herself during labour/ baby born blue/ takenn to ICU	Perinatal hypoxia, Microcephaly ASD Pulmoinary stenosis - not major	Birth
111	3	3		Caesarian	Pyloric stenosis	2months
112	5	3			Fractured tibia/Fell off bike	2 1/2 months
113	5	3	High blood pressure			
114	5	3			Cystitis/ baby excema/ caused by bubble bath	4yrs
115	5	3			FTT	4yrs
116	5	3	Sciatica			
117	5	3		FORCEPS		
118	5	3	Unspecified	Baby not breathing when born Induced labour		
119	4	3			Heart murmur	1yr
120	3	3	Urinary infections Poor appetite Weight loss	Long labour Forceps Baby small	Bronchiole pneumonia Hand burned	7months 3yrs
122	4	3		Forceps	Hernia/Lost weight Grommets	1yr 3yrs
124	5	3		Caesarian		
125	2	3	Toxaemia Nausea to 3 months	Caesarian	Recurrent UTI Renal scarring Reflux on one kidney Sight poor in one eye	7months 3yrs
126	5	3	Didn't gain weight			
129	5	3			Weight loss	1year

Record number	Code MEDHIST	Code COGMHIST	Problems during pregnancy	Problems at birth	Medical History	Age of admission
131	3	3			Convulsions/Epilin given for several months	8months
135	5	3	Bleeding		Croup	2yrs
137	5	3	Bleeding Placenta praevia	Presentation face up	Swallowed Brasso Cut under the chin	18months
138	5	3	Transverse presentation	Caesarian		
139	3	3			Intussusception Grommets	6months 5yrs
140	5	3			Failure to thrive	2wks 16months
141	5	3		Cord round the baby's neck	Plaster on leg/infectionin the bone	4yrs
142	5	3		Long labour Baby bruised Face up	Broken arm x 2	4.6yrs 5.11yrs
143	5	3		Forceps		
145	5	3		Long second stage of labour/ Forceps	Broken wrist	6yrs
146	5	3	Threatened miscarriage	Induced		
147	5	3		Breech Forceps	Splint for hip problems	1wk
150	3	3	Threatened miscarriage		Mild spina bifida	1wk and for regular checks
151	5	3		Caesarian		
152	5	3			Grommets/Adenoids ADD	6yrs 8yrs
153	5	3		Caesarian	Accidental injury	7yrs
154	5	3	Toxaemia	Caesarian		

Record number	Code MEDHIST	Code COGMHIST	Problems during pregnancy	Problems at birth	Medical History	Age of admission
155	5	3		Long labour SUD		
156	3	3			Bronchiolitis	8months
158	3	3	Not gaining weight	Small baby 4LB 1OZ	Grommets	9months
159	5	3	Toxaemia Proteinuria			
160	5	3		Face up/ long labour	Referred to growth clinic	4 I/2 yrs
161	5	3			Asthma	4yrs 10 months
162	2	3	High blood sugar		Osteomyelitis for 2 months	9 months
164	5	3		Forceps Long second stage of labour	Tonsillectomy	6yrs
165	4	3	High blood pressure during the last 2 weeks		High temperature / high blood glucose/ MRI scan for kidneys Croup	9 months 5yrs
166	5	3			Blocked tear ducts Circumcision	18 months 3yrs
167	5	3		Caesarian		
169	3	3			Asthma	2yrs
170	5	3			Temperature fir x 2	3yrs 5yrs
171	2	3			Recurrent UTI Renal scarring	1yr

Record number	Code MEDHIST	Code COGMHIST	Problems during pregnancy	Problems at birth	Medical History	Age of admission
172	5	3	Vomiting and nausea throughout pregnancy / poor weight gain Placenta Praevia	Long labour Caesarian	Slow growth	1yr
173	5	3	Low blood pressure Anaemia		Broken femur	1yr
174	5	3			Banged his head / fractured his skull	18month
176	5	3	Anaemic High blood pressure		Abcess on neck	2yrs 2 months
177	5	3			Croup Cyst on eye	18months 6yrs
179	5	3		Placenta praevia Haemorrhage Caesarian		
180	5	3	Lack of platelets Placenta Praevia	SUD		
182	5	3			Broken finger	8yrs
183	2	3		Forceps/Long second stage	Moderately severe asthma Frequent ear infections Grommets Broken arm	2 1/2 yrs 1yr 3yrs 10 months

Record number	Code MEDHIST	Code COGMHIST	Problems during pregnancy	Problems at birth	Medical History	Age of admission
184	3	3		Caesarian	Bronchiolitis	4months
186	5	3	Rhesus, but test for antibodies OK		Cut his head Jaw bone growing disproportionately	4yrs
187	5	3			Grommets Tonsils/Adenoids and grommets	5yrs 7 1/2 yrs
188	5	3	Pre cancerous growth in the cells of the cervix			
189	3	3			Asthma Convulsion Grommets	8months 18months 5yrs
190	5	3	High blood pressure from 4-5 months in hospital		Grommets	3yrs
191	5	3		Forceps/Long second stage	Broken arm	6yrs
192	5	3			Circumcision	2months
193	5	3		Caesarian		
195	5	3			Asthma Feet pointing inwards	4yrs 4yrs
196	5	3	High blood pressure Antibodies present		RTA/Broken leg + head injury	2 1/2 yrs
197	5	3			Viral infection	5yrs

Record number	Code MEDHIST	Code COGMHIST	Problems during pregnancy	Problems at birth	Medical History	Age of admission
198	3	3			Viral meningitis Grommets	6months 7yrs
199	3	3			Problems with sternum Asthma Suspected meningitis	6wks 18months 3yrs
200	5	3	Iron injections	Induced labour Foetal distress		
201	5	3		Long second stage	Virus/Diarrhoea/Vomiting for 5 days	6yrs
203	5	3	Anaemia			
204	5	3			Grommets Adenoids Tonsils	3yrs 4yrs 6yrs
205	5	3	Migraine/ nausea	Induced	Croup X 5	4yrs
206	3	3		Long labour Forceps Foetal distress	Checked following foetal distress for 18 months	From birth to 18 months
207	5	3			Squint/ minor operation	4yrs
208	3	3		Long labour Epidural Induction Forceps	Chronic diarrhoea - investigated- no cause stated.	1month
209	3	2	15 scans because baby was small	Short labour	Head injury - Brief L.O.C.	1yr 6yrs
210	5	3			Ear Infection	6yrs
211	5	3	High blood pressure. Wasn't gaining weight	Caesarian	Tonsillectomy Asthma	7yrs 7yrs

Record number	Code MEDHIST	Code COGMHIST	Problems during pregnancy	Problems at birth	Medical History	Age of admission
212	5	3			Irritable Hip x 2	6yrs
213	5	3			Asthma	4yrs
214	3	3			Bed wetting /tests for kidney function (duplex kidney) Tonsils and adenoids	4yrs 5yrs
215	5	3	Lost weight	Induced labour		
218	3	3		Caesarian Foetal distress	Repeated chest infections Infections Admitted to hospital once	2yrs
221	5	3	Pre-eclampsia	Epidural + Sedatives to keep blood pressure low. Induced labour	Failure to thrive	4 months
222	5	3			Appendicitis	6yrs
223	5	3	High blood pressure	Caesarian		
225	2	3	Small baby	Long second stage. Presentation face up	Major scald Recurrent ear infections Grommets	13months 15months
226	5	3		Caesarian	Urinary infection	5yrs
227	5	3	Appendicitis	Retained placenta		
230	3	3	Threatened miscarriage	Induced labour	Grommets Irritable hip Watery stools Nausea Lactose intolerance	5yrs 7yrs
233	5	3			Idiopathic thrombocytopenic puerpera	3yrs
234	5	3		Induced	Feeding problems	3months
235	3	3			Diarrhoea	3months

Record number	Code MEDHIST	Code COGMHIST	Problems during pregnancy	Problems at birth	Medical History	Age of admission
236	5	3	High blood pressure Foetal blood pressure high	Induced labour		
239	3	2			Epilepsy - Duration RX but at least 5 fits	8months
240	5	3	Threatened miscarriage			
241	5	3	Fainted alot during 1st half of pregnancy and at the end			
243	3	3			Recurrent ear infections/ treated with antibiotics	2yrs
244	5	3		Caesarian/ twins		
245	3?	3	Scan showed hole at the bottom of the spine		Problem with thyroid /checked every month grommets / balance problems. Can't grip a pencil	1week
246	5	3	High blood at the end of pregnancy			
247	5	3	Head engaged early - discomfort		RTA	
248	5	3	Hyperemesis / Lost weight			
249	1	2			Pierre Robins syndrome Cleft palate Pharyngeal anomaly Tracheostomy Gastrostomy	1yr 2yrs
252	5	3		Caesarian	Split mouth open Stitches needed	1yr
253	5	3	Bleeding			

Record number	Code MEDHIST	Code COGMHIST	Problems during pregnancy	Problems at birth	Medical History	Age of admission
254	5	3	Bleeding x 3 Blood clots in the placenta	Induced labour	Dysentery	3 1/2 yrs
255	2	1		Induced labour	Di George syndrome / Catch 22 Cleft palate Chromosomal anomaly	Since birth
256	1	1	Knew at 5 months pregnant that foetus had fuid on the brain	Caesarian	Encephalocele Hydrocephalus Major brain malformation	Since birth
257	4	3			Heart murmur / not treated	3yrs
258	3	3	Anaemic		Gastroenteritis Virus/vomitng Glandular fever	9months 7yrs
260	5	3		Baby distressed Forceps Drip		
261	5	3			Gastroenteritis	2yrs
264	5	3			Enuresis	
265	5	3			Telly cracked on his head Grommets	3yrs
266	3	3		Induced labour Long labour- 3 days	Gastroenteritis	1yr

<<TITLE INITIALS HEADTEACHER>>
<<SCHOOL>>
<<STREET>>
<<TOWN>>
<<POSTCODE>>
DATE

Dear <<TITLE headteacher>>,

The Growth and Development Study

During the last two years we have been carrying out a study of children born in Newcastle between April 1987 and March 1988 and now at school. Our aim was to relate their present growth and development to their growth in infancy, as previous studies had shown cognitive deficits and higher rates of school failure in children referred to hospital for growth faltering during the first two years of life.

The children were selected by using weight data routinely collected during infancy. Of the 326 children selected, half had a rate of weight gain in infancy which placed them in the slowest growing 5%, the other half were controls. We successfully traced and tested 257 of these children and their mothers using standard IQ tests and a reading test. We also collected data about the children's medical history and demographic data. In addition we measured the height and weight of a fifth of the children born in the same year in order to aquire normative data for growth of Newcastle children.

At follow up the children whose rate of weight gain was slow in infancy were found to be shorter and lighter, but there was no relationship between their early weight gain and current cognitive aptitude or reading ability.

This has been the largest study of its kind and would not have been possible without the help we received from heads, school secretaries and teachers. We are particularly grateful for your help at a time when changes in schools are adding to pressure of work.

Yours sincerely,

For research project
Dr. R.F. Drewett
Dr. C.M. Wright
Mrs. S.S. Corbett
Mrs. J Callum

6th December 1996

Mrs «surname»
«Street»
«district»
«county»
«postcode»

Dear Mrs «surname»

During the last two years we have been carrying out a follow up study of children born in Newcastle between April 1987 and March 1988. Our aim was to find out if there was a relationship between patterns of growth in infancy and present growth and development.

At follow up, children who grew slowly in infancy were still slightly shorter and lighter than other children, but they did just as well at school. This is very encouraging and as the largest study of its kind it is an important finding.

It would be impossible to carry out this kind of research without your help. We are very grateful for all the help you and «childsname» have given us.

Best wishes,

Sally Corbett

For research group
Dr. R.F. Drewett
Dr. C.M. Wright
Mrs. S.S. Corbett
Mrs. J. Callum

Recno	Case/control	Feeding problems reported on questionnaire
003	Case	Wouln't take milk/ couldn't suck/didn't like solids
006	Case	Vomiting-change to soya milk no problems with solids
011	Case	Faddy eater from 2 yrs old
014	Control	Weight gain slow until changed to bottle
015	Control	V. Hungry baby-switched from breast. Faddy eater
018	Control	Poor eater when on solids
020	Case	Vomiting-given gaviscon
021	Control	Colic
024	Case	Colic
026	Case	Picky eater/limited range of foods
029	Control	Aversion to lumpy foods. Constipation
032	Case	Problems with solids
033	Control	Spat out solid food
034	Case	Only took 2-3oz. Fed frequently
036	Case	Insufficient breastmilk and difficulties changing to cows milk
038	Control	Poor weight gain
046	Control	Hungry baby
047	Case	Vomiting/poor weight gain/switched to bottle/poor bottle feed/referred to hospital.
048	Case	Refused milk from 5wks/given juice instead. Wouldn't eat solids
051	Control	Vomiting/but eat alot
052	Control	Problems when starting solids
054	Case	Distressed when breast feeding/ difficulty sucking
056	Case	Only eats fresh vegetables
057	Control	Slow feeder
062	Case	Wasn,t getting enough breastmilk-changed to bottle. Fussy when eating solids
064	Case	Doesn't like milk, faddy, small portions, doesn't like the texture of meat.
066	Case (Oft)	Tube fed/couldn't suck/slow feeder/probslems chewing
068	Case	2hrly feeds/milk dried@2months wouldn't take bottle
069	Case	Vomiting
072	Control	Not getting enough breast milk
078	Control	Didn't like baby milk/had cereals from early on
079	Case	Small portions/picky
092	Control	Not getting enough/put on bottle twins
094	Case (Oft)	Vomiting/lost weight
097	Case	Changed from premium to plus because he was crying/not satisfied
100	Case (Oft)	Vomiting, slow to gain weight, fussy eater
107	Case	Projectile vomiting, didn't like lumpy solids
108	Case	Stopped feeding after a cold @ 7wks. Problems (faddy) until 5-6 yrs
109	Control	Not interested in food. Only ate enough to sate appetite
110	Case (Oft)	Fed by tube for two days until she took the bottle

111	Control	Vomiting for one week. Pyloric stenosis. Dehydrated
112	Control	Needed to top up feeds with bottles
115	Case	Lost weight. Wouldn't feed. Had to be force fed
117	Case	Sickly. Vomited. Colic. Faddy. Won't eat eggs.
122/n	Case	Didn't take solids until 5 months old. Had a hernia and lost weight
124	Control	Breastfeeding not successful. Switched to bottle. Ok after that
125/n	Case (Oft)	Takes a long time eating. Forgets to chew.
129	Case	Took a long time feeding. Weight dropped when she switched to solids
132	Control	She eats when she is hungry but tends to be faddy
140	Case	Wasn't taking any feeds. In hospital at 12 days
144	Control	Changed from breast to bottle as he was not gaining weight
145	Case	Mother had mastitis. Child faddy eater but is ok now
146	Case	Long feeding time. Faddy
160	Case	Poor weight gain whilst breastfeeding. Does not eat high calorie food. If she gets tired she won't eat. The more hungry she is the less she will eat.
162	Case (Oft)	Vomiting frequently
166	Case	Didn't take large amounts. Just little and often.
167/n	Case	Didn't take as much as elder sib. But gaining weight.
171/n	Case(Oft)	Wouldn't take any other form of milk.
174	Control	Faddy. Lost weight? Between 4 and 6 months
175	Case	Poor eater. Didn't take full bottles. Faddy. Would only eat puddings.
177	Control	Not specified
180	Case	Not a good sucker. Short interval between feeding. Ok with bottle.
183	Control (Oft)	Early feeding difficulties. Problems latching on to the nipple. This was resolved once she got home, but she remained colicky.
187	Case	Lazy feeder, slow
191	Control	Crying and sleeping difficulties, but gaining weight
194/n	Case	Difficulties latching on to the nipple, so changed to a bottle.
195	Control	Small quantities, ie. 1oz of milk at a time. Colic as well. Ok on solids.
198	Case	Vomiting. Ok when weaned.
201	Case	Problems chewing. Choked easily. Weaned at 4 months. Didn't like lumpy food.
203/n	Case	3 month colic
207/n	Case	Problems chewing
209	Case	Very sick. From 9 months to two years didn't gain weight, but eating all right.
212	Case	Took a long time to feed. Ok on solids.

214	Control	Projectile vomiting. Pyloric stenosis. Started solids at one year. Had measles at 18 months which affected appetite. Fussy eater. Tonsils swollen by three years , problems swallowing.
225	Case (Oft)	Weaned at 5 months. Faddy.
234	Case	Small amounts of milk taken. Colic.
235	Control	Stopped taking milk. Had diahorrea. Took a long time to finish bottles.
245/n	Case	Didn't chew very well
247	Case	Didn't take solids. Fussy eater. Food didn't interest her, she'd rather have a biscuit.
249	Case (Oft)	Small tongue, high palate and short jaw. Found it difficult to suck. Problems swallowing. Tissue at the back of the throat made swallowing difficult. Could not eat lumpy food and had to wash it down.
255	Case (Oft)	Cleft palate, only recently chewing food. Problems swallowing.
256	Case (Oft)	Tube fed. Bottle untile 2 years. Took liquidised solids at two and a half.
257	Case	Vomiting. Didn't like chewing. Problems with solids. Not a good eater. Doesn't like spicy food like curry. (child is asian)
262	Case	Mother was continually feeding her. She stopped this but the child continued to gain weight.
265	Control	Tube fed. Problems sucking. Couldn't take a whole bottle at a time.
266	Control	Vomiting when on milk

57 cases - 8 cases who reported that there was no problem but who went on to describe a problem
28 controls

Comparison of WISC-III^{UK} subtests

Standardised Subtest scores	*cases* *n = 105*	*controls* *n = 117*	*t*	*p*
Picture Completeion	8.09	8.7	1.50	0.13
Information	8.16	8.33	0.40	0.69
Coding	8.96	9.25	0.65	0.53
Similarities	8.42	8.78	0.68	0.51
Picture Arrangement	8.50	8.84	0.64	0.53
Arithmetic	8.30	9.40	2.66	0.008
Block Design	8.10	8.60	0.89	0.62
Vocabulary	8.20	8.60	0.77	0.58
Object Assembly	8.78	8.97	0.44	0.65
Comprehension	6.70	7.56	1.70	0.08