Investigating Statistical Approaches to Handling Missing Data in the Context of the Gateshead Millennium Study

Claire Ann Gordon

A Dissertation Submitted to the
University of Glasgow
for the degree of
Master of Science in Statistics

School of Mathematics and Statistics

December 2010

Abstract

A commonly occurring problem in all kinds of studies is that of missing data. These missing values can occur for a number of reasons, including equipment malfunctions and, more typically, subjects recruited to a study not participating fully. In particular, in a longitudinal study, one or more of the repeated measurements on a subject might be missing.

The way in which missing values are dealt with depends on the data analyst's experience with statistical techniques. The most common way in which data analysts proceed is to use the complete case analysis method, i.e. removing cases with missing values for any of the variables and running the analysis on the remaining cases. Although this method is very straightforward to implement and is used by the vast majority of data analysts, it can lead to biased results unless data are missing completely at random. Complete Case analysis can dramatically reduce the sample size of the study, as only those cases for which all variables are measured are included in the analysis. Therefore the complete case analysis method is "not generally recommended" (Diggle et al., 2002). Alternative approaches to the complete case analysis method involve filling in (or imputing) values for the incomplete cases, making "more efficient use of the available data" (Schafer, 1997).

The purpose of this thesis is to compare and contrast the results obtained from analysing the relationship between growth and feeding behaviour in the first year

of life using the complete case analysis and three imputation methods: single hot-decking, multiple hot-decking and the EM algorithm. The data used in this research come from the Gateshead Millennium Study, a prospective study of a cohort of just over 1,000 babies. In practical terms, the purpose of the work is to confirm the conclusions from the published complete-case analysis. It is of more theoretical interest to determine which imputation method is the most appropriate for dealing with missing data in this study.

Chapter 1 provides an introduction to the problem of missing data and how they may arise and a description of the Gateshead Millennium Study data, to which all the missing data methods will be applied. It concludes by giving the aims of this thesis.

Chapter 2 provides an in depth review of various missing data approaches and indicates which characteristics of the missing data have to be considered in order to determine which of these approaches can be employed to deal with the missing values. Also in Chapter 2, various aspects of the Gateshead Millennium Study data are reviewed. Measures of growth and feeding behaviour in the first year of life are described as these are important variables in the published analysis.

Chapter 3 assesses how complete the Gateshead Millennium Study data is by producing a detailed description of each of the questions in each of the questionnaires. This is achieved by examining the Wave Non-response, Section Non-response and Item Non-response for each of the six questionnaires.

Chapter 4 recreates the results from the complete case analyses for the relationship between development of growth and feeding in the first year of life which have already been performed and published in the paper - **How Does Maternal and Child Feeding Behaviour Relate to Weight Gain and Failure to Thrive? Data From a Prospective Birth Cohort** (Wright et al., 2006a). This chapter also gives insight as to whether or not it is appropriate to assume

that the missing data mechanism is MCAR and therefore whether or not it is reasonable to believe the results obtained from the complete case analysis.

Chapter 5 focusses on the various methods used to impute the missing values in the Gateshead Millennium Study data. This chapter begins by considering the EM Algorithm. It gives details of how the EM Algorithm was performed and the results obtained. In addition to the EM Algorithm, this chapter also considers the procedures and results for Single Imputation and Multiple Imputation by hot-decking. This chapter concludes by comparing the results of these methods to one another and also to the complete case analysis results from Chapter 4.

Finally, Chapter 6 provides a summary of the results from the various missing data methods applied and discusses various alternative methods which could also have been performed.

Acknowledgements

I would like to begin by giving a special thanks to Professor John McColl for all his support, encouragement, invaluable expertise and supervision during the production of this thesis. I would also like to thank both Professor John McColl and Professor Marian Scott for their empathy during stressful times.

I would like to say thank you to my external advisors from the PEACH Unit at Yorkhill hospital, Professor Charlotte Wright for her guidance, valuable knowledge and for allowing me the use of the Gateshead Millennium Study data and Dr Andrea Sherriff for her guidance and assistance.

Thanks and gratitude must also go to the Department of Statistics for giving me the opportunity to engage in this research and to the Engineering and Physical Sciences Research Council for funding me throughout this research.

Finally, I would like to say thanks to my family, friends at the University and members of the Department of Statistics for their constant support and encouragement.

Declaration

I confirm that this thesis is my own work.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Introduction

In longitudinal studies, experimental units, e.g. people or animals, are repeatedly measured over time (Diggle et al., 2002) which enables the direct study of change. At specified time points throughout the study, each experimental unit has a number of measurements taken on several variables of interest. This means that longitudinal studies can distinguish between changes over time within experimental units and differences among the experimental units in the study.

Longitudinal studies are most commonly prospective studies which involve following the experimental units forward in time, although the studies can also be retrospective which involves obtaining repeated measurements on experimental units through historical records. An example of a prospective study is a randomized clinical trial to compare different drug therapies in the treatment of schizophrenia, with measurements being taken at specified times throughout the length of the study (Diggle et al., 2002).

Since the experimental units are repeatedly measured over time, a number of

observations will be recorded for each experimental unit. The experimental units can be assumed to be independent of one another, but the repeated measurements on each experimental unit are likely to be correlated with one another and this must be taken into account when making inferences based on the data.

Missing values occur in longitudinal studies when one or more of the repeated measurements on an experimental unit within the study are incomplete. For example, referring back to the clinical trial which compares different drug therapies in the treatment of schizophrenia, missing values may occur due to a patient's early departure from the study. Missing values may arise for a number of possible reasons including:

- subjects moving away from the area
- subjects dying
- subjects discontinuing treatment due to adverse side effects
- subjects missing an appointment/not returning questionnaires
- records being lost

It is important in any study to consider why data are missing and whether or not missingness is related to the practical questions being investigated using the data. It is also important to deal with missing data in such a way that, as far as possible, the missing data do not lead to the results of the data analysis being biased.

Once again, the schizophrenia example is used to draw attention to the fact that missing data can lead to results being biased. If the missing data were to be ignored completely in the analysis of the data obtained during the trial, then the data analysts may find that one of the drugs is more effective in treating

schizophrenia than the others. This may not be the case if the missing data were taken into account, e.g. patients who have dropped out of the study may have had an adverse reaction to the drug in question so the analysis ignoring the missing data might be biased in favour of this treatment.

Three terms have been coined for the different mechanisms by which missing data may arise, depending on whether or not missingness is associated with the underlying values in the dataset (Rubin, 1976). The missing data mechanisms are Missing Completely at Random (MCAR), Missing at Random (MAR) and Not Missing at Random (NMAR). MCAR means that missingness does not depend on the missing or observed data, MAR means that missingness depends on the observed data but not the missing data and NMAR means that missingness depends on the missing data. The appropriate way to analyse the data is different depending on which of these missing data mechanisms are in operation.

In this thesis, the impact of missing data in longitudinal studies will be explored through the Gateshead Millennium Study.

1.2 The Gateshead Millennium Study

The Gateshead Millennium Study is a prospective cohort study of feeding and growth in infancy. This study was set up primarily to explore the relationship between development of growth and feeding in the first year. Babies born between 1 June 1999 and 31 May 2000 in the Gateshead area of northeast England were recruited to the study shortly after birth.

Within the recruitment year of the Gateshead Millennium Study, approximately two weeks in every three were assigned to be recruitment weeks and babies born in these pre-specified 34 recruitment weeks were eligible for recruitment to the

study. As well as the child being born in Gateshead in one of the pre-specified recruitment weeks, another criterion for recruitment to the study was that the mother of the child was a Gateshead resident at the time of delivery.

Of all births and multiple births in the 34 recruitment weeks, a total of 1029 (83%) babies of 1011 mothers were recruited to the study (shortly after the birth).

Mothers who agreed to participate in the study had a face-to-face interview shortly after recruitment, during which baseline information, including birthweight and socio-demographic data, was recorded. Participating parents also completed a questionnaire at recruitment and received postal questionnaires at 6 weeks, 4 months, 8 months, 12 months and 30 months to complete and return (**Appendix A**). As well as filling out and returning these questionnaires, parents were asked to keep weaning and finger food diaries which were part of the parent-held Personal Child Health Record (PCHR) which parents received at recruitment to the study. The Personal Child Health Record also included forms, which were to be completed by health professionals, in order to keep a record of the child's weight throughout their development.

In each of the six questionnaires, a wide range of feeding questions were asked including:

At present, how is your baby's appetite?

Very Good — Good — All Right — Poor — Very Poor

Each of the individual questionnaires also asked about different aspects of the mother and child. On the front of each questionnaire, parents were also asked to transcribe all weights recorded in the Personal Child Health Record since completing and returning the previous questionnaire.

As this is a longitudinal study, it is prone to non-response so a number of tactics were decided upon when designing the study to improve response rates and ensure

the success of the study, including media involvement, support from local health professionals, telephone reminders for questionnaire completion, newsletters and birthday cards. Although this would have reduced the number of non-responses, there are still a number of mothers who have not responded throughout the length of the study. **Table 1.1**, below, gives the number of respondents and the response rates for each of the individual questionnaires.

Questionnaires	**Number of Respondents**	**Response Rate (%)**
Newborn	1027	99.8
6 Week	831	80.8
4 Month	762	74.1
8 Month	676	65.7
12 Month	633	61.5
30 Month	491	47.7

Table 1.1. ***Questionnaire Response Rates***
The questionnaire response rate is calculated by dividing the number of respondents to each questionnaire by the total number of subjects recruited to the study (1029), multiplied by 100.

Table 1.1 shows that as time passes the number of respondents decreases, therefore the number of non-respondents increases.

Questionnaires no. of rows	**Newborn**	**6 Week**	**4 Month**	**8 Month**	**12 Month**
554	0	0	0	0	0
79	0	0	0	0	1
31	**0**	**0**	**0**	**1**	**0**
70	0	0	0	1	1
11	0	0	1	0	0
9	0	0	1	0	1
14	0	0	1	1	0
63	0	0	1	1	1
6	0	1	0	0	0
5	0	1	0	0	1
3	0	1	0	1	0
14	0	1	0	1	1
2	0	1	1	0	0
10	0	1	1	0	1
12	0	1	1	1	0
144	0	1	1	1	1
2	1	1	1	1	1

Table 1.2. ***Missing Data Pattern of Wave Non-response***
The no. of rows represent the number of mothers with that particular pattern of missing data across the five questionnaires. A value of 0 in the table corresponds to a questionnaire that has been returned and a value of 1 in the table corresponds to a questionnaire that has not been returned.

For example, looking at **Table 1.2**, 31 mothers returned the newborn, 6 week, 4 month and 12 month questionnaires but, for some reason or other, the 8 month questionnaire was not received. This is known as **Wave Non-response** which is defined as the unintended and temporary loss of cohort members as time passes. There could be a number of possible explanations for this including:

- the mothers did not return received questionnaire, either because they forgot, were too busy or decided they did not want to complete one at this time
- the mothers did not receive the questionnaires, e.g. because they had moved

away from the Gateshead area but had not sent forwarding addresses immediately

- the completed questionnaires were not received by the people in charge of the study e.g. lost in post

Non-response can also be looked at through **Section Non-response** and **Item Non-response**. **Section Non-response**, in the context of the Gateshead Millennium Study, is when a subject who completed and returned a questionnaire missed out or refused to answer a section of the questionnaire. **Item Non-response** is similar to **Section Non-response** with the difference being that each question is looked at individually to see which questions, if any, have not been answered. There are likely to be different reasons for section or item non-response as opposed to wave non-response. The most common are that:

- the mothers were confused about the meaning of the question
- the mothers found the question invasive or embarrassing.

This shows that there are various non-response types that need to be looked at.

1.3 Aim

A preliminary aim of this thesis is to assess how complete each data group is by producing a detailed description of the completeness of each question in the Newborn, 6 week, 4 month, 8 month, 12 month and 30 month questionnaires of the Gateshead Millennium Study. This can be difficult to implement since, as well as those who do not complete and return the questionnaires, there are mothers

who do not answer some questions or whole sections of the questionnaires.

The major aim of this thesis is to explore different approaches to handling missing data and their impact on the results of the analysis of data from the Gateshead Millennium Study. The various key analyses that have already been published (e.g. an analysis of variance for linear trend and a multiple linear regression for the relationship between feeding and weight gain from birth to 12 months) have used the complete-case analysis method. This method should only be used in certain circumstances as it can lead to biased results depending on the missing data mechanisms in operation. Therefore, it will be interesting to see how the results from the complete-case analyses compare with the results obtained from more complex missing data approaches, such as the EM algorithm, simple imputation and multiple imputation, and also to see how the more complex approaches compare to one another.

Chapter 2

Literature and Methods

2.1 Gateshead Millennium Study

The Gateshead Millennium Study is a prospective cohort study that was initially developed to explore the relationship between development of growth and feeding in the first year of life.

Feeding in the first year of life was assessed using a single appetite question which was asked in each of the six questionnaires. In this thesis, only five of the six questionnaires will be used - Newborn (3 days after birth), 6 Week, 4 Month, 8 Month and 12 Month questionnaires. Development of growth was assessed using the Thrive Index score (**Section 2.1.2**). Other factors which are used to explore the relationship between development of growth and feeding in the first year of life are Avoidant Eating Behaviour, Maternal Feeding Anxiety and Response to Food Refusal. These factors along with appetite and Thrive Index will be explained in the following sections.

2.1.1 Appetite

In each of the six questionnaires that mothers had to complete as part of the Gateshead Millennium study, a wide range of feeding questions were asked including:

At present, how is your baby's appetite?

Very Good — Good — All Right — Poor — Very Poor

This question was used to assess feeding in the first year of life as it is thought that early appetite determines feeding and also weight later in life. The data obtained for this question from each of the five questionnaires used is as follows:

Questionnaire Appetite	**Newborn**	**6 Week**	**4 Month**	**8 Month**	**12 Month**
Very Good	213	537	439	365	280
Good	353	193	219	188	226
All Right	262	17	26	49	58
Poor	22	2	5	4	10
Very Poor	38	-	-	4	4
Item Non-response	43 (4.2%)	82 (8.0%)	73 (7.1%)	66 (6.4%)	55 (5.3%)
Wave Non-response	2 (0.2%)	198 (19.2%)	267 (25.9%)	353 (34.3%)	396 (38.5%)

Table 2.1. ***Mothers Response to Appetite Question (Original)***

Although this question has been selected as being useable at every age to assess feeding in the first year of life, the appetite rates given by mothers in the Newborn questionnaire may not give an adequate representation of the child's/childrens' appetite as all mothers may not have had sufficient time to establish their child's/childrens'

appetite and some mothers may have nothing to base or compare their initial rating to i.e. this may be their first child and their first time feeding a baby.
Although the baby's appetite was originally rated on a 5-point scale, for the purpose of the analysis it has been converted to a 3-point scale as shown in **Table 2.2** (Wright et al., 2006a).

New Coding	**Original Coding**
Normal	Very Good
Borderline	Good
Low	All Right, Poor, Very Poor

Table 2.2. ***Coding of Appetite Question***

The reason the original 5-point scale has been converted to a 3-point scale is because the appetite rates reported by mothers who answered this question in each of the questionnaires were very skewed with only a small proportion of subjects falling into the 'All Right', 'Poor' and 'Very Poor' categories compared to the number of subjects in the 'Good' and 'Very Good' categories as shown in **Table 2.1**. This conversion also removes any question as to whether or not the appetite rate was reported accurately as the 'Poor' and 'Very Poor' categories were not in descending order in all of the questionnaires (reversed) i.e. parents could have possibly completed it thinking it was on a continuous scale hence marking 'Very Poor' instead of 'Poor' and vice versa.
Converting our data from a 5-point to a 3-point scale gives the following table:

Questionnaire Appetite	**Newborn**	**6 Week**	**4 Month**	**8 Month**	**12 Month**
Normal	213	537	439	365	280
Borderline	353	193	219	188	226
Low	322	19	31	57	72

Table 2.3. ***Mothers Response to Appetite Question (Converted)***

Table 2.3 shows that the majority of parents who answered this question in each of the questionnaires rate their child's/childrens' appetite as being 'Normal' except in the case of the Newborn questionnaire, which gives us reason to believe that the appetite rates recorded in the Newborn questionnaire are not an adequate representation of the child's/childrens' appetites and therefore should not be used to assess feeding in the first year of life.

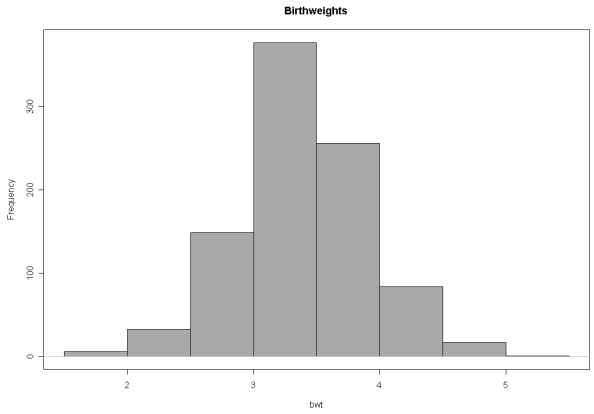
2.1.2 Thrive Index

During the first year of life, children in the UK are routinely weighed by primary care nurses in community based baby clinics. These routinely collected weights are recorded in parent-held Personal Child Health Records (PCHR) which mothers receive just after the birth of their child/children.

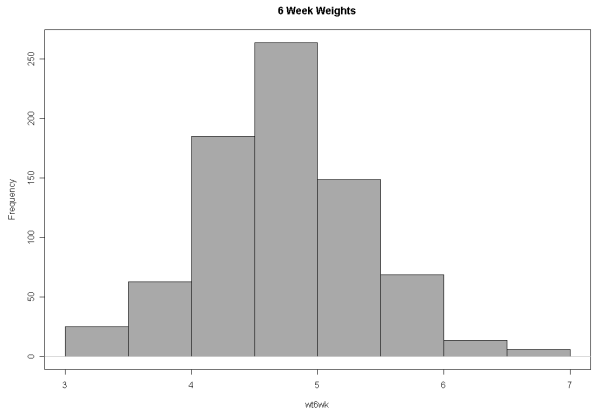
In the Gateshead Millennium Study, parents were asked to transcribe all weights recorded in the PCHR, since completing and returning the previous questionnaire, onto the front of each questionnaire as well as the date the measurement was taken.

At the age of 13 months, the children were weighed by the health professionals and a copy of the weight recording page from the PCHR was retrieved from parents by the health professional in order to check that the weights written on the front of the questionnaires by parents were identical to those in the clinics' records.

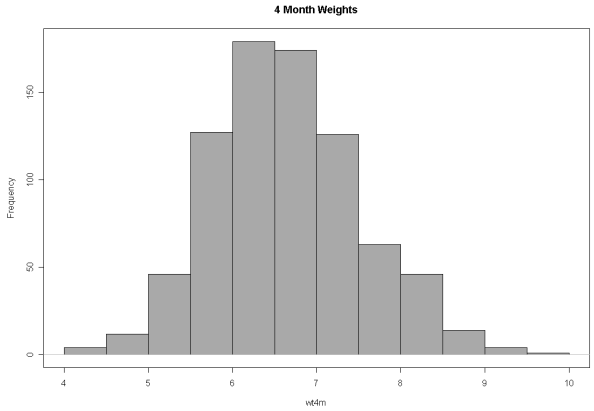
Once the routinely collected weights were cleaned and crosschecked, they were converted to Standard Deviation Scores (SDS) compared to the British 1990 growth reference (Freeman et al., 1995) using a Box-Cox transformation. The SD scores represent the difference between the actual weight and the population mean weight in units of the standard deviation. Converting raw weights to standard deviation scores is intended to result in the transformed data at any given age having an approximate standard Normal distribution with mean 0 and variance 1 in the reference population.



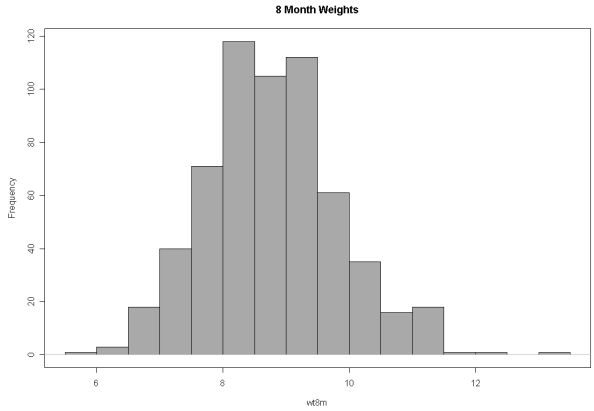
(a) Birthweights



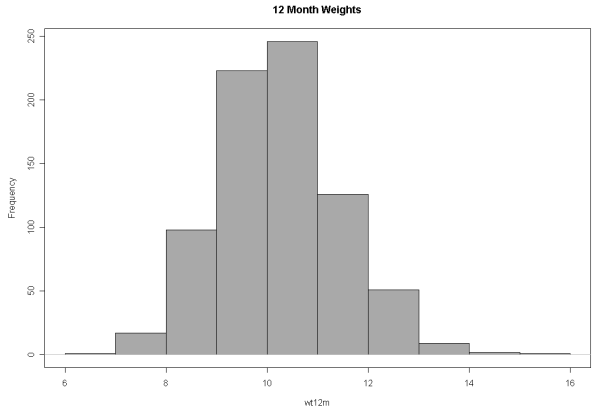
(b) 6 Week Weights



(c) 4 Month Weights

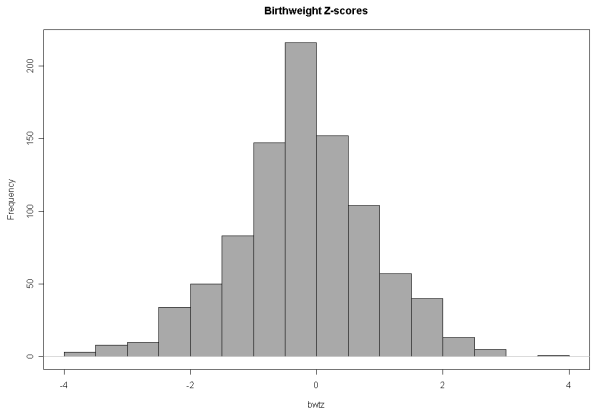


(d) 8 Month Weights

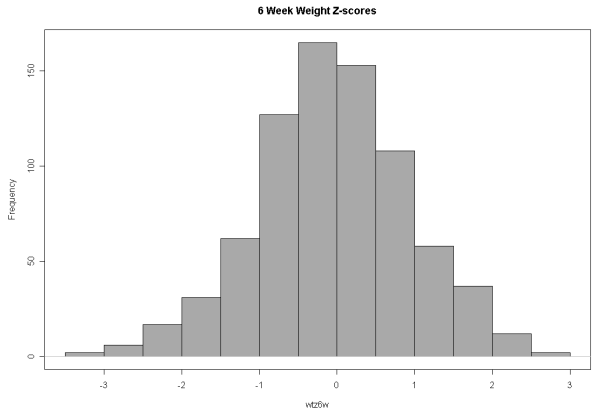


(e) 12 Month Weights

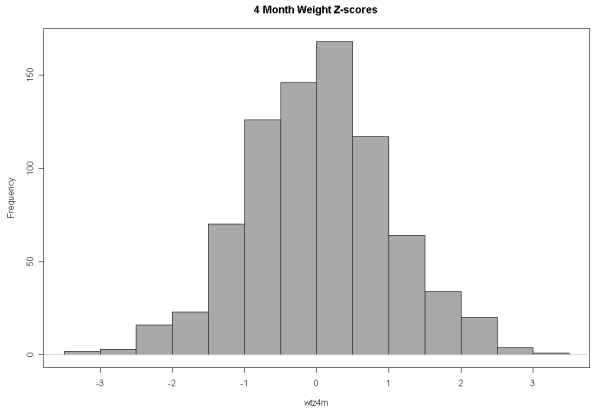
Figure 2.1. ***Histograms of Raw Weights***



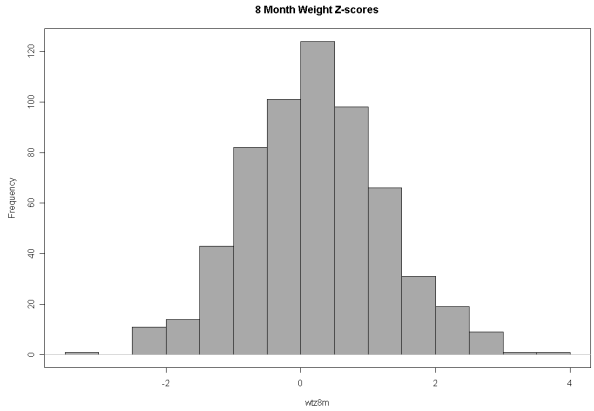
(a) Birthweight Z-scores



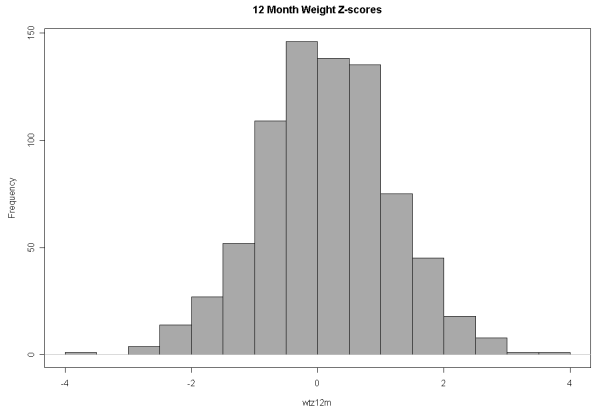
(b) 6 Week Weight Z-scores



(c) 4 Month Weight Z-Scores



(d) 8 Month Weight Z-Scores



(e) 12 Month Weight Z-Scores

Figure 2.2. ***Histograms of Weight Z-Scores***

Figure 2.1 shows the histograms of the raw weights recorded at each of the time points and **Figure 2.2** shows the histograms of the weight z-scores at each of the different time points.

Weight SD scores are used instead of the average of the weights because interest is in looking at a measure for the growth of each child in their first year of life and not the average weight of each child in their first year of life i.e. not an adequate measure as children are weighed at different times so taking the average weight would not give a fair representation.

Once the weights were converted to standard deviation scores, the Thrive Index scores were then calculated. The Thrive Index (TI) is defined by Wright et al. (2006a) as "a measure of the change in weight standard deviation score over time, conditional on initial weight, which adjusts for regression to the mean". This compares the child's actual weight SD score to their expected weight SD score. The TI score for birth to 12 months (TI0-12m) gives the growth of a child in their first year of life and is calculated by Wright et al. (2006b) using the following formula:

$$\mathbf{TI0-12m = wtz12m - 0.38 \times bwtz} \tag{2.1}$$

where wtz12m is the weight z-score at 12 months and bwtz is the birthweight z-score. The value of 0.38 is the regression coefficient from the complete-case analysis when wtz12m is regressed on bwtz. **Figure 2.3** illustrates how the formula used by Wright et al. (2006b) to calculate the TI score was found.

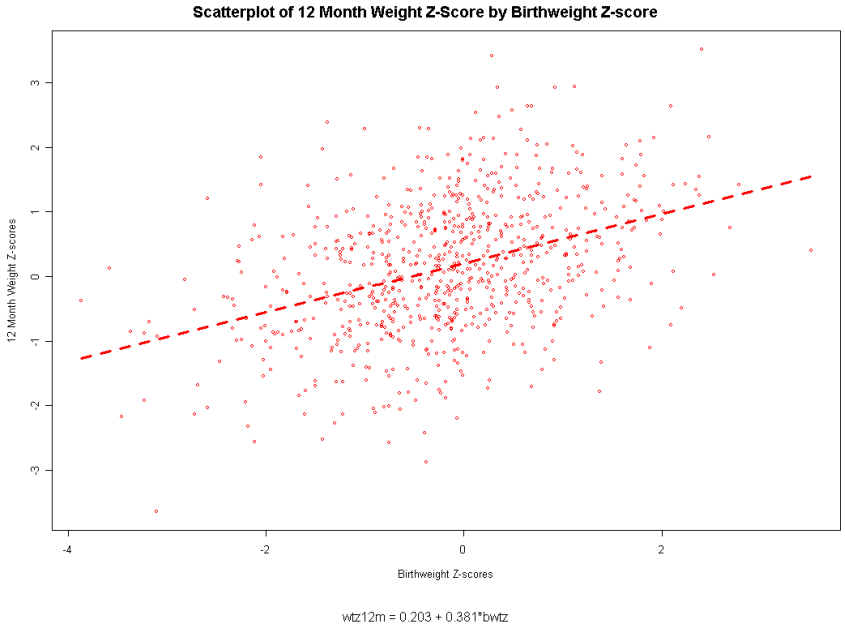


Figure 2.3. ***Scatterplot of 12 Month Weight Z-scores regressed on Birthweight Z-scores***

2.1.3 Avoidant Eating Behaviour

Avoidant Eating Behaviour (AEB) deals with the range of ways in which a child could resist being fed. In order to examine the extent to which children might resist, Wright et al. (2006a) identified eight questions, drawn from research and clinical experience, to devise scores for AEB. The questions posed to parents in order to establish Avoidant Eating Behaviour scores are as follows:

How often does your baby do the following when given food?

(a) Pushes food away — ***Rarely — Sometimes — Often***

(b) Turns head — ***Rarely — Sometimes — Often***

(c) Closes mouth — ***Rarely — Sometimes — Often***

(d) Gags — ***Rarely — Sometimes — Often***

(e) Holds food in mouth ***Rarely — Sometimes — Often***

(f) Spits ***Rarely — Sometimes — Often***

(g) Throws food ***Rarely — Sometimes — Often***

(h) Cries ***Rarely — Sometimes — Often***

An overall rating of avoidant eating behaviour was constructed by summing together the parents' responses to these questions. Each response was allocated a score in order to calculate avoidant eating behaviour. These scores are as follows:

- If response from parent is **Rarely**, a score of 0 is given
- If response from parent is **Sometimes**, a score of 1 is given and
- If response from parent is **Often**, a score of 2 is given

Once the overall rating of avoidant eating behaviour has been calculated, the (overall) scores are separated into low, medium and high categories as follows:

Avoidant Eating Behaviour	**Sum of Scores**
Low	0 - 1
Medium	2 - 5
High	> 5

Table 2.4. ***Coding of Avoidant Eating Behaviour Scores***

The data obtained from parents who responded to the questions relating to Avoidant Eating Behaviour in the 12 month questionnaire is as follows:

Avoidant Eating Behaviour Score	**12 Month Questionnaire**
Low	142
Medium	261
High	175

Table 2.5. ***Mothers Response to Avoidant Eating Behaviour Questions***

Table 2.5 shows that the highest frequency of children in the Gateshead Millennium Study have a medium Avoidant Eating Behaviour score after the parents have responded to the questions relating to Avoidant Eating Behaviour in the 12 month questionnaire.

2.1.4 Response To Food Refusal

Response to Food Refusal (RTFR) questions are a group of five questions, put to parents in the 8 month and 12 month questionnaires of the study, which examine how mothers responded when their child/children refused to eat a meal. This group of five questions devised to examine Response to Food Refusal was developed by Wright et al. (2006a) from previous research and from their own clinical studies.

The questions put to parents in order to generate a score for Response to Food Refusal are as follows:

If your baby does not finish a course, or part of a meal, what do you do?

(a) Encourage him/her to eat ***Rarely — Sometimes — Often***

(b) Make him/her eat ***Rarely — Sometimes — Often***

(c) Offer something else ***Rarely — Sometimes — Often***

If your baby does not finish a course, or part of a meal, what do you do after the meal?

(a) Offer same food again later ***Rarely — Sometimes — Often***

(b) Offer something else later ***Rarely — Sometimes — Often***

From the parents responses to these questions, an overall rating of Response to Food Refusal was constructed by summing together the five responses to the above questions. Each response was given a score in order to calculate Response to Food Refusal. These scores are as follows:

- For questions, encourage him/her to eat, offer something else, offer same food again later and offer something else later:
  - If response from parent is **Rarely**, a score of **0** is given
  - If response from parent is **Sometimes**, a score of **1** is given and
  - If response from parent is **Often**, a score of **2** is given
- For question, make him/her eat:
  - If response from parent is **Rarely**, a score of **0** is given
  - If response from parent is **Sometimes**, a score of **2*** is given and
  - If response from parent is **Often**, a score of **4*** is given

*** These are allocated higher scores to represent extreme responses (Wright et al., 2006a)**

Once the overall rating of Response to Food Refusal has been calculated, the overall scores are separated into low, medium and high categories as follows:

Response To Food Refusal	**Sum of Scores**
Low	0 - 3
Medium	4 - 5
High	> 5

Table 2.6. ***Coding of Response to Food Refusal Scores***

The data obtained from parents who responded to the questions relating to Response To Food Refusal in the 8 month and 12 month questionnaires is as follows:

RTFR Score	**8 Month Questionnaire**	**12 Month Questionnaire**
Low	302	269
Medium	240	241
High	63	66

Table 2.7. ***Mothers Response to Response To Food Refusal Questions***

From parents responses to the questions relating to Response to Food Refusal in the 8 month and 12 month questionnaires, **Table 2.7** shows that the highest frequency of parents in the study have a low Response to Food Refusal score at both 8 months and 12 months, suggesting they are not too worried about their child's/childrens' eating.

2.1.5 Maternal Feeding Anxiety

Maternal Feeding Anxiety (MFA) deals with how mothers cope with their child's/childrens' feeding times. Two questions, posed to parents in the 8 month and 12 month

questionnaires, were used to generate scores to examine mothers stress levels when feeding their child/children.

The questions put to parents in order to establish MFA scores are as follows:

Overall, is your baby feeding enough? ***Yes – Not Always – No***

At present, are feeding times for you usually:

Very Relaxed — Relaxed — OK — Stressful — Very Stressful

An overall rating of maternal feeding anxiety was constructed by summing together the parents' responses to these questions. Each response was allocated a score in order to calculate Maternal Feeding Anxiety. The scores for each response are as follows:

- For ***'Overall, is your baby feeding enough'*** question:
  - If response from parent is **Yes**, a score of **0** is given
  - If response from parent is **Not Always**, a score of **1** is given and
  - If response from parent is **No**, a score of **2** is given.
- For ***'At present, are feeding times for you usually'*** question:
  - If response from parent is **Very Relaxed**, a score of **0** is given
  - If response from parent is **Relaxed**, a score of **1** is given
  - If response from parent is **OK**, a score of **2** is given
  - If response from parent is **Stressful**, a score of **3** is given and
  - If response from parent is **Very Stressful**, a score of **4** is given

Once the overall rating of Maternal Feeding Anxiety has been calculated, the overall scores are separated into normal, borderline and high categories as follows:

Maternal Feeding Anxiety	**Sum of Scores**
Normal	0
Borderline	1
High	> 1

Table 2.8. ***Coding of Maternal Feeding Anxiety Scores***

The data obtained from parents who responded to the questions relating to Maternal Feeding Anxiety in the 12 month questionnaire is as follows:

Maternal Feeding Anxiety Score	**12 Month Questionnaire**
Normal	401
Borderline	123
High	54

Table 2.9. ***Mothers Response to Maternal Feeding Anxiety Questions***

Table 2.9 shows that the majority of mothers in the Gateshead Millennium Study have a normal Maternal Feeding Anxiety score after the parents have responded to the questions relating to Maternal Feeding Anxiety in the 12 month questionnaire. This suggests that most mothers are coping well with their child's/childrens' feeding times.

2.2 Approaches to Analysing Missing Data

2.2.1 Missing Data

2.2.1.1 Introduction

When faced with the problem of missing values, many researchers tend to use ad hoc methods to create a complete dataset from the incomplete dataset.

The simplest way of doing this is to include only experimental units that have no missing values for any of the variables used in the analysis. This is known as Complete Case analysis. This method is commonly used in many statistical software packages so that standard statistical methods for complete data can still be performed on incomplete datasets.

The complete case analysis method may be a satisfactory approach for dealing with missing values if the percentage of missing values in a large dataset is small and the bias is kept to a minimum. However, this is not usually the case and large amounts of data are discarded.

Another method which is frequently used when faced with the problem of missing values in a dataset is imputation. Imputation involves filling in plausible values for the missing ones in order to obtain an apparently completely observed dataset.

There are various imputation methods which can be performed to achieve the desired outcome of a completely observed dataset and some of these approaches will be discussed in **Section 2.2.3**.

As well as discussing the complete case analysis method and the various imputation methods, the Expectation-Maximization (EM) Algorithm will also be considered.

Before considering the different approaches to handling missing values, general

patterns of missing data and missing data mechanisms are reviewed because these characteristics of the missing data will influence which methods can be used to deal with the missing values.

Little and Rubin (2002) and Schafer (1997) are highly regarded for their work and achievements in the field of missing data and their books Statistical Analysis with Missing Data and Analysis of Incomplete Multivariate Data, respectively, are highly recommended by statisticians.

2.2.1.2 General Patterns of Missing Data

The Missing Data Pattern shows which values in the data matrix are observed and which are missing. Little and Rubin (2002) and Schafer (1997) both agree that it is very useful to be able to identify the pattern of missing data as the statistical method used to analyse the data depends upon the type of missing data pattern acquired. There are numerous patterns of missing data (Little and Rubin, 2002) but in this thesis we shall concentrate on two - monotone and general non-monotone missing data patterns in longitudinal studies.

Suppose measurements are taken on a number of subjects at specified times throughout the length of a study. A monotone missing data pattern occurs if a measurement for a particular subject is missing for a certain time point and for all successive time points. An example of this type of missing data pattern is shown in **Figure 2.4**.

This type of missing data pattern usually occurs if the subject drops out of the study. Thus, no additional measurements will be recorded after that time.

In mathematical terms, using Little and Rubin's (2002) notation, let $Y = [y_{ij}]$ where $i = 1, ..., n$ and $j = 1, ..., k$ denote an n x k completely observed dataset

Experimental Units	T1	T2	T3	T4	T5
1	0	0	0	0	0
2	0	0	0	0	0
3	0	0	0	0	1
4	0	0	0	0	1
5	0	0	0	1	1
6	0	0	0	1	1
7	0	0	1	1	1
8	0	0	1	1	1
9	0	1	1	1	1
10	0	1	1	1	1

Figure 2.4. ***Monotone Missing Data Pattern***
(0 = Observed, 1 = Missing)

where y_{ij} is the value of variable Y_j for subject i. For datasets including missing data, a missing data indicator matrix, $M = m_{ij}$ is defined such that, $m_{ij} = 1$ if y_{ij} is missing and $m_{ij} = 0$ if y_{ij} is present. This matrix, M, defines the pattern of missing data. Schafer (1997) states that, whenever a value y_{ij} is missing, y_{ik} must also be missing $\forall k > j$ for a data matrix to have a monotone missing data pattern. The ordering of experimental units in a monotone missing data pattern is very important in order to see if a pattern occurs in the data. Schafer (1997) and Little and Rubin (2002) both agree that monotone patterns of missing data most commonly arise in longitudinal studies as subjects drop out of the study before the end and do not return.

A general non-monotone missing data pattern may occur if a number of subjects miss a scheduled appointment at one or more of the specified times throughout the length of the study. In the general missing data pattern, missing data can occur anywhere in the data matrix as shown in **Figure 2.5**.

According to Little and Rubin (2002), this "haphazard" pattern of missing data most commonly occurs in surveys through Item Non-response and Diggle et al. (2002) believes that it is more difficult to deal with non-monotone missing data

Experimental Units	T1	T2	T3	T4	T5
1	0	0	1	0	0
2	1	0	1	0	1
3	0	1	1	1	0
4	0	0	0	0	1
5	0	1	1	1	0
6	0	1	0	1	1
7	1	1	1	1	1
8	0	0	0	1	1
9	0	1	0	1	1
10	0	1	0	1	0

Figure 2.5. ***Non-Monotone Missing Data Pattern***
(0 = Observed, 1 = Missing)

than monotone missing data "because of the wider variety of patterns of missing values which need to be accommodated". This type of missing data pattern is typically handled using imputation which will be discussed in **Section 2.2.3**.
As well as being able to determine which missing data pattern is in use, it is also useful to consider which type of missing data mechanism might be in operation. The Gateshead Millennium Study data suffers from a general non-monotone missing data pattern, as some mothers are not completing and returning the questionnaires at one or more of the pre-specified times, and so missing data can occur anywhere in the dataset.

2.2.1.3 Missing Data Mechanisms

As mentioned previously, there are several reasons why data may be missing and the Missing Data Mechanism shows the mechanism by which the missing data may have arisen. The Missing Data Mechanism in operation is dependent upon whether or not missingness is associated with the underlying values in the dataset. There are three different missing data mechanisms which may be encountered

depending on whether or not the fact that a particular value is missing is linked to the underlying values. These are Missing Completely at Random (MCAR), Missing at Random (MAR) and Not Missing at Random (NMAR). It is extremely valuable to consider which of the three missing data mechanisms might be in use as the appropriate statistical approach to analysing the data depends on the missing data mechanism in operation. If the process by which the missing data has arisen is ignored, the statistical technique used for the analysis of the data may often lead to biased and inefficient estimates.

Before the concept of missing data mechanisms was introduced by Rubin in 1976, the mechanism by which missing data may arise, depending on whether or not missingness is associated with the underlying values in the dataset, was very much ignored. Since then, Rubin's classification of Missing Data Mechanisms has been regarded as being "fundamental to the modelling of incomplete data" (Molenberghs and Kenward, 2007) and is in common use in the field of missing data with slightly different notation to that used in the original 1976 paper.

Following from **Section 2.2.1.2**, let $Y = (Y_{obs}, Y_{mis})$ is the complete data matrix where Y_{obs} represents the observed elements of Y and Y_{mis} denotes the missing elements of Y and M is the "missing data indicator matrix".

In terms of the Little and Rubin (2002) notation, the Missing Completely at Random (MCAR) assumption, "characterised by the conditional distribution of M given Y", assumes that:

$$f(M|Y, \phi) = f(M|\phi) \quad \forall \quad Y, \phi \tag{2.2}$$

where ϕ represents the unknown parameters of the model.

This means that the probability of a value being missing is unrelated to either the observed or unobserved elements of the data. For example, a patient leaves a longitudinal study because they move house (Little, 1995). This type of data

would be said to be missing completely at random since the reason the subjects values are missing does not depend on their previous results or on the results that would have been obtained if they had not left the study.

The assumption of MCAR can be checked by dividing recruits to the study into those included and not included in the analysis and then performing t-tests of mean differences on key variables in the dataset. If a non-significant result is obtained from the t-test i.e. no systematic difference between those included in the analysis and those not included in the analysis, then there is no evidence of a difference against the MCAR assumption of the missing data being a random sample of all of the data. It is possible that the MCAR assumption may hold and that no biased results will be obtained from the complete case analysis, but this can never be proven and depends on having informative factors available for the non-respondents. Little (1995) states that if any differences are found between those included in the analysis and those not included in the analysis i.e. the MCAR is not valid, then these differences will have important implications for the analysis and an alternative statistical technique involving imputing missing values will have to be chosen and used.

The second missing data mechanism, Missing at Random (MAR), is less restrictive than the MCAR assumption. If the dataset consists of a large number of variables, it is regarded as being the most plausible missing data mechanism. The Missing at Random assumption can be stated as follows:

$$f(M|Y,\phi) = f(M|Y_{obs},\phi) \quad \forall \quad Y_{mis},\phi \tag{2.3}$$

This means that the probability of a value being missing may be related to the observed elements of the data but not to the unobserved elements of the data. For example, a patient leaves a longitudinal study on their doctor's advice, based on their previously observed measurements (Little, 1995). This type of data would

be said to be missing at random since the reason the subject's values are missing depends on their earlier observed results and not on the results that would have been obtained if they had not been advised to leave the study.

There is currently no test available to check the MAR assumption, although Schafer (1997) suggests that, even if the missing data are not strictly missing at random, procedures using this type of missing data mechanism appear to produce better results than ad hoc procedures such as Complete Case Analysis as these procedures "remove all of the nonresponse bias explainable by Y_{obs}, whereas ad hoc procedures may not." MCAR is an 'ignorable missingness' process meaning the process that caused the missing data can be ignored. MAR can also be said to be an 'ignorable missingness' process if the analysis performed takes into account the dependence between the observed variables. Therefore, the process by which the missing data arises does not have to be accounted for when using the chosen estimation method.

The third missing data mechanism is Not Missing at Random (NMAR) which is a non-ignorable missingness process meaning that the actual mechanism which caused the missing data has to be examined and modelled appropriately. The term Not Missing at Random means that the probability of a value being missing depends on the observed and unobserved elements of the data. For example, a patient misses their appointment because they are feeling unwell (Little, 1995).

All the methods for handling missing data that have been implemented in this thesis assume that the missing data mechanism is ignorable, therefore the process that caused the missing data can be ignored.

Having now considered the characteristics of the missing data that will influence which methods can be used to deal with the missing values, Complete Case analysis, Imputation methods and the EM Algorithm will now be reviewed.

2.2.2 Complete Case Analysis

The most commonly used technique for dealing with missing data among non-statisticians is the method of complete case analysis. This is a very simple method for dealing with datasets that contain missing values, but the complete case analysis method is deemed as an "inadequate solution to the problem" by Diggle et al. (2002) and others alike and is "not generally recommended" as usually a large percentage of useful information is being discarded.

The complete case analysis method omits all cases with missing values from the analysis and only includes those cases for which all measurements are observed. For this reason, this method is only viable when the fraction of observations with missing values is small and the overall number of observations is large. The data analyst proceeds as if the cases removed from the analysis had never really been observed and so no provision for the missing data is made in the analysis (Schafer, 1997).

This method usually results in a considerable decrease in the number of cases which are available for analysis as it can only use subjects who have values for all of the variables involved in the analysis, but it has the important advantage of producing unbiased estimates for the parameters if the assumption that the data are MCAR holds i.e. the cases removed from the analysis are similar to those included in the analysis. Other obvious advantages of this method are that it is very easy to describe and also that standard complete data statistical analyses can be applied without any adjustments needing to be made as the data structure is as planned.

The disadvantages of this method arise from the conceivable loss of information in removing the incomplete cases from the analysis. If the MCAR assumption does not hold, this method can result in biased parameter estimates as it is ignoring potential systematic differences between the complete and incomplete cases. This

method also results in a significant loss in power and precision due to the reduced sample size.

This method can only be justified if the missing data mechanism in operation is MCAR. In addition to the missing data mechanism being MCAR, Schafer (1997) suggests that the complete case analysis method may be a satisfactory solution to the problem of missing data if the cases excluded from the analysis comprise of only a small percentage of all cases, 5%, say. However, Little and Rubin (2002) state that it is hard to create a general rule which can be used to validate the use of the complete case analysis method as the degree of bias and loss of precision depends not only on the fraction of complete cases and pattern of missing data, but also on the extent to which complete and incomplete cases differ.

In **Chapter 4**, the complete case analysis performed for the relationship between weight and appetite in the first year of life (complete case analysis results adapted from Table 4 of Wright et al. (2006a)) will be discussed as well as whether or not this method of analysis seems to be reasonable for the Gateshead Millennium Study data.

Schafer (1997) advises using imputation methods to substitute appropriate values for the incomplete cases rather than omitting the incomplete cases completely as these methods make "more efficient use of the available data". Harrell (2001) also agrees that "making up data for incomplete cases is better than throwing away real data".

Although alternative approaches to handling missing data should be considered in light of the problems arising with the Complete Case analysis method, not all of these alternative methods are better, as shall be seen in **Section 2.2.3**.

2.2.3 Imputation

A widely used technique for dealing with missing data is that of imputation. According to Little and Rubin (2002), "imputation is a general and flexible method for handling missing data problems" but it has a number of potential difficulties as Dempster and Rubin (1983) explain: "The idea of imputation is both seductive and dangerous. It is seductive because it can lull the user into a pleasurable state of believing that the data are complete after all, and it is dangerous because it lumps together situations where the problem is sufficiently minor that it can be legitimately handled in this way and situations where standard estimators applied to the real and imputed data have substantial biases" (Little and Rubin, 2002, page 59).

Imputation involves filling in (or imputing) values for the incomplete cases, usually using the observed values that are available. Unlike the Complete Case analysis method which removes any rows from Y that are not completely observed, leaving only Y_{obs}, imputation procedures produce complete datasets that have the same size as Y and so make more effective use of all of the observed data.

There are numerous imputation methods which can be used to handle missing data and these approaches can be applied in one of two ways - Single Imputation and Multiple Imputation. Both Single Imputation and Multiple Imputation methods will be considered in **Sections 2.2.3.1 and 2.2.3.2**, respectively.

2.2.3.1 Single Imputation

Single Imputation Methods

In this section, various Single Imputation (SI) methods will be considered as well as potential reasons why they should or should not be performed when imputing values for the missing values in the dataset. There are many single imputation approaches which can be used for imputing missing values. However, some of these procedures are better than others as shall be seen in this section. All Single Imputation methods theoretically rely on the assumption of the data being Missing at Random (MAR). This is a less restrictive assumption than the assumption of MCAR required for the complete case analysis and can be met using the observed data, in some way or another, to fill in values for the missing data.

The method of Single Imputation involves replacing each missing value in the dataset with one imputed value, creating a 'complete' dataset to which standard statistical techniques can be applied. The way in which the missing values are imputed depends upon which Single Imputation method has been chosen.

The methods of Single Imputation which are reviewed here are Last Observation Carried Forward (LOCF), mean imputation, regression imputation and hot deck imputation, although there are many other forms of Single Imputation as mentioned in Little and Rubin (2002). As well as describing these methods, reasons for and against their use will be given.

The Last Observation Carried Forward procedure involves filling in the missing values for a subject with their last recorded value for that particular measurement. For example, if the LOCF method was used for data in the form of **Figure 2.4**, Experimental Unit 3 would have its value at 'T4' used to fill in a value for 'T5' and Experimental Unit 5 would have its value at 'T3' used to fill in a value

for 'T4' and 'T5', etc. This is a simple way to deal with missing data in a longitudinal study although in many settings it is unrealistic as the majority of subjects' measurements will change through time (depending upon what is being assessed). Mean Imputation is another simple way of imputing values for missing values in a dataset. It involves estimating the missing values of a variable by the mean of the observed values for that variable. Thus, no additional information is being added as the overall mean will be identical whether the missing values have been imputed by the mean of the observed values or not. This leads to the standard errors being underestimated as the overall mean remains unchanged by the substitution of the missing values but the sample size has apparently increased. This method of imputation also distorts the distribution of the data, as it is imputing values at the centre of the distribution, and this reduces the apparent standard deviation which again affects the usual standard errors. Due to the reasons given above and the fact that this method "does not take into account, when producing the imputed value for a particular subject, any of the other information gathered on that subject" (Molenberghs and Kenward, 2007), this method is deemed problematic and therefore should not be used for imputation purposes.

Another Single Imputation method which is often used is Regression Imputation. As the name suggests, Regression Imputation uses regression to predict values for the missing entries of a variable based on other variables that have been measured for the subjects in the study. This method is better than mean imputation as it takes into account other information which has been collected on a subject when imputing a value for that subject. However, it does not solve the problem associated with mean imputation of underestimated standard errors as any values which have to be imputed will lie along the regression line. Again, this method is not really adding any additional information but it has apparently increased the sample size. If this was to be used as the method of choice, random variation

would have to be added to each imputed value to allow for fluctuations in the data from the regression line in order to solve the problem of underestimated standard errors.

Hot Deck Imputation, also known as hot-decking, is a well known technique for use in missing data problems. It involves replacing missing values by values obtained from "similar" subjects in the sample. This method of imputation is very common in survey settings and can involve complex schemes for selecting subjects that are "similar" for imputation purposes (Little and Rubin, 2002). The advantages of this method of imputation are that the imputed values do not distort the distribution of the data and it is good at preserving the variance structure.

From the imputation methods that have been reviewed in this section, it was decided that the best method to use for the Gateshead Millennium Study data was hot deck imputation as it is good at preserving the variance structure (Little, 1995). More details of how Hot Deck Imputation was achieved for the Gateshead Millennium Study data is given in **Section 5.2**. In the various imputation methods that have been reviewed, the values that have been observed in the dataset are used in some way or another to impute values for the missing observations. Once the missing data has been imputed using one of the imputation procedures, the now 'complete' dataset is analysed using one of the standard complete data methods of analysis, ignoring the fact that the missing data have been imputed i.e. treating them as real. Schafer (1997) warns that it is a "serious mistake to treat the imputed data as if they were real" and continue with the research without making adjustments/provisions for the fact that the missing data have been imputed because this will lead to invalid results as any standard errors or p-values obtained will fail to reflect the additional uncertainty required to account for the missing data being imputed.

Since Single Imputation does not account for imputation uncertainty, the standard errors and p-values of tests obtained are smaller than would be expected if imputation uncertainty was taken into account and subsequently any confidence intervals calculated will be narrower than expected. For this reason, Rao and Shao (1992) have formulated a special adjustment, the adjusted jackknife variance estimator, that will reflect the sampling variability in order to obtain precise standard deviations.

Adjusted Jackknife Variance Estimator

Using standard statistical techniques to analyse a 'completed' dataset, obtained from performing a particular imputation procedure e.g. hot deck imputation, does not allow for the true uncertainty due to non-response and therefore a further adjustment has to be made to account for this.

The special adjustment, which has been used in this thesis is the Adjusted Jackknife Variance Estimator, which gives the increase in variance due to the missing values being imputed. The formulae used to obtain the increase in variance will be viewed as though the missing data had been imputed using the hot decking procedure.

"Suppose, in a simple random sample of size n, r units respond and m do not respond to an item y. Consider the simplest form of hot deck imputation in which a simple random sample of size m is selected with replacement from the respondents to item y and the associated y-values are used as donors, that is, the imputed value $y_i{}^* = y_j$ for some $j \epsilon A_r$, where A_r denotes the sample of respondents. The imputed estimator of the population mean $\bar{Y}$ is $\bar{y}_I = \frac{1}{n}(r\bar{y}_r + m\bar{y}_m^*)$ where $\bar{y}_r$ is the mean of the respondents' values and $\bar{y}_m^*$ is the mean of the imputed values." (Rao and Shao, 1992)

The adjusted jackknife estimator of the variance of $\bar{y}_I$, which includes the increase

in variance due to non-response, is given by:

$$v_{JK} = \frac{n-1}{n}\sum_{j=1}^{n}[\bar{y}_I^a(-j) - \bar{y}_I]^2 \tag{2.4}$$

where

$$\bar{y}_I^a(-j) = (n-1)^{-1}[n\bar{y}_I - y_j - \frac{m(y_j - \bar{y}_r)}{r-1}] \quad when \quad j\epsilon A_r \tag{2.5}$$

and

$$\bar{y}_I^a(-j) = (n-1)^{-1}[n\bar{y}_I - {y_j}^*] \quad when \quad j\epsilon A_m \tag{2.6}$$

The adjusted jackknife variance estimator of $\bar{y}_I$ was calculated in **R** (R Development Core Team, 2010) using code based on Equations (2.4), (2.5) and (2.6). In **Section 5.2**, the Single Hot Deck Imputation (SHDI) method will be applied to the Gateshead Millennium Study data and the results from the analysis of variance for linear trends and multiple linear regressions using this method will be compared to the complete case analysis results.

2.2.3.2 Multiple Imputation

Since the concept of Multiple Imputation (MI) was introduced by Rubin (1978) around 30 years ago, it has become, according to Molenberghs and Kenward (2007), "an important and influential approach for dealing with the statistical analysis of incomplete data".

Multiple Imputation is an extension of the Single Imputation method as it involves replacing each missing value by two or more imputed values, creating

multiple 'completed' datasets to which standard statistical techniques can be applied, therefore resolving the main problem of estimating the true uncertainty due to non-response associated with Single Imputation. The way in which the missing values have been imputed depends upon which imputation method has been chosen.

The Multiple Imputation procedure assumes that the probability of a value being missing may be related to the observed elements of the data but not to the unobserved elements of the data i.e. that the missing data are Missing at Random. Since the Multiple Imputation method relies on the assumption of the data being MAR, the observed data can be used, in some way or another, to fill in values for the missing data.

In order to obtain parameter estimates which reflect the uncertainty that arises from imputing missing data using the Multiple Imputation method, the following three steps are required. The first step is to generate a number of 'completed' datasets, say D, by imputing values for each missing value D times. The second step is to analyse the D 'completed' datasets using the standard statistical technique that would have been used if the data had been complete. The third and final step is to combine the results of the D analyses found in step two to obtain a single parameter estimate which properly reflects the uncertainty due to non-response.

To generate the D 'completed' datasets required for step one, single imputation methods such as hot deck imputation could be used and repeated a number of times in order to create the multiple datasets.

The second step involves analysing the D 'completed' datasets using a standard statistical technique which produces D sets of results. The formulae required to combine the results of the D multiple datasets to obtain a single parameter estimate for step three are given below (Little and Rubin, 2002).

"Let $\hat{\theta}_d$ and W_d, $d = 1, ..., D$ be the complete-data estimates and their associated variances for an estimated parameter θ, calculated from D repeated imputations under one model."

The combined estimate from the D multiple datasets is:

$$\bar{\theta}_D = \frac{1}{D} \sum_{d=1}^{D} \hat{\theta}_d \tag{2.7}$$

which is a simple average of the D complete-data estimates.

The total variability associated with $\bar{\theta}_D$ is

$$T_D = \bar{W}_D + \frac{D+1}{D} B_D \tag{2.8}$$

where

$$\bar{W}_D = \frac{1}{D} \sum_{d=1}^{D} W_d \tag{2.9}$$

is the within-imputation variance and

$$B_D = \frac{1}{D-1} \sum_{d=1}^{D} (\hat{\theta}_d - \bar{\theta}_D)^2 \tag{2.10}$$

is the between-imputation component.

As well as deciding which imputation method to use to create the D 'completed' datasets, the number of multiple datasets, D, has to be specified. The number of multiple datasets required to obtain precise estimates of the parameters of interest depends on the fraction of information missing due to non-response, $\hat{\gamma}_D$,

where (Little and Rubin, 2002):

$$\hat{\gamma}_D = (1 + \frac{1}{D})\frac{B_D}{T_D} \tag{2.11}$$

Table 2.10 from Rubin (1987), page 114 gives the efficiencies achieved for different numbers of imputations and rates of missing information. The efficiency of a finite D imputation estimator relative to the fully efficient infinite D imputation estimator is approximately

$$(1 + \frac{\hat{\gamma}_D}{D})^{-1} \tag{2.12}$$

Values of this efficiency are listed in Table 2.10 for some possible values of D and γ. **Table 2.10** shows that there is little advantage in producing and analysing more than three to ten imputations, unless γ is exceptionally high.

		γ				
		0.1	**0.3**	**0.5**	**0.7**	**0.9**
D	**3**	97	91	86	81	77
	5	98	94	91	88	85
	10	99	97	95	93	92
	20	100	99	98	97	96

Table 2.10. ***Efficiency of Multiple Imputation (%)***

The method of Multiple Imputation reduces the increase in variance to negligible levels. Multiple Imputation also provides valid standard errors that take into account imputation uncertainty without having to use a further adjustment as

in Single Imputation. This method is also found to produce unbiased parameter estimates when the size of the sample is small and also when the rate of missing data is high.

In Section 5.3, the multiple hot deck imputation method will be applied to the Gateshead Millennium Study data and the results from the ANOVA for linear trend using this method will be compared to the complete case analysis and single hot deck imputation results.

2.2.4 EM Algorithm

The Expectation-Maximisation (EM) Algorithm is an iterative algorithm which is used to calculate maximum likelihood estimates in parametric models for incomplete data. It is a very "popular and remarkably simple method for maximum likelihood estimation in incomplete-data problems" (Meng and Rubin, 1991).

Dempster et al. (1977) provide a helpful introduction to the EM Algorithm as well as Schafer (1997), Little and Rubin (2002) and McLachlan and Krishnan (1997), who give comprehensive descriptions and applications of the algorithm.

The EM Algorithm approach, as with the Single and Multiple Imputation procedures, assumes that the missing data are Missing at Random. So, the observed data can be used in some way, or another, to fill in values for the missing data.

The basic idea behind the EM Algorithm is to replace each missing value by estimated values and estimate the parameters. The missing values are then re-estimated using the new, assumed correct, parameter estimates and the parameters are then re-estimated. This process continues until convergence has been reached.

Each iteration of the EM Algorithm consists of two steps, the Expectation step

(E-step) and the Maximisation step (M-step). The E-step calculates the conditional expectation of the complete data log-likelihood given the observed data and the parameter estimates, $\mathbf{E}[l(\theta|Y)|Y_{obs}, \theta^{(t)}]$, and the M-step finds the parameter estimates that maximise the complete data log-likelihood from the E-step. The E-step and M-step are repeated alternatively until convergence. Convergence is found when the difference between two iterations is arbitrarily small.

The EM Algorithm can be shown to converge reliably and it is also conceptually and computationally simple. The disadvantages of the EM Algorithm are that the rate of convergence can be very slow when there is a large amount of missing data and it does not always converge to the optimum. Another disadvantage of the EM Algorithm is that it does not provide an estimate of the observed variance-covariance matrix of the parameter estimates which is required to obtain confidence intervals for the parameter estimates. In order to obtain a numerically stable estimate of the variance-covariance matrix of the parameter estimates, the Supplemented EM (SEM) Algorithm can be used.

Supplemented EM Algorithm

The Supplemented EM Algorithm (Meng and Rubin, 1991) has been used in this thesis to obtain a "numerically stable estimate of the asymptotic variance-covariance matrix of the EM computed estimates" which reflects the true uncertainty due to non-response. The basic concept of the SEM algorithm is to "use the fact that the rate of convergence of EM is governed by the fraction of missing information to find the increased variability due to missing information to add to the complete-data variance-covariance matrix".

The Supplemented EM Algorithm can be used in this instance to find the observed variance-covariance matrix as the complete-data variance-covariance matrix is known.

Using Little and Rubin's (2002) notation (pages 191-192):

$$V_{obs} = V_{com} + \Delta V \tag{2.13}$$

where V_{obs} is the observed data variance-covariance matrix, V_{com} is the complete data variance-covariance matrix and $\Delta V = V_{com} DM(I - DM)^{-1}$ is the increase in variance due to missing data.

$$DM = i_{mis} i_{com}^{-1} = I - i_{obs} i_{com}^{-1} \tag{2.14}$$

where DM is the derivative of the EM mapping, $i_{com} = -D^{20} Q(\theta|\theta)|_{\theta=\theta^*}$ is the complete information, $i_{obs} = I(\theta|Y_{obs})|_{\theta=\theta^*}$ and $i_{mis} = -D^{20} H(\theta|\theta)|_{\theta=\theta^*}$ is the missing information at the converged value of θ.
$DM = i_{mis} i_{com}^{-1} = I - i_{obs} i_{com}^{-1}$ implies that $i_{obs}^{-1} = i_{com}^{-1}(I - DM)^{-1}$, that is

$$V_{obs} = V_{com}(I - DM) - 1 \tag{2.15}$$

where $V_{obs} = i_{obs}^{-1}$ and $V_{com} = i_{com}^{-1}$ are the variance-covariance matrices for the observed data and the complete data, respectively and I is the $d \times d$ identity matrix.
Meng and Rubin (1991) show how to evaluate DM using code for the E- and M-steps of the EM Algorithm in Section 3.3 of their paper **Using EM to Obtain Asymptotic Variance-Covariance Matrices: The SEM Algorithm**.
In **Section 5.1**, the EM Algorithm approach will be applied to the Gateshead Millennium Study data and the results from the analysis of variance for linear trends and multiple linear regressions using this method will be compared to the complete case analysis results.

Implementing the EM Algorithm and Supplemented EM Algorithm

R (R Development Core Team, 2010) is a free and widely used statistical language for statistical computing which has been used in this thesis to implement the methods of imputation described above.

The EM Algorithm approach for missing 12 month weights is implemented using functions from the **norm** library (Ported to R by Alvaro A. Novo. Original by Joseph L. Schafer, jls@stat.psu.edu, 2002) in **R**. This procedure begins by creating a data matrix containing the 12 month weight z-scores and the other variables which are to be used to estimate and impute the missing 12 month weight z-scores e.g. 8 month weight z-scores, 8 month and 4 month weight z-scores, etc. Once it has been decided which variables are going to be used to estimate and impute the missing 12 month weight z-scores, the **prelim.norm** function is used to sort the rows of the data matrix by the missing data pattern and to scale/centre the columns of the data matrix. It also calculates various quantities of the data matrix needed for input to the **em.norm** function. Once the **prelim.norm** function has been used, the output from this function is used as input to the **em.norm** function. The **em.norm** function uses Multivariate Normal models to obtain the maximum likelihood estimates of the parameters. The output from the **prelim.norm** function and the **em.norm** function as well as the data matrix are then used as input to the **imp.norm** function. This function creates a 'completed' data matrix with the missing elements of the original data matrix being imputed with simulated draws from a Multivariate Normal distribution given the observed data.

Since the procedure described above does not reflect the true uncertainty due to non-response, the Supplemented EM Algorithm approach is used. This procedure begins by calculating $\Delta V = V_{com} DM(I - DM)^{-1}$ which is the increase in variance due to missing data. DM which is the derivative of the EM mapping

(Equation 2.14) is evaluated in **R** using code created from the information given on page 192 of Little and Rubin (2002) and the **em.norm** function.

R was then used to calculate ΔV, once the complete data variance-covariance matrix and the identity matrix were entered into **R**. ΔV was then added to V_{com} to find V_{obs}, the observed data variance-covariance matrix (Equation 2.13).

The missing values in the data matrix are then imputed with draws from the estimated mean and a standard error associated with it, obtained from V_{obs}.

The EM Algorithm approach for missing appetite rates is implemented in a similar way to the method described above for the missing 12 month weight z-scores, except the **cat** library (Ported to R by Ted Harding and Fernando Tusell. Original by Joseph L. Schafer, 2004) in **R** is used.

Chapter 3

Completeness of Gateshead Millennium Study Data

A preliminary aim of this thesis is to evaluate how complete each data group is by creating a comprehensive description of the completeness of each question in each of the questionnaires of the Gateshead Millennium Study. This, at times, can be difficult to execute since, as well as those who do not complete and return the questionnaires, there are those who do not answer some questions or whole sections of the questionnaires. This is further complicated by the fact that there are some questions which only need to be answered by those who answered a specific response to the preceding question.

This detailed description of the completeness of the Gateshead Millennium Study data was achieved by examining various types of non-response - Wave Non-response, Section Non-response and Item Non-response - for the Newborn, 6 Week, 4 Month, 8 Month and 12 Month questionnaires.

3.1 Wave Non-response

Wave Non-response is the unintended and temporary loss of cohort members as time passes. This means that missing data can occur anywhere in the dataset as subjects may not complete and return one or more of the questionnaires throughout the length of the study, leading to a general non-monotone missing data pattern. **Table 1.2** shows the general non-monotone missing data pattern for the Gateshead Millennium Study Data.

Table 1.1 gives the number of respondents and the response rates for each of the questionnaires involved in the Gateshead Millennium Study. The number of respondents were those mothers who had completed and returned the individual questionnaire. The number of respondents for each questionnaire had to be checked thoroughly as there were some mothers who returned blank questionnaires, therefore should not be counted as a respondent. The response rate, which is a measure of Wave Non-response, is the number of respondents divided by the number of subjects who were recruited to the study i.e. 1,029 babies. Looking at the response rates, from **Table 1.1**, for each of the six questionnaires, the number of respondents decreases as time passes, therefore the number of missing values increases. This is only to be expected with a longitudinal study.

It is interesting to note that two mothers who agreed to participate in the study have dropped out before the Newborn questionnaire was sent out and they did not answer any of the subsequent questionnaires. These two families could be missing because they moved away from the area shortly after being recruited to the study and did not leave a forwarding address.

3.2 Section Non-response

In the context of the Gateshead Millennium Study, Section Non-response is where a subject who has completed and returned a questionnaire, has missed out or refused to answer a whole section of the questionnaire. This could be due to the mother not understanding the meaning of the questions in that section or it could be due to the mother finding the questions in that specific section too personal. As previously mentioned, each questionnaire asks a wide range of feeding questions and each individual questionnaire asks about different aspects of the mother and child. For this reason, each questionnaire is split into sections depending on the nature of the questions posed. It will be of interest to compare how complete each section of each questionnaire is and also to compare how complete sections which are repeated throughout the length of the study are.

Tables B.1 - B.6 of **Appendix B** give the response rates for the Newborn, 6 Week, 4 Month, 8 Month, 12 Month and 30 Month questionnaires, respectively. The response rates, which are a measure of Section Non-response, are found by creating an indicator variable, for each section within each questionnaire, which gives the total number of mothers who answered that particular section of the questionnaire i.e. ≥ 1 **Qu. answered by respondents to questionnaire** column of tables. A mother is regarded as having answered the section if they have answered one or more of the questions included in that section as in some of the sections, mothers were asked a question in which, if they responded "Yes", they were asked to answer the remaining questions in the section, whereas if they responded "No", they were asked to proceed to the next section of the questionnaire. The total number of mothers who answered a particular section of the questionnaire is then divided by the corresponding number of mothers who completed and returned the individual questionnaire and is also divided by the

number of subjects who were recruited to the study to give two measures of Section Non-response.

Looking closely at **Tables B.1 - B.6** of **Appendix B**, each section of each questionnaire is greater than or equal to 94% complete when Section Non-response is considered using the **% of Respondents** response rates, except Section B of the 12 Month questionnaire (**Figure A.6 of Appendix A**) which is only 25% complete. The reason that this section is only 25% complete is because this section of questions in the questionnaire was only to be answered by those mothers whose child/children had "started solids" since completing and returning the 8 month questionnaire. When considering Section Non-response using the **% of Recruits** response rates, each section of each questionnaire is approximately 65% complete. These are smaller than the response rates calculated using the **% of Respondents** response rates which is only to be expected as these response rates are calculated using all of the recruits to the study i.e. it includes those mothers who were recruited to the study but did not complete and return the questionnaire being considered.

Looking at the Section Non-response rates for the **General Feeding Questions** section which is repeated in every questionnaire of the study, **Table 3.1**, it is completed by 99% or above of respondents to the questionnaire being considered. When considering Section Non-response using the **% of Recruits** response rates, the response rates decrease as time passes. Therefore, the number of non-respondents increases, which again, is only to be expected with a longitudinal study.

Questionnaire	Section Number	Number who answered Section	% of Respondents to qu're	% of Recruits to Study
Newborn	B	1016	98.9	98.7
6 Week	C	831	100.0	80.8
4 Month	C	762	100.0	74.1
8 Month	C	676	100.0	65.7
12 Month	C	632	99.8	61.4

Table 3.1. ***Section Non-response for General Feeding Questions Section of the Gateshead Millennium Study***

3.3 Item Non-response

In the context of the Gateshead Millennium Study, Item Non-response is where a subject who has completed and returned a questionnaire, has missed out or refused to answer a particular question of the questionnaire. This, again, could be due to the fact the mother had found the question invasive or embarrassing or because they were confused about the meaning of the question.

It will be of interest to compare how complete each question of each questionnaire is and also to compare how complete each of the questions which are repeated throughout the length of the study are.

Tables C.1 - C.5 of **Appendix C** give the response rates for the Newborn, 6 Week, 4 Month, 8 Month and 12 Month questionnaires, respectively. The response rates, which are a measure of Item Non-response, are again found by dividing the number of mothers who answered the question by the corresponding number of mothers who completed and returned the individual questionnaire and also by the number of mothers who were recruited to the study to give two measures of Item Non-response.

Looking in detail at **Tables C.1 - C.5** of **Appendix C**, the completeness of each question in each of the questionnaires varies extensively when Item Non-response is considered using both the **% of Respondents** and **% of Recruits** response

rates. Looking at the Item Non-response rates for the **Appetite** question which is repeated in every questionnaire of the study, **Table 3.2**, it is completed by 96% or above of respondents to the questionnaire being considered. When considering Item Non-response using the **% of Recruits** response rates, the response rates decrease, from 95.6% in the Newborn questionnaire to 61.4% in the 12 month questionnaire, as time passes. Therefore, the number of non-respondents increases, which again, is only to be expected with a longitudinal study.

Questionnaire	**Question Number**	**Number who answered Appetite Question**	**% of Respondents to qu're**	**% of Recruits to Study**
Newborn	4	984	95.8	95.6
6 Week	18	826	99.4	80.3
4 Month	20	756	99.2	73.5
8 Month	24	669	99.0	65.0
12 Month	21	632	99.8	61.4

Table 3.2. ***Item Non-response for Appetite Question of the Gateshead Millennium Study***

Another important measure of Item Non-response, in this study, which has to be included is the conditional response rate which corresponds to those questions in the questionnaire which only have to be answered by those who have answered a specific response to the preceding question. The number of mothers who responded to these questions have to be checked thoroughly as some mothers did not answer the initial question but continued on to answer the following questions, therefore these mothers should be counted as respondents.

For example, **Table 3.3** shows the conditional response rates for questions 10 to 16 of the 6 Week questionnaire which only had to be answered by those mothers who answered "Yes, solids given" to question 9 of the questionnaire (See **Figure A.2 of Appendix A** for further details of questions 9 to 16). Of the 801 mothers who answered question 9, only 21 responded "Yes, solids given", so it is only these mothers who should answer questions 10 to 16. The conditional

response rate is calculated by the number of mothers who answered the question divided by the number of mothers who answered "Yes, solids given" to question 9 i.e. 21 mothers. If Item Non-response was to be measured using the **% of Respondents** and **% of Recruits** response rates for these questions, it would suggest that these questions were poorly answered i.e. very high non-response rate, when in fact they are 65% or above completed by those mothers who had to answer the questions.

6 week Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/831)**	**% of Recruits (/1029)**	**Conditional Response Rate (/21)**
9	801	96.4	77.8	
10	21	2.5	2.0	100.0
11a	21	2.5	2.0	100.0
11b	8	1.0	0.8	100.0
12a	13	1.6	1.3	61.9
12b	16	1.9	1.6	76.2
12c	18	2.2	1.7	85.7
13	20	2.4	1.9	95.2
14	21	2.5	2.0	100.0
15	19	2.3	1.8	90.5
16a	18	2.2	1.7	85.7
16b	17	2.0	1.7	81.0
16c	18	2.2	1.7	85.7
16d	18	2.2	1.7	85.7
16e	21	2.5	2.0	100.0

Table 3.3. ***Item Non-response Rates for Questions 10 - 16 of 6 Week Questionnaire***

The creation of these tables of Wave Non-response, Section Non-response and Item Non-response will be very useful as there is now documentation of how complete each questionnaire, each section in each questionnaire and each question in each questionnaire is for future reference by the project team.

Chapter 4

Complete Case Analysis

The Gateshead Millennium Study was initially set up by Wright et al. (2006a) to explore the relationship between development of growth and feeding in the first year of life. This was achieved by looking at the relationship between Thrive Index (**Section 2.1.2**) and Appetite rated at 6 weeks and 12 months (**Section 2.1.1**) as well as other factors which were known or thought to affect Thrive Index.

The complete case analyses were performed for and published in the **How Does Maternal and Child Feeding Behaviour Relate to Weight Gain and Failure to Thrive? Data From a Prospective Birth Cohort** paper by Wright et al. (2006a) in order to determine which variables were significantly related to Thrive Index in the first year of life. Each possible explanatory variable - Appetite rated at 6 weeks, Appetite rated at 12 months, Avoidant Eating Behaviour rated at 12 months, Maternal Feeding Anxiety rated at 12 months, Response to Food Refusal rated at 8 months and Response to Food Refusal rated at 12 months - is included in an analysis of variance for linear trend (Altman, 1991) in order to determine if that specific explanatory variable, on its own, is

significantly related to Thrive Index. All six potential explanatory variables are then included in a multiple linear regression to determine which variables, if any, are significantly related to Thrive Index when other explanatory variables are already included in the model.

Of the 1,029 babies originally recruited to the study, only 923 babies were eligible to be included in the published analysis. Of the 106 not included in the published analysis, 68 were born pre-term, 33 were Ultra-Orthodox Jews and 5 were Muslim infants. These 106 babies showed major differences in weight gain patterns and were removed from the analysis for this reason, irrespective of the completeness of their data. For the subsequent chapters in this thesis, only the 923 infants included in the published analysis will be dealt with.

In **Chapter 3**, the completeness of the Gateshead Millennium Study data was found by exploring Wave Non-response, Section Non-response and Item Non-response for each of the six questionnaires. The complete case analysis method is only valid under the MAR assumption if the proportion of missing data is small and the sample size is large. For this reason, the proportion of missing values in each of the variables used in the complete case analysis has to be investigated.

Variable	**Missing**	**Proportion**
TI0-12m	**149**	**0.16**
6 Week Appetite	**174**	**0.19**
12 Month Appetite	**345**	**0.37**
12 Month AEB	**345**	**0.37**
12 Month MFA	**345**	**0.37**
8 Month RTFR	**318**	**0.34**
12 Month RTFR	**347**	**0.38**

Table 4.1. % Missing for Variable used in the Complete Case Analyses

Table 4.1 shows those variables calculated from questions in the 8 month or 12 month questionnaires have a much higher proportion of missing values than those calculated from questions in earlier questionnaires suggesting that the results from the complete case analyses involving these variables will not be valid.

This is also suggested by the fact that, in order for a subject to be included in the complete case analysis, they have to have a value for both the response variable and the explanatory variable. **Table 4.2** gives the proportion of cases removed for each of the independent analysis of variance for linear trends. For all the analysis of variance for linear trends, the proportion of cases removed is 0.30 or above, so the complete case analysis appears to be an inappropriate method to use for the Gateshead Millennium Study data. This result would again be confirmed if the proportion of missing values from the multiple linear regression output was to be examined.

Model	**Cases Excluded**	**Proportion**
TI0-12m $\sim$ 6 Week Appetite	**245**	**0.27**
TI0-12m $\sim$ 12 Month Appetite	**354**	**0.38**
TI0-12m $\sim$ 12 Month AEB	**354**	**0.38**
TI0-12m $\sim$ 12 Month MFA	**354**	**0.38**
TI0-12m $\sim$ 8 Month RTFR	**341**	**0.37**
TI0-12m $\sim$ 12 Month RTFR	**356**	**0.39**

Table 4.2. % Cases Excluded in the Complete Case Analyses

Although it has been suggested that the results from the complete case analyses would not be valid due to the high proportions of missing values in the data, the analysis of the data using the complete case analysis method has been performed and any conclusions reached will be treated with caution. For an analysis of variance for linear trends, the Null hypothesis is that the mean Thrive Index

from birth to 12 months is the same for all levels of the explanatory variable i.e. no linear trend and the Alternative hypothesis is that the mean Thrive Index from birth to 12 months is decreasing linearly through the levels of the explanatory variable i.e. linear trend. The results from the complete case analyses adapted from Table 4 in Wright et al. (2006a) after some clarifications and corrections are as follows:

	TI, Mean (SD)	**n**	**p**^{a}	**p**^{b}
Appetite rated at 6 weeks				
Normal	0.28 (0.94)	484		
Borderline	0.06 (0.93)	176		
Low	- 0.25 (0.88)	18	0.001	0.010
Appetite rated at 12 months				
Normal	0.33 (0.92)	277		
Borderline	0.15 (0.96)	222		
Low	- 0.10 (0.96)	70	< 0.001	0.005
AEB rated at 12 months				
Low	0.33 (0.90)	140		
Medium	0.23 (0.96)	259		
High	0.08 (0.97)	170	0.017	
MFA rated at 12 months				
Normal	0.28 (0.94)	396		
Borderline	0.09 (0.91)	120		
Low	- 0.09 (1.03)	53	0.002	
RTFR rated at 8 months				
Low	0.28 (0.90)	290		
Medium	0.16 (1.01)	231		
High	0.04 (1.01)	61	0.048	
RTFR rated at 12 months				
Low	0.31 (0.91)	264		
Medium	0.16 (0.94)	237		
High	- 0.05 (1.10)	66	0.004	0.025

Table 4.3. Relationship Between Feeding and Eating Behaviour and Weight Gain from Birth to 12 Months

* values are mean (SD) Thrive Index from birth to 12 months
p^a gives the resulting p-values for the ANOVA for linear trends and p^b gives the p-values of the explanatory variables included in the Multiple Linear Regression including all other significant variables

From the six separate analysis of variance for linear trends, it was found that all of the explanatory variables were significantly related to weight gain from birth to 12 months (**Table 4.3**). However, when the multiple linear regression was performed (**Table 4.3**), only Appetite rated at 6 weeks, Appetite rated at 12 months and Response to Food Refusal rated at 12 months were significantly related to Thrive Index from birth to 12 months when added to the model together. The models obtained from the analyses of **Table 4.3** were proposed by Wright et al. (2006a) using their chosen method of model selection. The results of the analysis of variance for linear trends and multiple linear regressions obtained from using different approaches to handling missing data in this dataset, **Chapter 5**, will be compared to the results from the Complete Case Analysis (**Table 4.3**).

It will be of interest to see how the complete case analysis approach fares in comparison to more complicated missing data methods that are now available, such as Single Imputation and Multiple Imputation.

It has already been mentioned that for the complete case analysis method to be valid, the data is assumed to be MCAR i.e. the missing data are a random sample of all data so we would expect to see no systematic difference between those children included in the Complete Case analysis and those omitted from the Complete Case analysis. There is limited scope for comparing these two groups, but one interesting variable in the dataset that is recorded for virtually all the children is birthweight z-score. Looking at **Tables 4.4 and 4.5**, there appears to be a slight difference between those children included in the Complete Case analysis and those not included in the Complete Case analysis, with children not included in the analysis having a slightly lower birthweight z-score than those included in the analysis. This difference is found to be statistically significant when a two-sample t-test is performed ($p = 0.033$, $p = 0.001$), therefore the MCAR assumption is not valid and so the children included in the complete case analysis

are not representative of the entire cohort. This analysis highlights the fact that if we assume that the Gateshead Millennium Study data are MCAR, then the resulting means will be biased in favour of those children with higher birthweights so the results from the Complete Case analysis might not be representative of the population as a whole.

	N	**Mean**	**St. Dev.**
Included in CC Analysis	678	-0.16	1.08
Not Included in CC Analysis	244	-0.33	1.10

Table 4.4. Mean and Standard Deviations for Birthweight z-scores for children included and not included in the TI0-12M ∼ 6 Week Appetite Rates analysis

	N	**Mean**	**St. Dev.**
Included in CC Analysis	569	-0.11	1.06
Not Included in CC Analysis	353	-0.36	1.11

Table 4.5. Mean and Standard Deviations for Birthweight z-scores for children included and not included in the TI0-12M ∼ 12 Month Appetite Rates analysis

The above two-sample t-tests include all children included and not included in the Complete Case analysis and since it is likely that boys will weigh more than girls, if more boys than girls are included in the group of children that are included in the Complete Case analysis then this would artificially increase the birthweights in that group, leading to the apparent difference between the groups. **Table 4.6** shows that more boys than girls were recruited to the study, and therefore

more boys than girls were included in the group of children that are included in the Complete Case analysis. Wright et al. (2006a) mentions that those children not included in the Complete Case analysis tended to come from more deprived neighbourhoods than those children included in the Complete Case analysis so the Complete Case analysis might be biased in favour of those children from more affluent neighbourhoods. Therefore, examination of other variables to compare those included and not included in the Complete Case analysis, such as gender and deprivation, would be required in order to establish if the data were in fact MCAR.

Model	**Boys Included**	**Boys Excluded**	**Girls Included**	**Girls Excluded**
TI0-12m$\sim$ 6 Week Appetite	**341**	**124**	**337**	**121**
TI0-12m$\sim$12 Month Appetite	**286**	**179**	**283**	**175**

Table 4.6. Number of Boys and Girls included and not included in the Complete Case Analyses

Chapter 5

Missing Data Methods

The main purpose of this thesis is to explore different approaches to handling missing data and their impact on the results of the various key analyses which have already been performed.

The Gateshead Millennium Study was originally set up to explore the relationship between development of growth and feeding in the first year of life and the results of the complete case analyses which have been performed to assess the relationship between Thrive Index (TI0-12m) and appetite rated at 6 weeks and 12 months, as well as other factors which were known or thought to affect Thrive Index, have been published in the **How Does Maternal and Child Feeding Behaviour Relate to Weight Gain and Failure to Thrive? Data From a Prospective Birth Cohort** paper by Wright et al. (2006a).

The research team are now interested in looking at how the results from the complete case analyses change, if at all, when more complex missing data methods are implemented to impute the missing values which are contained in the Gateshead Millennium Study dataset. In particular, interest lies in the analysis of variance for linear trends examining the relationship between TI0-12m and

Appetite rated at 6 weeks and the relationship between TI0-12m and Appetite rated at 12 months as it has been suggested that infant weight gain in Britain is associated more with feeding and intrinsic characteristics of the infant than maternal factors. Therefore, the missing data methods have been applied to the Gateshead Millennium Study data to impute values for the missing TI0-12m scores, 6 week appetite rates and 12 month appetite rates. Thrive Index score for the growth of a child in their first year of life (TI0-12m) is calculated using birth and 12 month weight z-scores so instead of imputing the missing TI0-12m scores directly, the various missing data methods are used to impute the missing 12 month weight z-scores and these imputed values along with the observed values for birth and 12 month weight z-scores are used to calculate the TI0-12m scores. The missing data approaches of Single Hot Deck Imputation, Multiple Hot Deck Imputation and the EM Algorithm have been chosen to impute the missing values for 12 month weight z-scores, 6 week appetite rates and 12 month appetite rates as these methods can be used to impute missing values for both continuous and ordinal variables. As well as looking at the analysis of variance for linear trends examining the relationship between TI0-12m and Appetite rated at 6 weeks and the relationship between TI0-12m and Appetite rated at 12 months, it is also worth considering how the results of the multiple linear regressions change after imputation of the TI0-12m scores, 6 week appetite rates and 12 month appetite rates.

There are a number of possible ways in which the missing 12 month weight z-scores, appetites rated at 6 weeks and appetites rated at 12 months can be imputed using the agreed missing data approaches.

In order to calculate the missing TI0-12m scores, the missing 12 month weight z-scores can be imputed in a variety of ways using the weight z-scores at birth, 6 weeks, 4 months and 8 months i.e. 12 month weight z-scores can be imputed

using just birthweight z-scores or could be imputed using the birthweight and 6 week weight z-scores together, etc.

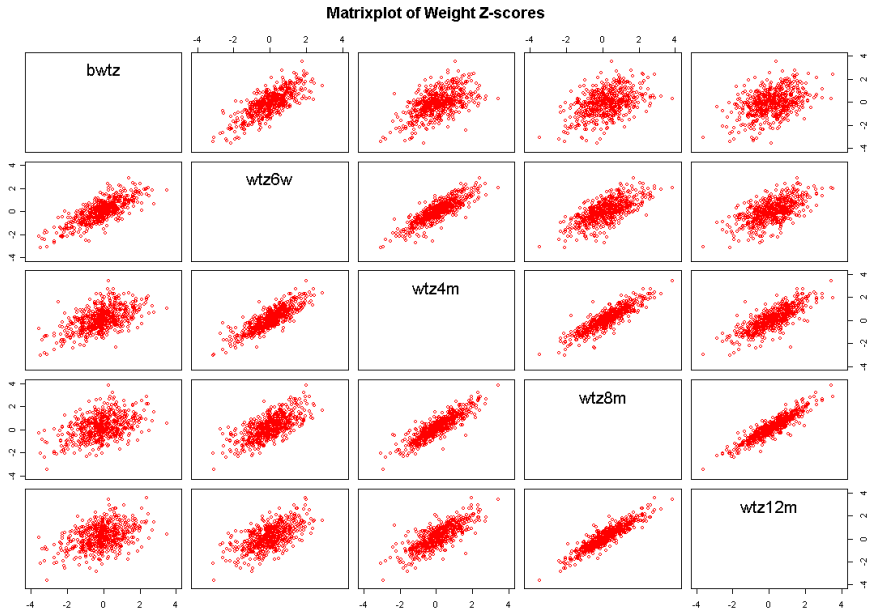


Figure 5.1. ***Matrixplot of Weight Z-scores***

	bwtz	wtz6wk	wtz4m	wtz8m	wtz12m
bwtz	1	0.752	0.555	0.435	0.392
wtz6wk	0.752	1	0.846	0.690	0.583
wtz4m	0.555	0.846	1	0.886	0.767
wtz8m	0.435	0.690	0.886	1	0.916
wtz12m	0.392	0.583	0.767	0.916	1

Figure 5.2. ***Pairwise Correlations for Weight Z-Scores***

Weight Z-scores	**Birth**	**6 Week**	**4 Month**	**8 Month**	**12 Month**
Number who had weight measured	923	780	794	601	774

Table 5.1. ***Number of Babies who had Weight Measured at Each Time Point***

Table 5.1 gives the number of infants who had their weights measured at each of the time points throughout the length of the study. If the 12 month weight

z-scores were imputed using the birthweight z-scores then more of the missing 12 month weight z-scores would be imputed than using, say, the 8 month weight z-scores for imputation purposes as there is a higher response rate for birthweight z-scores than 8 month weight z-scores. However, the correlation between birth and 12 month weight z-scores is 0.392 compared to 0.916 between 8 month and 12 month weight z-scores. Therefore, the imputed 12 month weight z-scores using the birthweight z-scores may not be as reliable as the imputed 12 month weight z-scores using the 8 month weight z-scores (**Figure 5.2**). Looking at **Figures 5.1** and **5.2**, weight z-scores are highly correlated with their neighbouring weight z-scores, therefore using the neighbouring weight z-score appears to be the best method for imputation purposes.

Appetite	**Birth**	**6 Week**	**4 Month**	**8 Month**	**12 Month**
Number who answered appetite qu.	888	749	689	610	578

Table 5.2. ***Number of Mothers who answered Appetite Question at Each Time Point***

The missing appetites rated at 6 weeks and 12 months could be imputed using appetites rated at birth, 6 weeks, 4 months, 8 months and 12 months although this may not be the best approach to use in this instance as early and late appetite rates are related to different aspects of feeding.

The remainder of this chapter focusses on the various imputation methods used to impute the missing values in the Gateshead Millennium Study data.

5.1 EM Algorithm

In this section, the EM Algorithm is used to estimate and impute the missing values of TI0-12m, appetite rated at 6 weeks and appetite rated at 12 months to investigate what effect these imputations have on the results of the complete case analyses (**Table 4.3**).

Example 5.1 - Imputing 12 Month Weight Z-Scores using Birthweight Z-Scores

To illustrate the use of the EM Algorithm, the missing 12 month weight z-scores are estimated using the birthweight z-scores and these imputed 12 month weight z-scores are used in Equation 2.1 to obtain the TI0-12m scores.

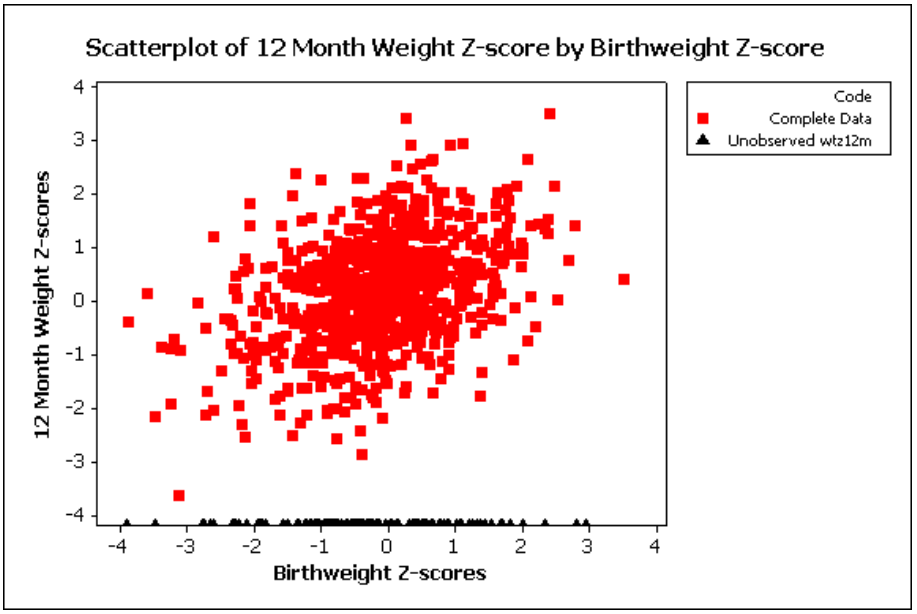


Figure 5.3. ***Scatterplot of Weight Z-scores***

The process of imputing the 12 month weight z-scores using the birthweight z-scores begins by sorting the data into a suitable form for input into the **R** (R Development Core Team, 2010) em.norm function. This is achieved by arranging

the data by its missing data pattern (Section 2.2.1.2). The missing data pattern for this example is shown in **Table 5.3**.

Variable no. of rows	**bwtz (x)**	**wtz12m (y)**
774	0	0
149	0	1

Table 5.3. ***Missing Data Pattern***
The no. of rows represent the number of mothers with that particular pattern of missing data across the weight z-scores. A value of 0 in the table corresponds to an observed weight z-score and a value of 1 in the table corresponds to an unobserved weight z-score.

Table 5.3 and **Figure 5.3** show that of the 923 babies included in the study, 774 had observed values for both birth and 12 month weight z-scores and 149 had observed birthweight z-scores but their 12 month weight z-scores were missing. Once the data has been arranged by its missing data pattern, we then run the EM Algorithm using the **R** (R Development Core Team, 2010) em.norm function until convergence. **Table 5.4** shows the parameter estimates at each iteration.

t	μ_x	μ_y	σ_x^2	σ_{xy}	σ_y^2	ρ_{xy}
1	-0.2037	0.13752	1.17837	0.36286	1.07318	0.32267
2	-0.2037	0.12921	1.17837	0.43182	1.07631	0.38344
3	-0.2037	0.12628	1.17837	0.44515	1.07827	0.39492
4	-0.2037	0.12551	1.17837	0.44776	1.07892	0.39711
5	-0.2037	0.12532	1.17837	0.44828	1.0791	0.39754
6	-0.2037	0.12528	1.17837	0.44838	1.07914	0.39762
7	-0.2037	0.12527	1.17837	0.44841	1.07915	0.39764
8	-0.2037	0.12527	1.17837	0.44841	1.07916	0.39764
9	-0.2037	0.12527	1.17837	0.44841	1.07916	0.39764
10	**-0.2037**	**0.12527**	**1.17837**	**0.44841**	**1.07916**	**0.39764**
11	-0.2037	0.12527	1.17837	0.44841	1.07916	0.39764
∞	-0.2037	0.12527	1.17837	0.44841	1.07916	0.39764

Table 5.4. Iterations of the EM Algorithm

Table 5.4 shows that it takes 10 iterations for the EM Algorithm to converge to the maximum likelihood estimates of the parameters. **Table 5.4** also shows that the two parameters relating to the birthweight z-scores, μ_x and σ_x^2, converge in a single step regardless of the starting value because there are no missing values for the birthweight z-scores so the maximum likelihood estimates are the sample mean and sample variance of the birthweight z-scores, respectively.

The maximum likelihood estimates obtained from the EM Algorithm (**Table 5.4**) are then used in the equation, $\mathbf{E}(Y|X = x) = \alpha + \beta x$, to impute a single value for each of the missing 12 month weight z-scores.

The imputed 12 month weight z-scores are calculated as follows:

Let Y = 12 month weight z-score, X = birthweight z-score and

$\mathbf{E}(Y|X = x) = \alpha + \beta x$, where $\alpha = \mathbf{E}(Y) - \frac{\rho_{XY}\sqrt{Var(Y)}}{\sqrt{Var(X)}}\mathbf{E}(X)$ and $\beta = \frac{\rho_{XY}\sqrt{Var(Y)}}{\sqrt{Var(X)}}$.
Using the maximum likelihood estimates of the parameters from **Table 5.4**, α, β and $\mathbf{E}(Y|X = x) = \alpha + \beta x$ are as follows:

$\beta = \frac{0.3976431\sqrt{1.0791573}}{\sqrt{1.1783663}} = 0.3805359$ and

$\alpha = 0.1252685 - (\beta \times -0.2037053) = 0.2027857$

so the missing 12 month weight z-scores are imputed using the following formula:

$$\mathbf{E}(Y|X = x) = 0.2027857 + (0.3805359 \times x)$$

where x is the birthweight z-score corresponding to the missing 12 month weight z-score.

The scatterplot of the 12 month weight z-scores against the birthweight z-scores, **Figure 5.4**, shows the birthweight and 12 month weight z-scores for the 774 babies whose birthweights and 12 month weights were observed and also shows the birthweight z-scores and imputed 12 month weight z-scores for the 149 babies whose birthweights were observed but their 12 month weights were not.

As can be seen from **Figure 5.4**, the EM Algorithm is estimating and imputing the missing 12 month weight z-scores along the regression line which means that the imputed Thrive Index will be 0.2028 in all 149 missing cases. The Supplemented EM Algorithm or Multiple Imputation will need to be used in conjunction with the EM results to allow for fluctuations in the data from the regression line.

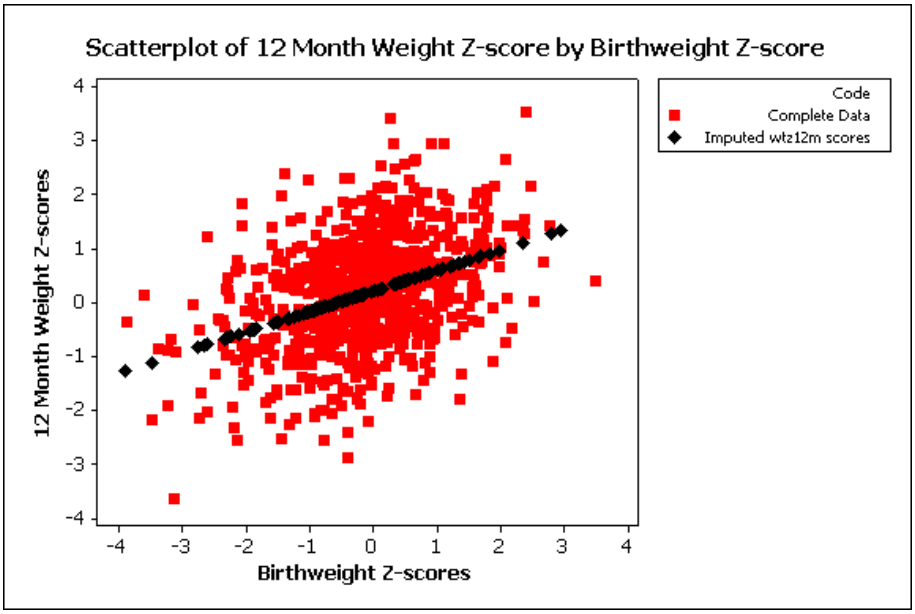


Figure 5.4. ***Scatterplot of Weight Z-scores***

As mentioned previously, the EM Algorithm does not produce estimates of the observed covariance matrix which are needed to obtain confidence intervals for the parameter estimates. In order to obtain estimates of the observed covariance matrix, the Supplemented EM Algorithm is used. The Supplemented EM Algorithm can be used in this instance to find the desired variance-covariance matrix as the complete-data asymptotic variance-covariance matrix is known.

	TI, Mean (SD)	**n**
Appetite rated at 6 weeks		
Normal	0.27 (0.89)	537
Borderline	0.07 (0.88)	193
Low	- 0.23 (0.86)	19
Appetite rated at 12 months		
Normal	0.33 (0.92)	280
Borderline	0.15 (0.95)	226
Low	- 0.09 (0.95)	72

Table 5.5. Mean (SD) values for Thrive Index from birth to 12 months not accounting for the missing 12 month weight z-scores being estimated and imputed using the EM Algorithm

	TI, Mean (SD)	**n**
Appetite rated at 6 weeks		
Normal	0.27 (0.94)	537
Borderline	0.07 (0.96)	193
Low	- 0.23 (0.99)	19
Appetite rated at 12 months		
Normal	0.33 (0.95)	280
Borderline	0.15 (0.97)	226
Low	- 0.09 (0.98)	72

Table 5.6. Mean (SD) values for Thrive Index from birth to 12 months accounting for the missing 12 month weight z-scores being estimated and imputed using the EM Algorithm

Table 5.5 shows the means and standard deviations for the Gateshead Millennium Study data when the 12 month weight z-scores have been imputed using the birthweight z-scores. These results have not taken into account the fact that the missing values in the analysis have been estimated and imputed using the EM Algorithm so the standard deviations will be underestimated and the results of any subsequent analyses will be invalid e.g. if an analysis of variance was to be performed using these standard errors, the p-values would be smaller than expected and so the analysis of variance may give a significant result when in fact there is a non-significant result. **Table 5.6** shows the means and standard deviations for the Gateshead Millennium Study data when the 12 month weight z-scores have been imputed using the birthweight z-scores once the SEM algorithm has been used to take account of the fact that the missing values in the dataset have been estimated and imputed via the EM algorithm.

Returning to the full missing data problem in the Gateshead Millennium Study, the 12 month weight z-scores are estimated and imputed in a variety of ways using **R**'s (R Development Core Team, 2010) em.norm function and the 6 week and 12 month appetite rates are estimated and imputed in a variety of ways using **R**'s (R Development Core Team, 2010) em.cat function. The em.norm function uses multivariate normal models to obtain the maximum likelihood estimates of the parameters and the em.cat function uses log linear models to obtain the maximum likelihood estimates of the parameters.

Appendix D.1 shows the results of the analyses of variance for linear trends for all of the possible ways in which TI0-12m, Appetite rated at 6 weeks and Appetite rated at 12 months can be estimated and imputed using the EM Algorithm.

Although all of the possible ways of imputing the missing values have be employed, it was decided, by the research team, that appetite rated at 4 months should be used to impute appetite rated at 6 weeks as these appetite rates were

related to milk feeding and appetite rated at 8 months should be used to impute appetite rated at 12 months as these appetite rates were related to solid feeding. Appetite rated at birth was not used to impute any of the missing appetite rates that are of interest as they may not give an adequate representation of the child's/childrens' appetite as all mothers may not have had sufficient time to establish their child's/childrens' appetite and some mothers may have nothing to base or compare their initial rating to. It was not discussed with the research team the best way in which to impute the missing 12 month weight z-scores but it is apparent that the best method would be to use the 8 month weight z-scores as the observed 12 month weight z-scores and 8 month weight z-scores are highly correlated (**Figure 5.2**).

The results for the six separate analysis of variance for linear trends and the multiple linear regression for the chosen EM imputation model, in accordance with the research team, are given in **Table 5.7**.

	TI, Mean (SD)	**n**	**p**^{a}	**p**^{b}
Appetite rated at 6 weeks				
Normal	0.26 (0.97)	663		
Borderline	0.03 (0.94)	234		
Low	- 0.29 (0.93)	26	0.0001	0.0150
Appetite rated at 12 months				
Normal	0.27 (0.95)	433		
Borderline	0.14 (0.97)	388		
Low	0.01 (1.01)	101	0.0055	0.0043
AEB rated at 12 months				
Low	0.34 (0.90)	142		
Medium	0.22 (0.96)	261		
High	0.07 (0.99)	175	0.0123	
MFA rated at 12 months				
Normal	0.29 (0.95)	401		
Borderline	0.08 (0.91)	123		
Low	- 0.10 (1.02)	54	0.0012	
RTFR rated at 8 months				
Low	0.28 (0.90)	302		
Medium	0.17 (1.03)	240		
High	0.06 (1.01)	63	0.0677	
RTFR rated at 12 months				
Low	0.31 (0.91)	269		
Medium	0.16 (0.96)	241		
High	- 0.04 (1.10)	66	0.0050	0.0224

Table 5.7. Relationship Between Feeding and Eating Behaviour and Weight Gain from Birth to 12 Months* using EM Algorithm

* values are mean (SD) Thrive Index from birth to 12 months
p^a gives the resulting p-values for the ANOVA for linear trends and p^b gives the p-values of the explanatory variables included in the Multiple Linear Regression including all other significant variables

When the missing 12 month weight z-scores were imputed using the 8 month weight z-scores and the appetites rated at 6 weeks and 12 months were imputed using the appetites rated at 4 months and 8 months, respectively, it was found, from the six separate analysis of variance for linear trends, that all of the explanatory variables except RTFR rated at 8 months were significantly related to weight gain from birth to 12 months (**Table 5.7**). When the multiple linear regression was performed (**Table 5.7**), Appetite rated at 6 weeks, Appetite rated at 12 months and Response to Food Refusal rated at 12 months were significantly related to Thrive Index from birth to 12 months when added to the model together. Comparing these results to the results obtained for the complete case analyses, we can see that they are fairly similar with the only difference being that RTFR rated at 8 months was not significantly related to weight gain from birth to 12 months when the missing TI0-12m scores, appetites rated at 6 weeks and appetites rated at 12 months were estimated and imputed via the EM Algorithm.

All 923 subjects, eligible to be included in the published analysis (**Chapter 4**), have been included in the analysis once the EM Algorithm has been used to estimate and impute the missing TI0-12m scores, appetite rated at 6 weeks and appetite rated at 12 months.

	N	**Mean**	**St. Dev.**
Included in CC Analysis	678	-0.16	1.08
Further Included in Analysis	245	-0.33	1.10

Table 5.8. Mean and Standard Deviations for Birthweight z-scores for children included in the TI0-12M $\sim$ 6 Week Appetite Rates complete case analysis and those further included after Imputation via the EM Algorithm

	N	Mean	St. Dev.
Included in CC Analysis	569	-0.11	1.06
Further Included in Analysis	354	-0.36	1.11

Table 5.9. Mean and Standard Deviations for Birthweight z-scores for children included in the TI0-12M $\sim$ 12 Month Appetite Rates complete case analysis and those further included after Imputation via the EM Algorithm

Looking at **Tables 5.8 and 5.9**, those children who have been further included in the analysis, using the EM Algorithm approach to handling missing data, appear to have lower birthweight z-scores than those children included in the Complete Case analysis. This difference is found to be statistically significant when a two-sample t-test is performed (p = 0.03271, p = 0.00088).

	wtz12m
bwtz	0.428
wtz6wk	0.574
wtz4m	0.742
wtz8m	0.9
wtz12m	1

Figure 5.5. ***Pairwise Correlations for Weight Z-Scores after Imputing 12 Month Weight Z-scores using the EM Algorithm***

Figure 5.5 shows that after the 12 month weight z-scores have been estimated and imputed using the EM Algorithm approach, the correlation structure is fairly similar to that of **Figure 5.2**, showing that the EM Algorithm approach preserves the correlation between variables.

5.2 Single Hot Deck Imputation

The method of Single Hot Deck Imputation implemented in this thesis is based on Example 4.8: Hot Deck Within Adjustment Cells of Little and Rubin (2002). It involves splitting the subjects into groups depending on their previous appetite measurements so that similar responding subjects are in the same group. Missing values within each group are then replaced by recorded values from respondents in the same group via simple random sampling with replacement. This approach was performed using the **impute** function in the **Hmisc** library (Harrell, F. E. and with contributions from many other users, 2007) in **R** after some manipulation of the data i.e. splitting the subjects into groups depending on their previous appetite rates.

Example 5.2 - Imputing 12 Month Appetite Rates using 8 Month Appetite Rates

To illustrate the use of the Single Hot Deck Imputation method, the missing 12 month appetite rates are imputed using the 8 month appetite rates.
In order to obtain reasonable imputed values for the missing 12 month appetite rates in the dataset, babies in the Gateshead Millennium Study are split into groups depending on their 8 month appetite rate, 'Normal', 'Borderline' or 'Low'. Babies with missing 12 month appetite rates in each of the groups will be imputed by a 12 month appetite rate from a respondent in the same group.

	TI, Mean (SD)	**n**
Appetite rated at 12 months		
Normal	0.29 (0.94)	334
Borderline	0.18 (0.95)	228
Low	- 0.08 (0.95)	92

Table 5.10. Mean (SD) values for Thrive Index from birth to 12 months not accounting for the missing 12 month weight z-scores being imputed using SHDI

	TI, Mean (SD)	**n**
Appetite rated at 12 months		
Normal	0.29 (0.94)	334
Borderline	0.18 (0.96)	228
Low	- 0.08 (0.97)	92

Table 5.11. Mean (SD) values for Thrive Index from birth to 12 months accounting for the missing 12 month weight z-scores being imputed using SHDI

Table 5.10 shows the means and standard deviations for the Gateshead Millennium Study data when the 12 month appetite rates have been imputed using the 8 month appetite rates. These results have not taken into account the fact that the missing values in the analysis have been imputed using Single Hot Deck Imputation so the standard deviations will be underestimated and the results of any subsequent analyses will be invalid. **Table 5.11** shows the means and standard deviations for the Gateshead Millennium Study data when the 12 month appetite

rates have been imputed using the 8 month appetite rates once the adjusted jackknife variance estimator (Rao and Shao, 1992) has been used to account for the true uncertainty due to non-response.

Returning to the full missing data problem in the Gateshead Millennium Study, the 12 month weight z-scores, the 6 week appetite rates and the 12 month appetite rates are imputed in a variety of ways.

Appendix D.2 shows the results of the analyses of variance for linear trends for all of the possible ways in which TI0-12m, Appetite rated at 6 weeks and Appetite rated at 12 months can be imputed.

Although all of the possible ways of imputing the missing values have been employed, it was decided that appetite rated at 4 months should be used to impute appetite rated at 6 weeks, appetite rated at 8 months should be used to impute appetite rated at 12 months and the missing 12 month weight z-scores should be imputed using the 8 month weight z-scores as they are highly correlated (**Figure 5.2**).

The results for the six separate analysis of variance for linear trends and the multiple linear regression for the chosen SHDI model, in accordance with the research team, are given in **Table 5.12**.

	TI, Mean (SD)	**n**	**p**^{a}	**p**^{b}
Appetite rated at 6 weeks				
Normal	0.24 (0.96)	620		
Borderline	0.05 (0.94)	226		
Low	- 0.29 (0.87)	21	0.0007	0.0248
Appetite rated at 12 months				
Normal	0.34 (0.93)	312		
Borderline	0.13 (0.97)	274		
Low	- 0.10 (0.95)	82	0.0001	0.0028
AEB rated at 12 months				
Low	0.34 (0.91)	142		
Medium	0.22 (0.96)	261		
High	0.07 (0.97)	175	0.0106	
MFA rated at 12 months				
Normal	0.28 (0.95)	401		
Borderline	0.08 (0.91)	123		
Low	- 0.11 (1.03)	54	0.0011	
RTFR rated at 8 months				
Low	0.27 (0.90)	302		
Medium	0.15 (1.00)	240		
High	0.05 (1.01)	63	0.0553	
RTFR rated at 12 months				
Low	0.31 (0.92)	269		
Medium	0.15 (0.94)	241		
High	- 0.05 (1.10)	66	0.0031	0.0165

Table 5.12. Relationship Between Feeding and Eating Behaviour and Weight Gain from Birth to 12 Months* using SHDI

* values are mean (SD) Thrive Index from birth to 12 months
p^a gives the resulting p-values for the ANOVA for linear trends and p^b gives the p-values of the explanatory variables included in the Multiple Linear Regression including all other significant variables

When the missing 12 month weight z-scores were imputed using the 8 month weight z-scores and the appetites rated at 6 weeks and 12 months were imputed using the appetites rated at 4 months and 8 months, respectively, it was found, from the six separate analysis of variance for linear trends, that all of the explanatory variables except RTFR rated at 8 months were significantly related to weight gain from birth to 12 months (**Table 5.12**). When the multiple linear regression was performed (**Table 5.12**), Appetite rated at 6 weeks, Appetite rated at 12 months and Response to Food Refusal rated at 12 months were significantly related to Thrive Index from birth to 12 months when added to the model together. Comparing these results to the results obtained for the complete case analyses, we can see that they are fairly similar with the only difference being that RTFR rated at 8 months was not significantly related to weight gain from birth to 12 months when the missing TI0-12m scores, appetites rated at 6 weeks and appetites rated at 12 months were imputed via the Single Hot Deck Imputation missing data method.

Unlike the EM Algorithm method, the Single Hot Deck Imputation procedure does not include all 923 subjects in the analysis once the missing values for TI0-12m scores, appetite rated at 6 weeks and appetite rated at 12 months have been imputed. It further includes in the analysis only those children who are most like the children included in the Complete Case analysis and does not include those children with significantly lower birthweight z-scores.

Figure 5.6 shows that after the 12 month weight z-scores have been imputed using the Single Hot Deck Imputation approach, the correlation structure is approximately the same as the correlation structure in **Figure 5.2**, showing that the Single Hot Deck Imputation is better at preserving the correlation between the variables than the EM Algorithm.

	wtz12m
bwtz	0.374
wtz6wk	0.582
wtz4m	0.757
wtz8m	0.911
wtz12m	1

Figure 5.6. ***Pairwise Correlations for Weight Z-Scores after Imputing 12 Month Weight Z-scores using SHDI***

5.3 Multiple Hot Deck Imputation

The method of Multiple Hot Deck Imputation implemented in this thesis involves repeating the Single Hot Deck Imputation method (**Section 5.2**) a number of times to create multiple 'completed' datasets to which standard statistical techniques can be applied and which allows us to obtain a single parameter estimate which properly reflects the uncertainty due to non-response. The results obtained from analysing each of the multiple datasets using a standard statistical technique are combined using the formulae given in **Section 2.2.3.2**, to obtain a single parameter estimate. For the Gateshead Millennium Study data, 10 'completed' datasets were created. The rates of missing information are 0.27 and 0.38 for the two analyses of interest, TI0-12m $\sim$ 6 Week Appetite and TI0-12m $\sim$ 12 Month Appetite, respectively. The efficiency of the Multiple Imputation method for the two analyses of interest when 10 'completed' datasets are created is 99% from **Table 2.10**.

Appendix D.3 shows the results of the analyses of variance for linear trends for all of the possible ways in which TI0-12m, Appetite rated at 6 weeks and Appetite rated at 12 months can be imputed.

Although all of the possible ways of imputing the missing values have been employed, as mentioned previously, it was decided that appetite rated at 4 months

should be used to impute appetite rated at 6 weeks, appetite rated at 8 months should be used to impute appetite rated at 12 months and the missing 12 month weight z-scores should be imputed using the 8 month weight z-scores as they are highly correlated (**Figure 5.2**).

	Number of Cases Included	**F-Statistic**	**P-Value**
Analysis 1	569	12.99	0.0003
Analysis 2	578	12.41	0.0006
Analysis 3.1	640	12.45	0.0008
Analysis 3.2	714	14.91	0.0002
Analysis 4.1	668	11.25	0.0018
Analysis 4.2	799	10.48	0.0034

Table 5.13. Table of Results for TI0-12m $\sim$ 12 Month Appetite Rate.

Notes on Table 5.13

Analysis 1 is the complete case analysis. **Analysis 2** is the analysis where only wtz12m is imputed. **Analysis 3.1** is the analysis where only 12 month appetite is imputed using 8 month appetite. **Analysis 3.2** is the analysis where only 12 month appetite is imputed using 8 month, 4 month and 6 week appetite. **Analysis 4.1** is where both wtz12m and 12 month appetite using 8 month appetite are imputed. **Analysis 4.2** is where both wtz12m and 12 month appetite using 8 month, 4 month and 6 week appetite are imputed.

Table 5.13 shows the results of the analysis of variance for linear trend for the relationship between TI0-12m and 12 Month Appetite rate by imputing the data in different ways. When the number of cases included in the analysis increases compared to the number of cases included in the Complete Case Analysis (**Analysis 1**), the p-value also increases except in **Analysis 3.2**. Although the

p-values have increased compared to the p-value for the Complete Case Analysis, appetite rate at 12 months is still significantly related to TI0-12m in all of the analyses.

The Complete Case Analysis appears to be giving a more positive outcome than is justified. This is likely to mean that the children for whom we have all their data available are not representative of the cohort. The children who have their appetite rate, weight z-score at 12 months or both imputed in **Analysis 2** - **Analysis 4.2** appear to be different from the children who were included in the Complete Case Analysis. For this reason, we will have to look at these children and investigate why they are different from the children included in the Complete Case Analysis i.e. did they drop out of study early due to low weights/appetites?, were they from different social classes?, etc.

These children were found, through exploratory statistics, to be different from the children included in the Complete Case Analysis because they were from a lower social class and had lower birthweights. Of the children not included in the analysis, more were likely to have missing appetite rates than missing weights. This could be due to the fact that in some of the questionnaires, the answers to the appetite question were not in descending order (**Section 2.1.1**) so some parents, especially those from lower social classes, may have been a bit confused and therefore did not answer the question.

The results for the six separate analysis of variance for linear trends and the multiple linear regression for the chosen MHDI model, in accordance with the research team, are given in **Table 5.14**.

	TI, Mean (SD)	**n**	**p**^{a}	**p**^{b}
Appetite rated at 6 weeks				
Normal	0.26 (0.96)	627		
Borderline	0.05 (0.93)	220		
Low	- 0.27 (0.86)	20	0.0004	0.0315
Appetite rated at 12 months				
Normal	0.34 (0.93)	312		
Borderline	0.15 (0.99)	274		
Low	- 0.09 (0.95)	82	0.0001	0.0046
AEB rated at 12 months				
Low	0.34 (0.91)	142		
Medium	0.22 (0.96)	261		
High	0.07 (0.97)	175	0.0111	
MFA rated at 12 months				
Normal	0.28 (0.95)	401		
Borderline	0.09 (0.91)	123		
Low	- 0.11 (1.02)	54	0.0012	
RTFR rated at 8 months				
Low	0.28 (0.90)	302		
Medium	0.17 (1.01)	240		
High	0.07 (1.03)	63	0.0661	
RTFR rated at 12 months				
Low	0.31 (0.92)	269		
Medium	0.15 (0.94)	241		
High	- 0.05 (1.10)	66	0.0033	0.0153

Table 5.14. Relationship Between Feeding and Eating Behaviour and Weight Gain from Birth to 12 Months using MHDI

* values are mean (SD) Thrive Index from birth to 12 months
p^a gives the resulting p-values for the ANOVA for linear trends and p^b gives the p-values of the explanatory variables included in the Multiple Linear Regression including all other significant variables

As for the EM Algorithm and Single Hot Deck Imputation methods, the Multiple Hot Deck Imputation method found that all of the explanatory variables except RTFR rated at 8 months were significantly related to weight gain from birth to 12 months (**Table 5.14**). When the multiple linear regression was performed, (**Table 5.14**) the only variables to be significantly related to Thrive Index from birth to 12 months, when added to the model together, were Appetite rated at 6 weeks, Appetite rated at 12 months and Response to Food Refusal rated at 12 months. Comparing these results to the results obtained for the complete case analyses, we can see that they are fairly similar with the only difference being that RTFR rated at 8 months was not significantly related to weight gain from birth to 12 months when the missing TI0-12m scores, appetites rated at 6 weeks and appetites rated at 12 months were imputed via the Multiple Hot Deck Imputation missing data method.

Chapter 6

Discussion and Conclusions

6.1 Conclusions

The Gateshead Millennium Study is a prospective cohort study of feeding and growth in infancy. This study was set up primarily to explore the relationship between development of growth and feeding in the first year of life. Babies born between 1 June 1999 and 31 May 2000 in the Gateshead area of northeast England were recruited to the study shortly after birth.

Within the recruitment year of the Gateshead Millennium Study, approximately two weeks in every three were assigned to be recruitment weeks and babies born in these pre-specified 34 recruitment weeks were eligible for recruitment to the study. As well as the child being born in Gateshead in one of the pre-specified recruitment weeks, another criterion for recruitment to the study was that the mother of the child was a Gateshead resident at the time of delivery.

Of all births and multiple births in the 34 recruitment weeks, a total of 1029 (83%) babies of 1011 mothers were recruited to the study (shortly after the birth).

Mothers who agreed to participate in the study had a face-to-face interview

shortly after recruitment, during which baseline information, including birthweight and socio-demographic data, was recorded. Participating parents also completed a questionnaire at recruitment and received postal questionnaires at 6 weeks, 4 months, 8 months, 12 months and 30 months to complete and return. A wide range of feeding questions were asked in each of the questionnaires as well as questions about different aspects of the mother and child. On the front of each questionnaire, parents were also asked to transcribe all weights which the child had measured since completing and returning the previous questionnaire.

The main objective of this thesis was to explore different approaches to handling missing data and their impact on the results of the various key analyses which have already been performed and published for the Gateshead Millennium Study data in the **How Does Maternal and Child Feeding Behaviour Relate to Weight Gain and Failure to Thrive? Data From a Prospective Birth Cohort** paper by Wright et al. (2006a).

Missing data is a commonly occurring problem which can lead to biased and possibly misleading non-significant results if the missing data are not dealt with in the correct manner. For this reason, it is important to consider why the data are missing and whether or not missingness is related to the practical questions being investigated using the data.

There are several reasons why, in certain studies, missing values may occur and the missing data mechanism (**Section 2.2.1.3**) shows the mechanism by which the missing data may have arisen. There are three different missing data mechanisms which may be encountered depending on whether or not the fact that a particular value is missing is linked to the underlying values. These are Missing Completely at Random (MCAR), Missing at Random (MAR) and Not Missing at Random (NMAR). The statistical approach used to impute the missing data

is different depending on which of these missing data mechanisms are in operation. Another characteristic of missing data that will influence which statistical method can be used to analyse the data is the missing data pattern. This shows which values in the data matrix are observed and which are missing. In **Section 2.2.1.2**, two patterns of missing data were considered, monotone and general non-monotone missing data patterns. The Gateshead Millennium Study data suffers from a general non-monotone missing data pattern, as some of the mothers are not completing and returning the questionnaires at any one or more of the pre-specified times, and so missing data can occur anywhere in the dataset. The type of missing data pattern was taken into account when deciding which approaches to handling missing data to use.

In **Chapter 3**, the extent of missing data was evaluated by creating a comprehensive description of the response rate to each of the questions in each of the questionnaires. The extent of the missing data in the Gateshead Millennium Study is not as large as the fraction of missing data that would be expected in a routine longitudinal study as a number of tactics were decided upon when designing the study to improve response rates and to ensure the success of the study.

In **Chapter 4**, the complete case analyses that were performed for and published in the **How Does Maternal and Child Feeding Behaviour Relate to Weight Gain and Failure to Thrive? Data From a Prospective Birth Cohort** paper by Wright et al. (2006a), in order to determine which variables were significantly related to Thrive Index in the first year of life, were repeated and used to assess whether or not there was any evidence against the Missing Completely at Random assumption. Each possible explanatory variable - Appetite rated at 6 weeks, Appetite rated at 12 months, Avoidant Eating Behaviour rated at 12 months, Maternal Feeding Anxiety rated at 12 months, Response to

Food Refusal rated at 8 months and Response to Food Refusal rated at 12 months - is included in an analysis of variance for linear trend in order to determine if that specific explanatory variable, on its own, is significantly related to Thrive Index. All six potential explanatory variables are then included in a multiple linear regression to determine which variables, if any, are significantly related to Thrive Index when other explanatory variables are already included in the model. From the six separate analysis of variance for linear trends, it was found that all of the explanatory variables were significantly related to weight gain from birth to 12 months. However, when the multiple linear regression was performed, only Appetite rated at 6 weeks, Appetite rated at 12 months and Response to Food Refusal rated at 12 months were significantly related to Thrive Index from birth to 12 months when added to the model together. When assessing whether or not there was any evidence against the Missing Completely at Random assumption in **Chapter 4**, it was found that the complete case analysis method may not be an appropriate way in which to analyse the Gateshead Millennium Study data as the missing data are not a random sample of all of the data i.e. the MCAR assumption is questionable, and so the above results from the Complete Case analysis might not be representative of the population as a whole and should be treated with caution. For this reason, a number of alternative methods were used which rely on the assumption of the data being Missing at Random. This is a less restrictive assumption than the assumption of Missing Completely at Random required for the complete case analysis and can be met using the observed data to fill in values for the missing data.

In **Chapter 5**, various missing data methods were used to impute the missing values in the Gateshead Millennium Study. The various missing data methods considered were Single Hot Deck Imputation, Multiple Hot Deck Imputation and the EM Algorithm. The variables with missing data, Thrive Index and Appetite

rated at 6 weeks and 12 months, required for the analyses of variance for linear trends were imputed in a variety of ways using the above missing data methods. It was suggested that Appetite rated at 6 weeks should be imputed using Appetite rated at 4 months as these appetite rates are related to milk feeding and Appetite rated at 12 months should be imputed using Appetite rated at 8 months as these appetite rates are related to solid feeding. The Thrive Index for growth of a child in their first year of life is calculated using birth and 12 month weight z-scores. So, instead of imputing the missing Thrive Index scores directly, the various missing data methods were used to impute the missing 12 month weight z-scores and these imputed values were used along with the observed values for birth and 12 month weight z-scores to calculate the Thrive Index scores. It was suggested that the missing 12 month weight z-scores be imputed using the observed 8 month weight z-scores.

The results for the relationship between Thrive Index from birth to 12 months and appetite rated at 6 weeks (TI0-12M $\sim$ 6 Week Appetite Rates) and for the relationship between Thrive Index from birth to 12 months and appetite rated at 12 months (TI0-12M $\sim$ 12 Month Appetite Rates) using the different approaches to handling missing data are shown in **Figures 6.1** and **6.2**, respectively.

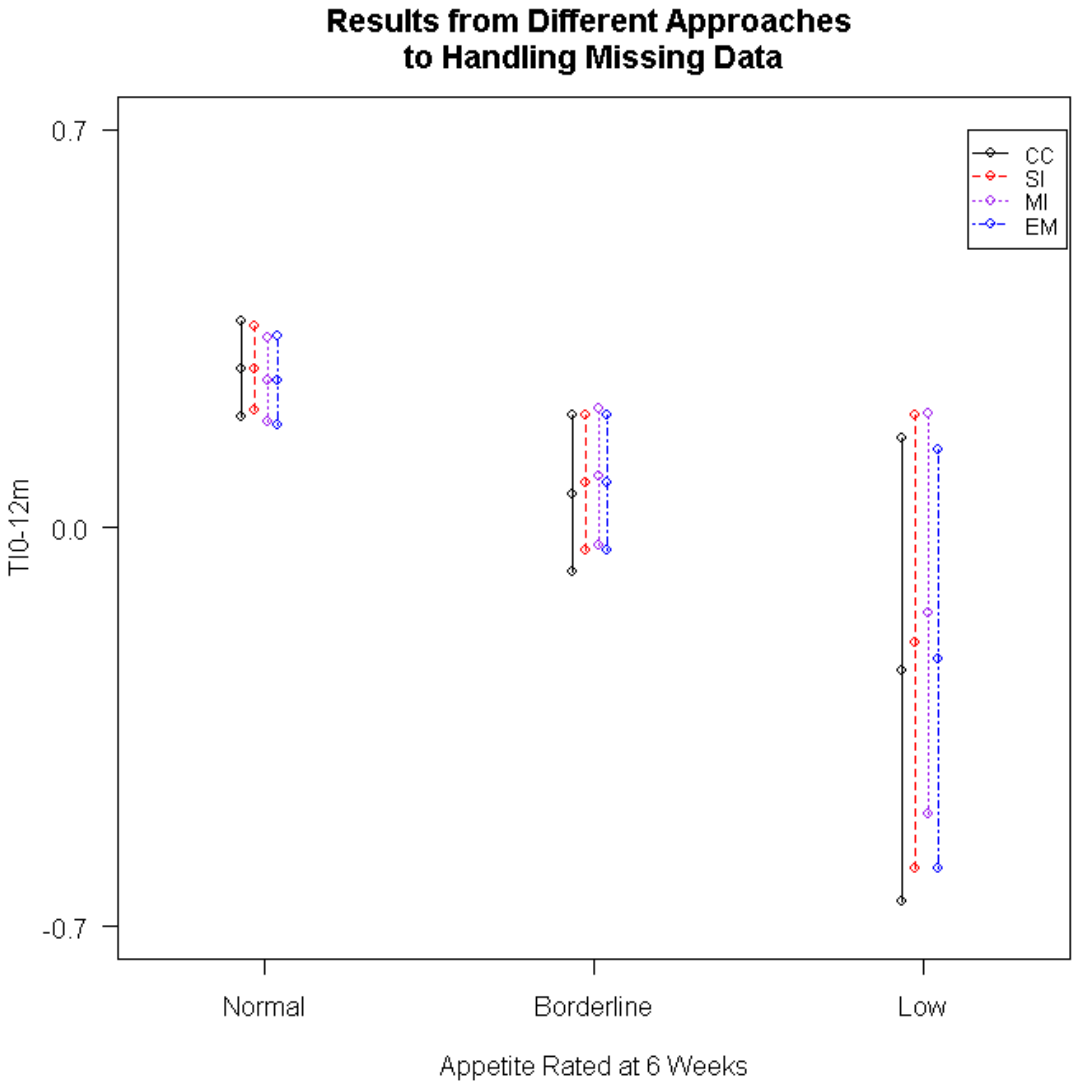


Figure 6.1. ***Results for TI0-12M ∼ 6 Week Appetite Rates***

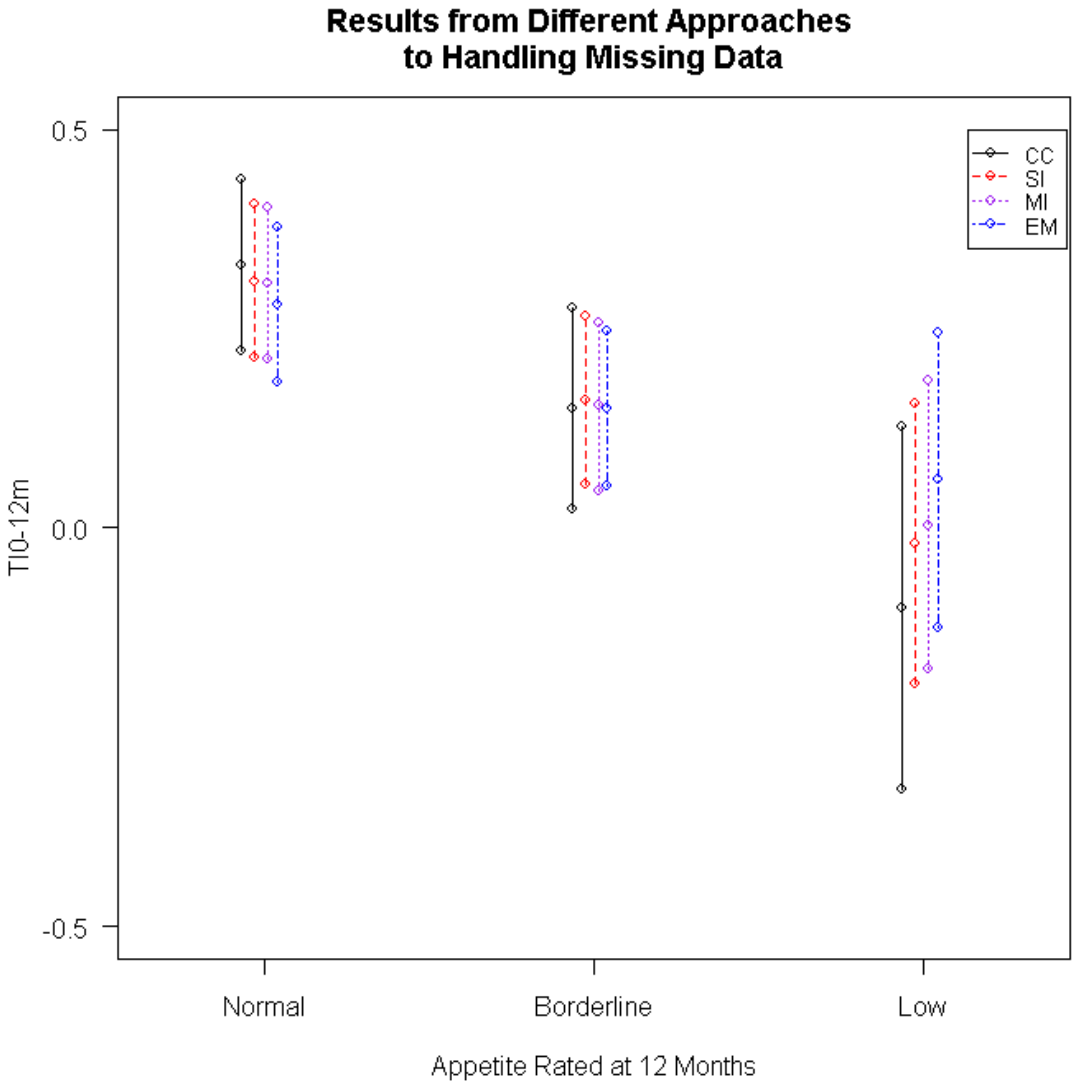


Figure 6.2. ***Results for TI0-12M ∼ 12 Month Appetite Rates***

In **Figures 6.1** and **6.2**, the mean TI0-12m scores for each of the missing data methods within each level of appetite rate (Normal, Borderline and Low), are fairly similar as are the associated 95% confidence intervals. For most of the imputation methods, the 95% confidence intervals for the 'Normal' appetite rates are narrower than the 95% confidence intervals for the 'Borderline' appetite rates

which are in turn narrower than the 95% confidence intervals for the 'Low' appetite rates since there are fewer subjects whose appetite has been rated as 'Low' compared to the number of subjects whose appetite has been rated as 'Normal'. Since all of the missing data techniques used give reasonably similar results, it is feasible to use any of the methods for the Gateshead Millennium Study data. From the Complete Case analysis assumption checking, it has been suggested that the complete case analysis method is not the best way to analyse the Gateshead Millennium Study data as the MCAR assumption is questionable and so an alternative missing data method needs to be used. However, when the alternative methods have been implemented, the results are qualitatively the same as those obtained using the Complete Case analysis method.

Although the results of all of the missing data methods tried are similar, I would suggest using the Multiple Hot Deck Imputation method as it captures the variability in the data due to imputation more effectively than the other methods without having to carry out further calculations, such as the ones required for Single Hot Deck Imputation and the EM Algorithm, to estimate the true uncertainty due to non-response i.e. the Multiple Hot Deck Imputation method is computationally efficient.

6.2 Limitations

The Gateshead Millennium Study was a well designed and thought-out study. The research team employed a number of strategies to improve response rates and ensure the success of the study, including media involvement, support from local health professionals, telephone reminders for questionnaire completion and newsletters. In spite of these efforts to maintain a high level of response, there

was an increasing rate of attrition with a reduction in the questionnaire response rates as time went on (**Table 1.1**).

There is some question as to the suitability of the methods for handling missing data which have been used to impute the missing values in the Gateshead Millennium Study data.

The EM Algorithm (**Section 2.2.4**) does not produce precise estimates for the standard deviation when the missing values are estimated and imputed initially, and therefore the Supplemented EM Algorithm has to be used to obtain the increase in variance due to the missing values being estimated and imputed, hence producing precise standard deviations which account for the additional uncertainty that arises from estimating and imputing the missing data. As mentioned previously, the SEM Algorithm involves a number of difficult steps for calculating the increase in variance due to imputation uncertainty and so it may be worthwhile using another missing data approach.

The Single Hot Deck Imputation method involves filling in one value for every missing value. The now 'complete' dataset is analysed using one of the standard statistical techniques, ignoring the fact that the missing data have been imputed. As with the EM Algorithm method, the results obtained from analysing the 'complete' dataset using standard statistical techniques do not reflect the additional uncertainty that arises from imputing the missing data and therefore a further adjustment has to be made to account for this. The special adjustment used in this instance is the Adjusted Jackknife Variance Estimator (**Section 2.2.3.1**). Once again, calculating the Adjusted Jackknife Variance Estimator to give the increase in variance due to non-response, as with the Supplemented EM Algorithm for the EM Algorithm, could cause problems if the data analyst is not confident in implementing statistical techniques. Therefore, the method of Multiple Imputation (**Section 2.2.3.2**) is by far the best method to use as no further

computation is required in addition to the initial calculations to produce precise estimates for the mean and standard deviation.

The imputation methods used in this thesis are imputing the missing appetite rates and weights using the observed appetite rates and weights of children from birth to 12 months, respectively as using the variables of interest from the research team's original analysis. It was mentioned in Chapter 4 that children not included in the Complete Case analysis (Table 4.3) tended to come from more deprived neighbourhoods and that the gender of the children included and not included in the Complete Case analysis may lead to an apparent difference between the groups in terms of their birthweights. For this reason, the prediction models for the imputation methods should include predictors for the missing appetite rates and weights which are known to affect the appetite rates and weights e.g. gender and deprivation should be included in the prediction models as well as other factors suggested by Wright et al. (2006b). Including more predictors in our prediction models would lead to more complicated patterns of missing data but would produce imputes which are better than those obtained from prediction models with smaller numbers of predictors. Multiple Imputation using Chained Equations (Carpenter and Kenward, 2005) could be used to perform this analysis. All of the imputation methods which have been implemented in this thesis rely theoretically on the assumption of the data being Missing at Random (MAR). Although there is currently no test available to check that the MAR assumption holds for this dataset, there is no reason to believe that the missing data are Not Missing at Random (NMAR).

6.3 Further Work

The imputation methods reviewed here are clearly not the only ones available. Little and Rubin (2002) mention several others which may be of interest (some of which have already been discussed above, in Chapter 2 and Section 6.2). Although some of these other missing data approaches could be used to impute the missing values for the Gateshead Millennium Study data, we have qualitatively confirmed the results of the complete case analyses using the SHDI, MHDI and the EM Algorithm methods, even though the MCAR assumption required for the complete case analysis is in doubt and the proportion of missing data is moderately high.

In this thesis, we were only interested in imputing the missing values for Appetite rated at 6 weeks and 12 months and Thrive Index from birth to 12 months, but it may also be of interest to impute the missing values for the other factors which are related to Thrive Index i.e. Avoidant Eating Behaviour, Maternal Feeding Anxiety and Response to Food Refusal, to investigate what effect these imputations have on the results of the analyses of variance for linear trends and the multiple linear regressions.

The Gateshead Millennium Study was initially set up to explore the relationship between weight gain and appetite, but since its introduction it has been used to analyse other aspects of the children. For this reason, it may be of interest to apply the missing data techniques used in this thesis to the other analyses which have been performed, in order to discover if the results found would change after imputing the missing values.

Since the Gateshead Millennium Study data is now being used to explore the relationship between other variables relating to children, it may be of interest to

produce an overall dataset which has all of the missing values for all of the variables imputed so that it can be used by future researchers who want to analyse certain aspects of the children. If this overall dataset was to be created, adjustments may have to be made to the model used to impute the missing values to include the mechanism which caused the missing data.

Appendices

Appendix A

Questionnaires

Figure A.1. ***Recruitment Questionnaire***

MILLENNIUM BABY STUDY **Feeding and Growth Study**

ID No. ☐

Recruitment Form

Today's date ☐☐ day ☐☐☐ month ☐☐ year

1. Mother's name — Personal name ..... Family name .....
2. Address .....
   .....
   Postcode ..... Telephone number .....
3. Which Council do you pay your Council Tax to? .....
4. Mother's date of birth ☐☐ day ☐☐☐ month ☐☐ year
5. No. of previous children .....
6. Length of gestation ..... weeks
7. Mode of delivery — Normal / Suction / Forceps / Planned Caesarean / Emergency Caesarean
8. Place of birth .....
9. Number of babies — Single / twins / three or more
10. Baby's name — Personal name ..... Family name .....
11. Baby's date of birth ☐☐ day ☐☐☐ month ☐☐ year
12. Sex — Male / female
13. Birth weight ..... g
14. Father's name — Personal name ..... Family name .....
15. Father's date of birth ☐☐ day ☐☐☐ month ☐☐ year
16. Would you describe yourself as
    Caucasian / Indian Sub-continent / Other Asian / Afro-Caribbean / Other:
17. Would you describe the baby's father as
    Caucasian / Indian Sub-continent / Other Asian / Afro-Caribbean / Other:
18. What language do you speak at home? .....
19. What religion would you describe your family as?
    None / Christian / Orthodox Jewish / Non-orthodox Jewish / Muslim / Other:
20. Community midwife .....
21. Clinic .....
22. G.P. .....
23. Would you like help filling in the questionnaires? .....

Figure A.2. ***Newborn Questionnaire***

Feeding and growth study: Your new-born baby

ID No.

Please write the date you complete this: ___/___/___

This questionnaire asks you about you and your new baby.

If, rather than a single baby, you have twins or triplets, please answer the questionnaire in relation to each baby on a different copy of the questionnaire.

If for any reason you do not wish to fill in this questionnaire yourself, you might prefer the research assistant to ask you the questions and fill in the form for you. Please ask and we are happy to help.

How to fill in the questionnaire

1. Some questions on the following pages can be answered simply by putting a tick in the box next to the answer that applies to you.

   Example *Yes* ☐ *No* ☐

2. Some questions on the following pages can be answered by circling the response that applies to you.

   Example *Not at all---Occasionally---Frequently*

   If you really feel that you are in-between two of the descriptions, you can indicate this by circling the dotted line.

Section A: Milk feeding

1. How did you feed your baby at birth? **(tick one only)**
   Breast feeding ☐ *Bottle feeding* ☐ *Both* ☐

2. How are you feeding your baby at the moment? **(tick one only)**
   Breast feeding ☐ *Bottle feeding* ☐ *Both* ☐

Section B: General feeding questions

3. So far, how do you rate how well your baby sucks?
   Strong---Average---Weak---No opinion

4. So far, how do you rate your baby's appetite?
   Very good---Good---All right---Very poor---Poor---No opinion

5. So far, do you think your baby is feeding enough?
   Yes---Not always---No

6. Are feeding times for you:
   Very relaxed---Relaxed---All right---Stressful---Very stressful

7. Are feeding times for your baby:
   Very relaxed---Relaxed---All right---Stressful---Very stressful---Can't tell

8. So far, has your baby been easy to feed?
   Very easy---Easy---All right---Difficult---Very difficult

9. So far, has your baby had any trouble with any of the following:

   (a) Sucking *Not at all---Occasionally---Frequently*
   (b) Swallowing *Not at all---Occasionally---Frequently*
   (c) Choking *Not at all---Occasionally---Frequently*

10. Does any of the following describe your baby at present?

    (a) Has to be woken up for feeds *Not at all---Occasionally---Frequently*
    (b) Sleeps during feeds *Not at all---Occasionally---Frequently*
    (c) Cries during feeds *Not at all---Occasionally---Frequently*
    (d) Slow feeder *Not at all---Occasionally---Frequently*
    (e) Not satisfied *Not at all---Occasionally---Frequently*

11. Has your baby posseted at all yet (brought up small vomits)?
    Rarely---Sometimes---Often

12. Has your baby vomited at all yet (brought up most or all of feed)?
    Rarely---Sometimes---Often

13. Do you see your baby as being:
    Very thin---Thin---Average---Chubby---Fat

Section C: Looking into the future

These are some questions about how you expect to look after your baby in future.
Please tick the most appropriate response to each question.

14. Some mothers think babies should be fed only when they seem hungry. Other mothers feed their baby whenever they think they need it (for example, if too long a time has passed since the last feeding). Would you

Feed your baby when hungry? ☐ *Feed your baby whenever you think your baby needs it* ☐

Something in between? ☐

15. Many mothers think that a baby's regular feeding should not be delayed and will wake their baby up to feed if it is past the regular time. Would you

Let your baby sleep and ignore the time ☐ *Wake your baby up if is late for feeding?* ☐

Something in between? ☐

16. What about when your baby cries? If there is no obvious reason for the crying (your baby is not wet, is not hurt), would you

Try to feed your baby? ☐ *Try to calm your baby by other means, without changing the feeding?* ☐

Something in between? ☐

17. Some mothers worry if other people such as friends or relatives think their baby is not gaining enough weight or is too thin. If this happened to you, would you

Encourage your baby to eat? ☐ *Continue with your usual feeding routine?* ☐

Something in between? ☐

18. Suppose your baby has just been fed and about half an hour later becomes fussy and irritable. Would you

Not feed in between regular feedings just because your baby is fussy? ☐ *Try to feed your baby again?* ☐

Something in between? ☐

19. Supposing you were in the middle of watching your favourite TV programme or doing something else you really enjoy, and it was your baby's normal feeding time. If your baby seemed content, would you

Finish what you were doing and then feed your baby? ☐ *Stop what you were doing to feed your baby?* ☐

Something in between? ☐

20. Sometimes when babies get older they seem not to like new foods. Would you

Persist in offering your baby the new food for at least a week before giving up? ☐ *Only try once or twice and then try another food?* ☐

Something in between? ☐

21. When babies are sick with a cold or the flu, they often lose their appetite. If this happened, would you

Try to encourage your baby to eat? ☐ *Wait until your baby felt like eating normally again, even if you felt that your baby was not getting enough?* ☐

Something in between? ☐

22. Older children often refuse to eat everything they are given at a meal. Would you

Permit your child to refuse? ☐ *Encourage your child to eat everything?* ☐

Something in between? ☐

Is there anything else you would like to say about feeding your baby?

If so, **please give details below:**

Section D: General information

23. Please look down the list and state whether you have any of the qualifications listed. Start at the top of the list and tick all the ones that you have passed. **(tick all that apply)**

(a)	*Degree (or degree level qualification)*	☐	(f)	*NVQs*	☐
(b)	*Nursing qualifications*	☐	(g)	*No formal qualifications*	☐
(c)	*'A' levels*	☐	(h)	*Not yet finished education*	☐
(d)	*Scottish highers*	☐	(i)	*Did not go to school*	☐
(e)	*'O' level passes/GCSE/CSE/GNVQ*	☐	(j)	*Other qualifications (please state)*	☐

24. Does anyone in your household earn a wage at present?

Yes ☐ *No* ☐

25. Are you **(tick one only)**

Married, living with husband ☐ *Living with partner* ☐

Single/separated, living with parents ☐ *Single, living alone* ☐

Other ***(please tick and specify)*** ☐

26. Does your household own or rent your house or flat? **(tick one only)**

Owns with mortgage/loan/outright ☐ *Rents* ☐ *Rent free* ☐

27. (a) Does anyone in your household own a car?

Yes ☐ *No* ☐

27. (b) If no, do you have the use of a car?

Yes ☐ *No* ☐

Was there anything you intended to go back to and complete? Please check.

When you have finished please give the questionnaire to the researcher, even if you were not able to answer all of it.

We will be in touch when your baby is six weeks old.

Thank you very much for your help.

Figure A.3. ***6 Week Questionnaire***

Feeding and growth study: Your baby at six weeks

ID No:

This questionnaire asks about you and your baby. If, for any reason, your baby is no longer with you, please tick the box below and return the questionnaire to us so we do not trouble you further.

My baby is no longer with me ☐

The baby's regular carers should fill in this questionnaire. Generally this will be the baby's mother and father, but there may be others who look after the baby such as the baby's grandparents or childminders and it is fine to ask them to help answer the questions.

If you would prefer the research assistant to fill in the form for you, just let us know. We can ask you the questions over the phone, or arrange a home visit.

How to fill in the questionnaire

1. Some questions on the following pages can be answered simply by putting a tick in the box next to the answer that applies to you.

   Example *Yes* ☑ *No* ☐

2. Some questions on the following pages can be answered by circling the response that applies to you.

   Example *Not at all---Occasionally---Frequently*

   If you really feel that you are in-between two of the descriptions, you can indicate this by circling the dotted line.

3. Usually after answering each question you go on to the next one unless a box you have ticked has an arrow next to it with an instruction to go to another question.

   Example *Yes* ☑ → Go to Question 5

   No ☐

WEIGHTS

Please fill in below all your baby's weights written in your Personal Child Health Record since birth. The weight recording page is normally near the end of your record.

We particularly need a weight of when your baby is at least six weeks old. If your baby hasn't been weighed since the age of six weeks, you could either make a special visit to the clinic, or else your baby is due an important doctor's check at eight weeks: the weight from this will be fine. Please write it in below before returning the questionnaire.

Date	Weight (kg)	Weight (lb/oz)

Date	Weight (kg)	Weight (lb/oz)

Section A: Milk feeding

1. How is your baby being fed at the moment? **(tick one only)**
   Breast feeding ☐ *Bottle feeding* ☐ *Both* ☐

2. At the moment, is your baby being fed on demand or generally at set times? **(tick one only)**
   On demand ☐ *It depends* ☐ *Generally set times* ☐

3. How often does your baby have milk feeds each day now? ..... *times per day*

4. (a) Does your baby ever have baby milk in a bottle at present (apart from expressed breast milk)?
   Not at all---Occasionally---Frequently

   (b) If so, how many ounces does your baby usually take? **(tick one only)**
   Less than 2oz (60ml) ☐ *2-4oz (60-120ml)* ☐ *4-6oz (120-180ml)* ☐ *More than 6oz (180ml)* ☐

5. At present, how long does it take to feed your baby?
   Less than 5 mins ☐ *5-15 mins* ☐ *15-25 mins* ☐ *25-35 mins* ☐ *More than 35 mins* ☐

If you have stopped breast feeding since completing the last questionnaire, please answ questions 6 to 8. If not, please go to question 9.
Never breastfed ☐→ please go to Question 9
Still breastfeeding ☐→ please go to Question 9

Only answer the questions in this box if you have stopped breast feeding since filling in the last questionnaire

6. How old was your baby when you last breast fed him/her? **(tick one only)**
   Less than one week ☐ *1-2 weeks* ☐ *3-4 weeks* ☐ *5-6 weeks* ☐ *7-8 weeks* ☐

7. Would you have liked to continue breast feeding for longer?
   Yes---Possibly---No

8. Please circle the most appropriate response to each statement about your reasons for stopping breast feeding.

   I stopped breast feeding because:

   (a) My baby was not gaining weight
   Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

   (b) My baby seemed hungry
   Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

   (c) I wasn't producing enough milk
   Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

   (d) I had cracked/sore nipples/mastitis etc.
   Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

   (e) I thought it was the right time
   Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

   (f) My family and friends told me to
   Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

   (g) Other **(please write the reasons)**

Section B: Weaning

9. Since the last questionnaire, your baby might have been given baby foods such as cereal, rusks or any other kind of solid food, including home made foods. Please indicate whether your baby has had solids.
   No, solids not given yet ☐→ please go to Section C, Page 4
   Yes, solids given ☐→ please answer questions below

If Yes, have you remembered to fill in the weaning diary in your Personal Child Health Record?

Only answer the questions in this box if your baby has been given solids

10. How old was your baby the very first time solid food of any kind was offered? ..... *weeks old*

11. (a) Since then has your baby had solid foods **(tick one only)**
    Not at all ☐ *Occasionally* ☐ *Regularly* ☐

    (b) If regularly, when did your baby first take solids everyday? ..... *weeks old*

12. Was your baby given any of these foods yesterday? **(tick all that apply)**

	Not at all	*Once*	*More than once*
(a) Home made weaning foods	☐	☐	☐
(b) Tinned/jarred weaned foods	☐	☐	☐
(c) Dried weaning foods	☐	☐	☐

13. How many times per day does your baby have solid foods at present? ..... *times*

14. I thought by baby started weaning
    Too early---At just the right time---Too late

15. Was there any attempt to delay giving your baby solids?
    No---By a few days---A week or two---Two to four weeks---Over four weeks

16. Please circle the most appropriate response to each statement about the reasons for starting weaning.

    I started solid food because:

    (a) My health visitor or doctor advised me to
    Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

    (b) A book or leaflet suggested I should
    Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

    (c) My family and friends told me to
    Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

    (d) I thought it was the right time
    Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

    (e) My baby seemed hungry
    Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

Section C: General feeding questions

17. At present, how well does your baby suck?
Strong---Average---Weak

18. At present, how is your baby's appetite?
Very good---Good---All right---Very poor---Poor

19. Overall, is your baby feeding enough?
Yes---Not always---No

20. At present, are feeding times for you usually:
Very relaxed---Relaxed---All right---Stressful---Very stressful

21. At present, are feeding times for your baby usually:
Very relaxed---Relaxed---All right---Stressful---Very stressful---Can't tell

22. At present, is your baby easy to feed?
Very easy---Easy---All right---Difficult---Very difficult

23. At present, does your baby have any trouble with the following:
(a) Sucking *Not at all---Occasionally---Frequently*
(b) Swallowing *Not at all---Occasionally---Frequently*
(c) Choking *Not at all---Occasionally---Frequently*

24. At present, do any of the following describe your baby?
(a) Has to be woken up for feeds *Not at all---Occasionally---Frequently*
(b) Sleeps during feeds *Not at all---Occasionally---Frequently*
(c) Cries during feeds *Not at all---Occasionally---Frequently*
(d) Slow feeder *Not at all---Occasionally---Frequently*
(e) Not satisfied *Not at all---Occasionally---Frequently*

25. At present, does your baby posset (bring up small amounts of feed)?
Rarely---Sometimes---Often

26. At present, does your baby vomit (bring up most or all of feed)?
Rarely---Sometimes---Often

27. At present, is your baby
Very thin---Thin---Average---Chubby---Fat

Is there anything else you would like to say about feeding your baby?
If so, **please give details below:**

Section D: Baby's illnesses

28. Has your baby seen the doctor due to illness, either at home or at the surgery? **(tick one only)**
No ☐ *Once* ☐ *More than once* ☐

29. Since birth has your baby had any of the following? **(tick all that apply)**

	No did not have	*Yes but did not see doctor*	*Yes and saw/spoke to a doctor*
(a) Diarrhoea and vomiting	☐	☐	☐
(b) Cough/cold	☐	☐	☐
(c) Ear ache/infection/discharge	☐	☐	☐
(d) Rash	☐	☐	☐
(e) Chest infection/difficulty breathing	☐	☐	☐
(f) An accident	☐	☐	☐

30. Has your baby ever been admitted to hospital? **(tick one only)**
No ☐ *Once* ☐ *More than once* ☐

31. Please describe for each admission

Age of baby (weeks)	*Reason for admission*	*Number of nights in hospital*

Section E: Baby's behaviour

INSTRUCTIONS: Please read carefully before starting:

This section asks lots of questions about how your baby behaves most of the time. Don't think too hard about the answer - tick the response that seems most true for your baby. As you read each description of the baby's behaviour below, please indicate how often your baby did this during the **LAST WEEK** (the past seven days) by circling one of the numbers as illustrated below.

1 = never	2 = very rarely	3 = less than half of the time	4 = about half the time	5 = more than half the time	6 = almost always	7 = always

If you have not seen your baby in the situation described during the last week, move to next section.

If you have seen your baby in this situation during the last week, but your baby never engaged in the behaviour listed, circle 1 = "Never".

Please be sure to circle a number for every item where you saw your baby in the situation.

FEEDING

During feeding, how often did baby:

32. *lie or sit quietly?* 1 2 3 4 5 6 7
33. *squirm or kick?* 1 2 3 4 5 6 7
34. *wave arms?* 1 2 3 4 5 6 7
35. *fuss or cry when s/he had enough to eat?* 1 2 3 4 5 6 7

Did your baby have to wait for food or liquids during the last week?

No ☐→ go to question 39
Yes ☐

If yes, how often did baby:

36. *seem not bothered?* 1 2 3 4 5 6 7
37. *show mild fussing?* 1 2 3 4 5 6 7
38. *cry loudly?* 1 2 3 4 5 6 7

SLEEPING

How often did baby:

39. *seem angry (crying and fussing) when you left her/him in the cot?* 1 2 3 4 5 6 7
40. *seem contented when left in the cot?* 1 2 3 4 5 6 7
41. *cry or fuss before going to sleep for naps?* 1 2 3 4 5 6 7

Before falling asleep at night during the last week, how often did baby:

42. *show no fussing or crying?* 1 2 3 4 5 6 7

During sleep, how often did baby:

43. *toss about in the cot?* 1 2 3 4 5 6 7
44. *move from the middle to the end of the cot?* 1 2 3 4 5 6 7
45. *sleep in one position only?* 1 2 3 4 5 6 7

After sleeping, how often did baby:

46. *fuss or cry immediately?* 1 2 3 4 5 6 7
47. *play quietly in cot?* 1 2 3 4 5 6 7
48. *coo for periods of 5 minutes or longer?* 1 2 3 4 5 6 7
49. *cry if someone doesn't come within a few minutes?* 1 2 3 4 5 6 7

BATHING AND DRESSING

When being dressed or undressed during the last week, how often did baby:

50. *wave his/her arms or kick?* 1 2 3 4 5 6 7
51. *squirm and/or try to roll away?* 1 2 3 4 5 6 7
52. *smile or laugh?* 1 2 3 4 5 6 7

1 = never	2 = very rarely	3 = less than half of the time	4 = about half the time	5 = more than half the time	6 = almost always	7 = always

Has your baby been put into bath water in the last week? *No* ☐→ go to question 56
Yes ☐

If yes, how often did baby:

53. *smile?* 1 2 3 4 5 6 7
54. *splash or kick?* 1 2 3 4 5 6 7
55. *turn body and/or squirm?* 1 2 3 4 5 6 7

When face was washed, how often did baby:

56. *smile or laugh?* 1 2 3 4 5 6 7
57. *fuss or cry?* 1 2 3 4 5 6 7

Has your baby's hair been washed in the last week? *No* ☐→ go to question 60
Yes ☐

If yes, how often did baby:

58. *smile or laugh?* 1 2 3 4 5 6 7
59. *fuss or cry?* 1 2 3 4 5 6 7

DAILY ACTIVITIES

How often during the last week did baby:

60. *cry or show distress at a loud sound (blender, vacuum cleaner, etc)?* 1 2 3 4 5 6 7 Not seen
61. *cry or show distress at a change in parents' appearance (glasses off, shower cap on, etc.)?* 1 2 3 4 5 6 7 Not seen
62. *cry after startling?* 1 2 3 4 5 6 7 Not seen

When being held, how often did baby:

63. *squirm, pull away or kick?* 1 2 3 4 5 6 7

When placed on his/her back, how often did baby:

64. *fuss or protest?* 1 2 3 4 5 6 7
65. *smile or laugh?* 1 2 3 4 5 6 7
66. *lie quietly?* 1 2 3 4 5 6 7
67. *wave arms or kick?* 1 2 3 4 5 6 7
68. *squirm and/or turn body?* 1 2 3 4 5 6 7

Has baby been placed in an infant seat or seat? *No* ☐→ go to question 73
Yes ☐

If yes, how often did baby:

69. *wave arms or kick?* 1 2 3 4 5
70. *squirm and turn body?* 1 2 3 4 5
71. *lie or sit quietly?* 1 2 3 4 5
72. *show distress at first, then quiet down?* 1 2 3 4 5

73. Have you returned from being away and ba was awake? *No* ☐→ go to question 74
Yes ☐

If yes, how often did baby:

smile or laugh? 1 2 3 4 5

Have any of the following soothing techniques been tried on baby in the last weeks? If so, how often did the method succeed in soothing baby?

74. *rocking* 1 2 3 4 5 Did not try
75. *holding* 1 2 3 4 5 Did not try
76. *singing or talking* 1 2 3 4 5 Did not try
77. *walking with baby* 1 2 3 4 5 Did not try
78. *giving baby a toy* 1 2 3 4 5 Did not try
79. *showing baby something to look at* 1 2 3 4 5 Did not try
80. *patting or gently rubbing some part of baby's body* 1 2 3 4 5 Did not try
81. *offering a feed* 1 2 3 4 5 Did not try
82. *changing baby's position* 1 2 3 4 5 Did not try
83. *other (please specify)* 1 2 3 4 5

Was there anything you intended to go back to and complete? Please check.

Please make sure you have filled in the weights on the first page.

Who completed this questionnaire? **(tick all that apply)**

Baby's mother ☐ *Baby's father* ☐ *Baby's grandparent* ☐

Nanny ☐ *Childminder* ☐ *Nursery* ☐

Other **(please tick and specify)** ☐ :

How old is your baby now? ..... *weeks and* ..... *days*

Please write the date you complete this ___ / ___ / ___

It would help us in our record keeping if you write your name here .....

If the name or address on the envelope was not correct or incomplete, or if you expect to move house in the near future and know your new address, it would help us if you could write it below:

.....

.....

.....

.....

When you have finished please return the questionnaire in the enclosed envelope even if you were not able to answer all of it.

Thank you very much for your help.

We will be back in touch with you when your baby is ~~three~~ four months old.

Dr. Kathryn Parkinson
Community Child Health
University of Newcastle
13 Walker Terrace
Gateshead
Tyne & Wear
NE8 1EB

Tel: Tyneside (0191) 4776000

Figure A.4. ***4 Month Questionnaire***

Feeding and growth study: Your baby at four months

ID No:

This questionnaire asks about you and your baby. If, for any reason, your baby is no longer with you, please tick the box below and return the questionnaire to us so we do not trouble you further. My baby is no longer with me ☐

The baby's regular carers should fill in this questionnaire. Generally this will be the baby's mother and father, but there may be others who look after the baby such as the baby's grandparents or childminders and it is fine to ask them to help answer the questions.

Any information you give us will be helpful. It will be treated in complete confidence, stored securely, and there will be nothing to identify you on this questionnaire unless you choose to put your name on it. However, do not feel you have to answer any questions you are uncomfortable with.

If you would prefer the research assistant to ask you the questions and fill in the form for you, just let us know. We can ask you the questions over the phone, or arrange a home visit.

How to fill in the questionnaire

1. Some questions on the following pages can be answered simply by putting a tick in the box next to the answer that applies to you.

   Example *Yes* ✓ *No* ☐

2. Some questions on the following pages can be answered by circling the response that applies to you.

   Example *Not at all* *(Occasionally)* *Frequently*

   If you really feel that you are in-between two of the descriptions, you can indicate this by circling the dotted line.

3. Usually after answering each question you go on to the next one unless a box you have ticked has an arrow next to it with an instruction to go to another question.

   Example *Yes* ✓ → Go to Question 5

   No ☐

WEIGHTS

Please fill in below all your baby's weights written in your Personal Child Health Record since filling in the last questionnaire. The weight recording page is normally near the end of your record.

Date	Weight (kg)	Weight (lb/oz)

Date	Weight (kg)	Weight (lb/oz)

Section A: Milk feeding

1. How is your baby being fed at the moment? **(tick one only)**
   Breast feeding ☐ *Bottle feeding* ☐ *Both* ☐

2. At the moment, is your baby being fed on demand or generally at set feeding times? **(tick one only)**
   On demand ☐ *It depends* ☐ *Generally set times* ☐

3. How often does your baby have milk feeds each day now?
   (please write in number of times) ..... *times per day*

4. (a) Does your baby ever have baby milk in a bottle at present (apart from expressed breast milk)?
   Not at all---Occasionally---Frequently

   (b) If so, how many ounces does your baby usually take? **(tick one only)**
   Less than 2oz (60ml) ☐ *2-4oz (60-120ml)* ☐ *4-6oz (120-180ml)* ☐
   6-8oz (180-240ml) ☐ *More than 8oz (240ml)* ☐

If you have stopped breast feeding since completing the last questionnaire, please answer questions 5 to 7. If not, please go to question 8.

Never breast fed ☐ → please go to Question 8
Stopped breast feeding since completing last questionnaire ☐ → please go to Questions 5-7
Still breast feeding ☐ → please go to Question 8

Only answer the questions in this box if you have stopped breast feeding since filling in the last questionnaire

5. How old was your baby when you last breast fed him/her? *7-8 weeks* ☐ *9-10 weeks* ☐
   11-12 weeks ☐ *13-14 weeks* ☐ *15-16 weeks* ☐ *17-18 weeks* ☐

6. Would you have liked to continue breast feeding for longer?
   Yes---Possibly---No

7. Please circle the most appropriate response to each statement about your reasons for stopping breast feeding.

   I stopped breast feeding because:

   (a) My baby was not gaining weight
   Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

   (b) My baby seemed hungry
   Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

   (c) I wasn't producing enough milk
   Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

   (d) I had cracked/sore nipples/mastitis etc.
   Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

   (e) I thought it was the right time
   Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

   (f) I was returning to work
   Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

   (g) My family and friends told me to
   Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

   (h) Other **(please write the reasons)**

Section B: Weaning

8. Since the last questionnaire, has your baby been given solid food of any kind?

   No, solids not given yet ☐ → please go to Section C
   Yes, solids given ☐ → please go to questions 9 - 17 below

♦ **If yes, have you remembered to fill in the weaning diary in your Personal Child Health Record?** ♦

9. How old was your baby the very first time solid food of any kind was offered? ..... *weeks old*

10. (a) Since then has your baby had solid foods **(tick one box)**
    Not at all ☐ *Occasionally* ☐ *Regularly* ☐

    (b) If regularly, when did your baby first take solids everyday? ..... *weeks old*

11. Was your baby given any of these foods yesterday? **(tick all that apply)**

	Not at all	*Once*	*More than once*
(a) Home made weaning foods	☐	☐	☐
(b) Tinned/jarred weaning foods	☐	☐	☐
(c) Dried weaning foods	☐	☐	☐

12. How many times per day does your baby have solid foods at present? ..... *times*

13. I thought by baby started weaning
    Too early---At just the right time---Too late

14. Was there any attempt to delay giving your baby solids?
    No---By a few days---A week or two---Two to four weeks---Over four weeks

15. Please circle the most appropriate response to each statement about the reasons for starting weaning.

    I started solid food because:

    (a) My health visitor or doctor advised me to
    Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

    (b) A book or leaflet suggested I should
    Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

    (c) My family and friends told me to
    Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

    (d) I thought it was the right time
    Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

    (e) My baby seemed hungry
    Strongly agree---Agree---Uncertain---Disagree---Strongly disagree

16. Has it been easy to wean your baby onto solid food?
    Very easy---Easy---All right---Difficult---Very difficult

17. Please circle the appropriate response to each statement about weaning your baby.

    (a) Will not take solids *Rarely---Sometimes---Often*
    (b) Will only take certain solids *Rarely---Sometimes---Often*
    (c) Uninterested in food *Rarely---Sometimes---Often*
    (d) Prefers drinks to food *Rarely---Sometimes---Often*
    (e) Cries during feeds *Rarely---Sometimes---Often*

Section C: General feeding questions

18. At present, how long does it take to feed your baby?
Less than 5 mins ☐ *5-15 mins* ☐ *15-25 mins* ☐ *25-35 mins* ☐ *More than 35 mins* ☐

19. At present, how well does your baby suck?
Strong---Average---Weak

20. At present, how is your baby's appetite?
Very good---Good---All right---Very poor---Poor

21. Overall, is your baby feeding enough?
Yes--Not always---No

22. At present, are feeding times for you usually:
Very relaxed---Relaxed---All right---Stressful---Very stressful

23. At present, are feeding times for your baby usually:
Very relaxed---Relaxed---All right---Stressful---Very stressful---Can't tell

24. At present, is your baby easy to feed?
Very easy---Easy---All right---Difficult---Very difficult

25. At present, does your baby have any trouble with any of the following:
(a) Sucking *Not at all---Occasionally---Frequently*
(b) Swallowing *Not at all---Occasionally---Frequently*
(c) Choking *Not at all---Occasionally---Frequently*

26. At present, do any of the following describe your baby?
(a) Has to be woken up for feeds *Not at all---Occasionally---Frequently*
(b) Sleeps during feeds *Not at all---Occasionally---Frequently*
(c) Cries during feeds *Not at all---Occasionally---Frequently*
(d) Slow feeder *Not at all---Occasionally---Frequently*
(e) Not satisfied *Not at all---Occasionally---Frequently*

27. At present, does your baby posset (bring up small amounts of feed)?
Rarely---Sometimes---Often

28. At present, does your baby vomit (bring up most or all of feed)?
Rarely---Sometimes---Often

29. At present, is your baby
Very thin---Thin---Average---Chubby---Fat

Is there anything else you would like to say about feeding your baby?
If so, **please give details below**

Section D: Accidents

30. Has your baby ever had a serious fall? **(tick one only)**
No never ☐ → Please go to Section E
Once ☐ → Please go to Question 31
More than once ☐ → Please go to Question 31

Can you tell us more about that fall? If your baby has had more than one fall, tell us about the most serious one.

31. How old was your baby when the accident happened? **(tick one only)**
Less than 1 month old ☐ *1-2 months* ☐ *2-3 months* ☐
3-4 months ☐ *More than 4 months* ☐

32. Where did your baby fall from? **(tick one only)**
A bed or sofa ☐ *A table or worktop* ☐ *Someone's arms* ☐
Something else **(please tick box and specify)** ☐:

33. How far did your baby fall? **(tick one only)**
Less than 1 foot (30 cm) ☐ *1-2 feet (½ metre)* ☐ *2-3 feet (1 metre)* ☐
More than 3 feet (1 metre) **(please tick box and specify)** ☐:

34. What sort of surface did they fall onto? **(tick one only)**
Padded ☐
Soft (e.g. carpet, grass) ☐
Firm (e.g. wood, vinyl, carpet tiles) ☐
Hard (e.g. concrete, asphalt) ☐

35. Can you tell us in your own words how the fall happened?

36. Was your baby injured at all? **(tick one only)**
No ☐ *Bruising* ☐ *Cut or graze* ☐ *Broken, bone fracture* ☐
Concussion, head injury ☐ *Other* **(please tick box and specify)** ☐:

37. Where was your baby's injury? **(tick one only)**
No injury ☐ *Head or neck* ☐ *Body* ☐ *Arms or legs* ☐

38. (a) Did your baby receive any medical help after the fall? **(tick one only)**
None ☐ *Telephone advice only* ☐ *Attended casualty* ☐
Saw GP ☐ *Admitted to hospital* ☐

38. (b) If admitted to hospital, how many nights did your baby spend there?
(please write in number of nights) ..... nights spent in hospital

Section E: Child's illnesses

39. Since the age of six weeks, has your baby seen the doctor due to illness, either at home or at the surgery? **(tick one only)**
No ☐ *Once* ☐ *More than once* ☐

40. Since the age of six weeks, has your baby had any of the following? **(tick all that apply)**

	No did not have	*Yes but did not see doctor*	*Yes and saw/spoke to a doctor*
(a) Diarrhoea and vomiting	☐	☐	☐
(b) Cough/cold	☐	☐	☐
(c) Ear ache/infection/discharge	☐	☐	☐
(d) Rash	☐	☐	☐
(e) Chest infection/difficulty breathing	☐	☐	☐
(f) Other **(please describe)**	☐	☐	☐

.....

41. Since filling in the questionnaire at six weeks, has your baby been admitted to hospital? **(tick one only)**
No ☐ *Once* ☐ *More than once* ☐

42. Please describe for each admission

Age of baby (weeks)	*Reason for admission*	*Number of nights in hospital*

Section F: Life events

We hope that you and your family's lives are going well at the moment and will continue to do so. However, all of us experience upsetting events occasionally. It will be helpful for us to know more about any such experiences that you may have had in the last 12 months in order to investigate what sorts of family circumstances may (or may not) influence a baby's feeding and development. Like all the information we collect, this will be kept entirely confidential. Please indicate below if you prefer not to complete this section.

I do not wish to complete this section ☐ → please go to last page

43. Has anyone close to you died within the last year? **(tick all that apply)**
Partner ☐ *Child* ☐ *Parent* ☐ *Other close family or friend* ☐

44. Have you experienced any of the following difficulties with relationships during the last year? **(tick all that apply)**
Arguments with partner ☐ *Unfaithfulness* ☐ *Separation / getting back together* ☐
Divorce ☐ *Domestic violence* ☐

45. Have any of the following people been seriously ill or injured during the last year? **(tick all that apply)**
Self ☐ *Partner* ☐ *Child* ☐ *Parent* ☐ *Other close family* ☐

46. Have you suffered any financial problems during the last year? **(tick all that apply)**
Loan called in / bailiffs ☐ *Decrease in income* ☐ *General money worries* ☐

47. Has either you or your partner suffered work problems during the last year? **(tick all that apply)**
You *Changed job* ☐ *Left job* ☐ *Lost job* ☐ *Job demotion / disciplined* ☐
Your partner *Changed job* ☐ *Left job* ☐ *Lost job* ☐ *Job demotion / disciplined* ☐

48. Have you been involved with the police in any of the following ways during the last year? **(tick all that apply)**
Arrested / convicted of crime ☐ *Victim of crime* ☐ *Victim of police brutality* ☐

49. Have you had either of the following difficulties with parenting during the last year? **(tick each if applicable)**
Becoming single parent ☐ *Disagreement over child custody / access / child support* ☐

50. Have you taken on responsibility for sick or elderly relatives within the last year?
Yes ☐ *No* ☐

51. Have you been involved in either of the following over the last year?
Car accident ☐ *Other major accident / disaster* ☐

52. Has any other unpleasant event happened in your life over the past year that you would like to tell us about? If so, **please write details below**

53. *None of the above events have happened to me in the last 12 months* ☐

Who completed this questionnaire? **(tick all that apply)**

Baby's mother ☐ *Baby's father* ☐ *Baby's grandparent* ☐

Nanny ☐ *Childminder* ☐ *Nursery* ☐

Other **(please tick and specify)** ☐ :

How old is your baby now? ..... *weeks and* ..... *days*

Please write the date you complete this ___ / ___ / ___

If the name or address on the envelope was not correct or incomplete, or if you expect to move house in the near future and know your new address, it would help us if you could write it below:

.....

.....

.....

Was there anything you intended to go back to and complete? Please check.

It would help us in our record keeping if you write your name here .....

Remember to fill in the weaning diary in your Personal Child Health Record when your baby starts having solids. If you have already filled in the weaning diary, please return the pink copy with this questionnaire.

Please check that you have filled in the table on page one with any weight records you have of your baby since filling in the last questionnaire.

When you have finished please return the questionnaire in the enclosed envelope even if you were not able to answer all of it. Please send the pink copy of the weaning diary from your Personal Child Health Record if you have already filled it in.

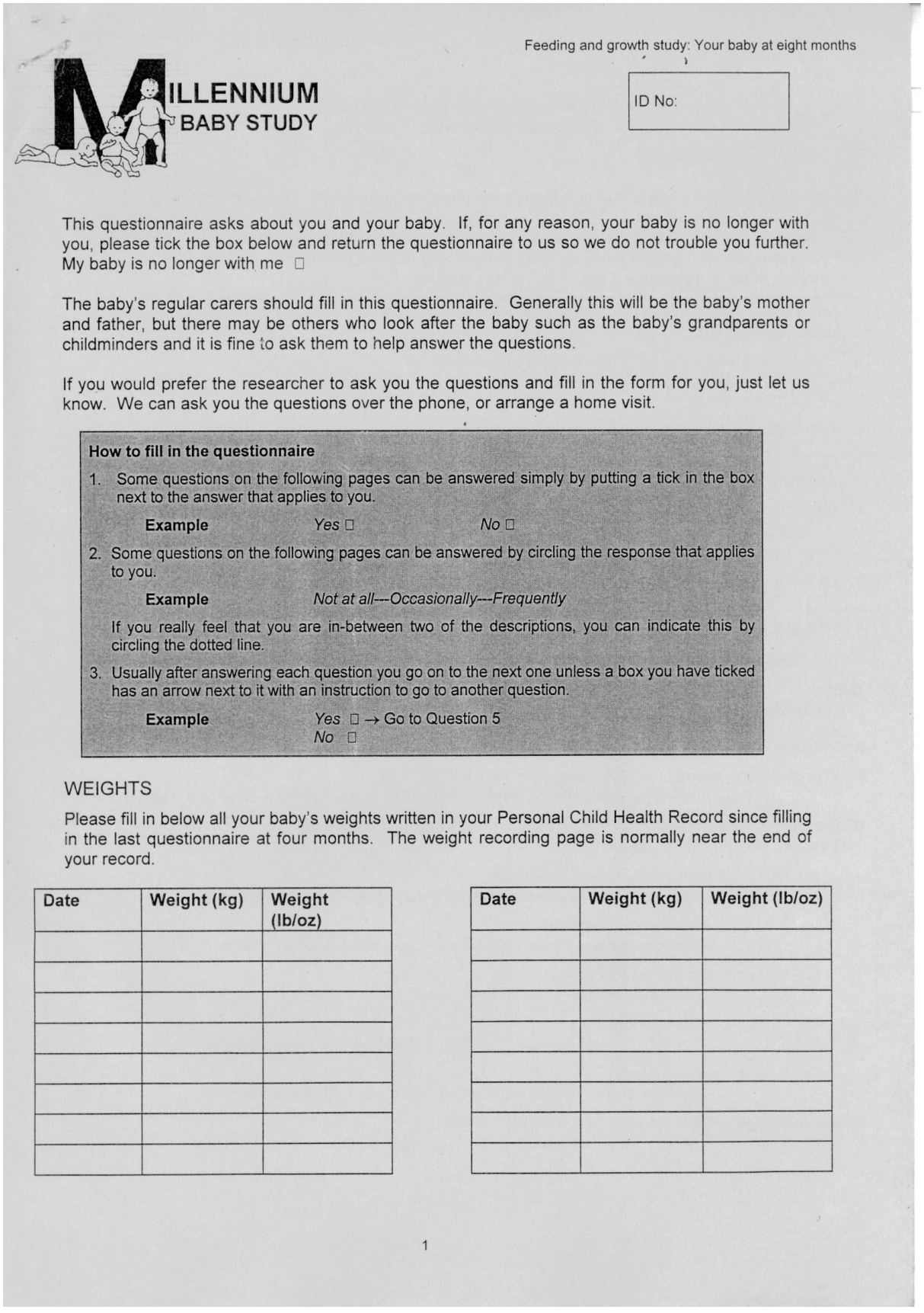
Thank you very much for your help.

We will be back in touch with you when your baby is eight months old.

Dr. Kathryn Parkinson, Community Child Health, University of Newcastle, 13 Walker Terrace, Gateshead, NE8 1EB

Tel: Tyneside (0191) 4776000

Figure A.5. ***8 Month Questionnaire***



MILLENNIUM BABY STUDY

Feeding and growth study: Your baby at eight months

ID No:

This questionnaire asks about you and your baby. If, for any reason, your baby is no longer with you, please tick the box below and return the questionnaire to us so we do not trouble you further.
My baby is no longer with me ☐

The baby's regular carers should fill in this questionnaire. Generally this will be the baby's mother and father, but there may be others who look after the baby such as the baby's grandparents or childminders and it is fine to ask them to help answer the questions.

If you would prefer the researcher to ask you the questions and fill in the form for you, just let us know. We can ask you the questions over the phone, or arrange a home visit.

How to fill in the questionnaire

1. Some questions on the following pages can be answered simply by putting a tick in the box next to the answer that applies to you.

   Example *Yes* ☐ *No* ☐

2. Some questions on the following pages can be answered by circling the response that applies to you.

   Example *Not at all---Occasionally---Frequently*

   If you really feel that you are in-between two of the descriptions, you can indicate this by circling the dotted line.

3. Usually after answering each question you go on to the next one unless a box you have ticked has an arrow next to it with an instruction to go to another question.

   Example *Yes* ☐ → Go to Question 5
   No ☐

WEIGHTS

Please fill in below all your baby's weights written in your Personal Child Health Record since filling in the last questionnaire at four months. The weight recording page is normally near the end of your record.

Date	Weight (kg)	Weight (lb/oz)

Date	Weight (kg)	Weight (lb/oz)

1

Section A: Milk feeding

1. Which milk is your baby being fed at the moment? **(tick all that apply)**
   Breast □ *Formula* □ *Cow's milk (doorstep)* □ *None* □
   Other □ ***please specify:***

2. How often does your baby have milk feeds each day now? ..... *times per day*

If you never breastfed or stopped before completing the last questionnaire, go to Section B. If you have stopped breast feeding since completing the last questionnaire (aged four months), answer questions 3 and 4 in the box below.

Answer the questions in this box if you have stopped breast feeding since completing the last questionnaire.

3. How old was your baby when you last breast fed him/her?
   Less than 17 weeks □ *17-20 weeks* □ *21-24 weeks* □
   25-28 weeks □ *29-32 weeks* □ *33-36 weeks* □
4. Would you have liked to continue breast feeding for longer? *Yes---Possibly---No*

Section B: Weaning

5. Has your baby started solids since receiving the last questionnaire (aged four months)?
   Yes, solids started since filling in last questionnaire □ → answer questions 6 - 11
   No, solids started before filling in last questionnaire □ → Section C

♦ **If yes, have you remembered to return the weaning diary in your Personal Child Health Record?** ♦

6. How old was your baby the very first time solid food of any kind was offered? ..... *weeks old*
7. (a) Since then has your baby had solid foods
   Not at all □ *Occasionally* □ *Regularly* □
   (b) If regularly, when did your baby first take solids everyday? ..... *weeks old*
8. I thought my baby started weaning *Too early---At just the right time---Too late*
9. Was there any attempt to delay giving your baby solids?
   No---By a few days---A week or two---Two to four weeks---Over four weeks
10. Please circle the most appropriate response to the following statements:
    I started solid food because:
    (a) My health visitor or doctor advised me to
    Strongly agree---Agree---Uncertain---Disagree---Strongly disagree
    (b) A book or leaflet suggested I should
    Strongly agree---Agree---Uncertain---Disagree---Strongly disagree
    (c) My family and friends told me to
    Strongly agree---Agree---Uncertain---Disagree---Strongly disagree
    (d) I thought it was the right time
    Strongly agree---Agree---Uncertain---Disagree---Strongly disagree
    (e) My baby seemed hungry
    Strongly agree---Agree---Uncertain---Disagree---Strongly disagree
11. Has it been easy to wean your baby onto solid food?
    Very easy---Easy---All right---Difficult---Very difficult

Section C: General feeding questions

12. Was your baby given any of these foods yesterday? **(tick all that apply)**

	Not at all	*Once*	*More than once*
(a) Home made weaning foods	☐	☐	☐
(b) Tinned/jarred weaning foods	☐	☐	☐
(c) Dried weaning foods	☐	☐	☐
(d) Family foods	☐	☐	☐

13. How many times per day does your baby have solid foods at present? ..... *times*

14. Which of the following statements describes your baby's feeding most accurately? **(tick one only)**
    (a) *Generally still needs to be fully fed* ☐
    (b) *Generally needs to be fed but eats food with fingers* ☐
    (c) *Generally eats with spoon but needs help* ☐
    (d) *Generally eats without help* ☐

♦ **If your baby has started to feed him or herself, have you remembered to fill in the finger feeding diary in your Personal Child Health Record?** ♦

15. When did your baby start reaching out for foods?
    Not yet ☐ *4-5 months* ☐ *5-6 months* ☐ *6-7 months* ☐ *7-8 months* ☐ *8-9 months* ☐

16. When was your baby first given finger foods (food your child can pick up and feed to themselves)?
    Not yet ☐ *4-5 months* ☐ *5-6 months* ☐ *6-7 months* ☐ *7-8 months* ☐ *8-9 months* ☐

17. How often does your baby eat finger foods?
    None ☐ *Once a day* ☐ *2-3 times a day* ☐ *4 or more times a day* ☐

18. At present, how does your baby drink?
    Mainly drinks from feeder cup ☐ *Mainly drinks from bottle* ☐ *Mainly drinks from breast* ☐

19. At present, how long does it take to give your baby a meal?
    Less than 5 mins ☐ *5-15 mins* ☐ *15-25 mins* ☐ *25-35 mins* ☐ *More than 35 mins* ☐

20. At present, do any of the following describe your baby? **(please circle most appropriate response to each question)**

(a) Hungry for foods	*Rarely---Sometimes---Often*
(b) Loves food	*Rarely---Sometimes---Often*
(c) Cannot fill him/her	*Rarely---Sometimes---Often*
(d) Will not take solids	*Rarely---Sometimes---Often*
(e) Eats a limited variety of food	*Rarely---Sometimes---Often*
(f) Uninterested in food	*Rarely---Sometimes---Often*
(g) Prefers drinks to food	*Rarely---Sometimes---Often*
(h) Slow feeder	*Rarely---Sometimes---Often*
(i) Cries during feeds	*Rarely---Sometimes---Often*

21. Some babies have difficulties being fed. Does your baby do any of the following when given food? **(please circle most appropriate response to each question)**

(a)	Pushes food/spoon away	*Rarely---Sometimes---Often*
(b)	Turns head away repeatedly	*Rarely---Sometimes---Often*
(c)	Closes mouth when offered food	*Rarely---Sometimes---Often*
(d)	Can't chew solid foods	*Rarely---Sometimes---Often*
(e)	Gags on food	*Rarely---Sometimes---Often*
(f)	Holds food in mouth	*Rarely---Sometimes---Often*
(g)	Spits food out	*Rarely---Sometimes---Often*
(h)	Throws food	*Rarely---Sometimes---Often*
(i)	Cries/screams during meals	*Rarely---Sometimes---Often*

22. If your baby does not finish a course, or part of a meal, what do you do?

(a)	Encourage him/her to eat	*Rarely---Sometimes---Often*
(b)	Make him/her eat the food	*Rarely---Sometimes---Often*
(c)	Offer something else	*Rarely---Sometimes---Often*

23. If your baby does not finish a course, or part of a meal, what do you do after the meal?

(a)	Offer the same food again later	*Rarely---Sometimes---Often*
(b)	Offer something else later	*Rarely---Sometimes---Often*
(c)	Offer nothing else until the next meal	*Rarely---Sometimes---Often*

24. At present, how is your baby's appetite?
Very good---Good---All right---Very poor---Poor

25. Overall, is your baby feeding enough?
Yes--Not always---No

26. At present, are feeding times for you usually:
Very relaxed---Relaxed---All right---Stressful---Very stressful

27. At present, are feeding times for your baby usually:
Very relaxed---Relaxed---All right---Stressful---Very stressful---Can't tell

28. At present, is your baby easy to feed?
Very easy---Easy---All right---Difficult---Very difficult

29. At present, does your baby vomit?
Rarely---Sometimes---Often

30. At present, is your baby
Very thin---Thin---Average---Chubby---Fat

31. Is there anything else you would like to say about feeding your baby?
If so, **please give details below**

Section D: Accidents

32. Has your baby had a serious fall since the last questionnaire? **(tick one only)**

No never	☐ → Please go to Section E
Once	☐ → Please go to Question 33
More than once	☐ → Please go to Question 33

Can you tell us more about that fall? If your baby has had more than one fall, tell us about the most serious one.

33. How old was your baby when the accident happened? **(tick one only)**
Less than 4 months old ☐ *4-5 months* ☐ *5-6 months* ☐ *6-7 months* ☐
7-8 months ☐ *More than 9 months* ☐

34. Where did your baby fall from? **(tick one only)**
A bed or sofa ☐ *A table or worktop* ☐ *Someone's arms* ☐
Something else **(please tick box and specify)** ☐:

35. How far did your baby fall? **(tick one only)**
Less than 1 foot (30 cm) ☐ *1-2 feet (½ metre)* ☐ *2-3 feet (1 metre)* ☐
More than 3 feet (1 metre) **(please tick box and specify)** ☐:

36. What sort of surface did they fall onto? **(tick one only)**

Padded	☐
Soft (e.g. carpet, grass)	☐
Firm (e.g. wood, vinyl, carpet tiles)	☐
Hard (e.g. concrete, asphalt)	☐

37. Can you tell us in your own words how the fall happened?

38. Was your baby injured at all? **(tick one only)**
No ☐ *Bruising* ☐ *Cut or graze* ☐ *Broken, bone fracture* ☐
Concussion, head injury ☐ *Other* **(please tick box and specify)** ☐:

39. Where was your baby's injury? **(tick one only)**
No injury ☐ *Head or neck* ☐ *Body* ☐ *Arms or legs* ☐

40. (a) Did your baby receive any medical help after the fall? **(tick one only)**
None ☐ *Telephone advice only* ☐ *Attended casualty* ☐
Saw GP ☐ *Admitted to hospital* ☐

(b) If admitted to hospital, how many nights did your baby spend there?
(please write in number of nights) ..... nights spent in hospital

Section E: Your baby's illnesses

41. Since the age of four months, has your baby seen the doctor due to illness, either at home or at the surgery? **(tick one only)**

No ☐ *Once* ☐ *More than once* ☐

42. Since the age of four months, has your baby had any of the following? **(tick all that apply)**

	No did not have	*Yes but did not see doctor*	*Yes and saw/spoke to a doctor*
(a) Diarrhoea and vomiting	☐	☐	☐
(b) Cough/cold	☐	☐	☐
(c) Ear ache/infection/discharge	☐	☐	☐
(d) Rash	☐	☐	☐
(e) Chest infection/difficulty breathing	☐	☐	☐
(f) Other **(please describe)**	☐	☐	☐

.....

43. Since filling in the questionnaire at four months, has your baby been admitted to hospital? **(tick one only)**

No ☐ *Once* ☐ *More than once* ☐

44. Please describe each admission

Age of baby (months)	*Reason for admission*	*Number of nights in hospital*

Section F: Baby's behaviour

INSTRUCTIONS: Please read carefully before starting:

This section asks lots of questions about how your baby behaves most of the time. Don't think too hard about the answer - tick the response that seems most true for your baby. As you read each description of the baby's behaviour below, please indicate how often your baby did this during the **LAST WEEK** (the past seven days) by circling one of the numbers as illustrated below.

1 = never	2 = very rarely	3 = less than half of the time	4 = about half the time	5 = more than half the time	6 = almost always	7 = always

If you have not seen your baby in the situation described during the last week, please circle NA (for Not Applicable) and move to next question.

If you have seen your baby in this situation during the last week, but your baby never engaged in the behaviour listed, circle 1 = "Never".

Please be sure to circle a number for every item where you saw your baby in the situation.

FEEDING

During feeding, how often did baby:

45.	*lie or sit quietly?*	1 2 3 4 5 6 7
46.	*squirm or kick?*	1 2 3 4 5 6 7
47.	*wave arms?*	1 2 3 4 5 6 7
48.	*fuss or cry when s/he had enough to eat?*	1 2 3 4 5 6 7
49.	*fuss or cry when given a disliked food*	1 2 3 4 5 6 7

If your baby had to wait for food or liquids during the last week, how often did baby:

50.	*seem not bothered?*	NA 1 2 3 4 5 6 7
51.	*show mild fussing?*	NA 1 2 3 4 5 6 7
52.	*cry loudly?*	NA 1 2 3 4 5 6 7

If your baby was given a new food or liquid, how often did the baby:

53.	*accept it immediately?*	NA 1 2 3 4 5 6 7
54.	*reject it by spitting out, closing mouth, etc.?*	NA 1 2 3 4 5 6 7
55.	*not accept it no matter how many times offered?*	NA 1 2 3 4 5 6 7

SLEEPING

How often did baby:

56.	*seem angry (crying and fussing) when you left her/him in the cot?*	1 2 3 4 5 6 7
57.	*seem contented when left in the cot?*	1 2 3 4 5 6 7
58.	*cry or fuss before going to sleep for naps?*	1 2 3 4 5 6 7

Before falling asleep at night during the last week, how often did baby:

59.	*show no fussing or crying*	1 2 3 4 5 6 7

1 = never	2 = very rarely	3 = less than half of the time	4 = about half the time	5 = more than half the time	6 = almost always	7 = always

During sleep, how often did baby:

60.	*toss about in the cot?*		1 2 3 4 5 6 7
61.	*move from the middle to the end of the cot?*		1 2 3 4 5 6 7
62.	*sleep in one position only?*		1 2 3 4 5 6 7

After sleeping, how often did baby:

63.	*fuss or cry immediately?*		1 2 3 4 5 6 7
64.	*play quietly in cot?*		1 2 3 4 5 6 7
65.	*coo and vocalise for periods of five minutes or longer?*		1 2 3 4 5 6 7
66.	*cry if someone didn't come within a few minutes?*		1 2 3 4 5 6 7

BATHING AND DRESSING

When being dressed or undressed during the last week, how often did baby:

67.	*wave his/her arms or kick?*		1 2 3 4 5 6 7
68.	*squirm and/or try to roll away?*		1 2 3 4 5 6 7
69.	*smile or laugh?*		1 2 3 4 5 6 7

If your baby has been put into bath water in the last week, how often did baby:

70.	*smile?*	NA	1 2 3 4 5 6 7
71.	*laugh?*	NA	1 2 3 4 5 6 7
72.	*splash or kick?*	NA	1 2 3 4 5 6 7
73.	*turn body and/or squirm*	NA	1 2 3 4 5 6 7

When face was washed, how often did baby?

74.	*smile or laugh?*		1 2 3 4 5 6 7
75.	*fuss or cry?*		1 2 3 4 5 6 7

If hair was washed, how often did baby?

76.	*smile or laugh?*	NA	1 2 3 4 5 6 7
77.	*fuss or cry?*	NA	1 2 3 4 5 6 7

DAILY ACTIVITIES

How often during the last week did baby:

78.	*cry or show distress at a loud sound (blender, vacuum cleaner, etc.)?*	NA	1 2 3 4 5 6 7
79.	*cry or show distress at a change in parents' appearance (glasses off, shower cap on, etc.)?*	NA	1 2 3 4 5 6 7
80.	*when in a position to see the television set, look at it for 2-5 minutes?*	NA	1 2 3 4 5 6 7
81.	*when in a position to see the television set, look at it for 5 minutes or longer?*	NA	1 2 3 4 5 6 7
82.	*protest at being put in a confining place (infant seat, play pen, car seat, etc)?*	NA	1 2 3 4 5 6 7
83.	*cry after startling?*	NA	1 2 3 4 5 6 7

When being held, how often did baby:

84.	*squirm, pull away or kick?*		1 2 3 4 5 6 7

1 = never	2 = very rarely	3 = less than half of the time	4 = about half the time	5 = more than half the time	6 = almost always	7 = always

If placed on his/her back, how often did baby:

85.	*fuss or protest?*	NA	1	2	3	4	5	6	7
86.	*smile or laugh?*	NA	1	2	3	4	5	6	7
87.	*lie quietly?*	NA	1	2	3	4	5	6	7
88.	*wave arms or kick?*	NA	1	2	3	4	5	6	7
89.	*squirm and/or turn body?*	NA	1	2	3	4	5	6	7

When baby wanted something, how often did baby:

90.	*become upset when baby could not get what was wanted?*	NA	1	2	3	4	5	6	7
91.	*have tantrums (crying, screaming, face red, etc.) when baby did not get was wanted?*	NA	1	2	3	4	5	6	7

If your baby has been placed in an infant seat or car seat, how often did baby:

92.	*wave arms or kick?*	NA	1	2	3	4	5	6	7
93.	*squirm and turn body?*	NA	1	2	3	4	5	6	7
94.	*lie or sit quietly?*	NA	1	2	3	4	5	6	7
95.	*show distress at first, then quiet down?*	NA	1	2	3	4	5	6	7

If you returned from being away and baby was awake, how often did baby:

96.	*smile or laugh?*	NA	1	2	3	4	5	6	7

If introduced to a strange person, how often did baby:

97.	*cling to a parent?*	NA	1	2	3	4	5	6	7
98.	*refuse to go to the stranger?*	NA	1	2	3	4	5	6	7
99.	*hang back from the stranger?*	NA	1	2	3	4	5	6	7
100.	*never "warm up" to the stranger?*	NA	1	2	3	4	5	6	7
101.	*approach the stranger at once?*	NA	1	2	3	4	5	6	7
102.	*smile or laugh?*	NA	1	2	3	4	5	6	7

If introduced to a dog or cat, how often did baby:

103.	*cry or show distress?*	NA	1	2	3	4	5	6	7
104.	*smile or laugh?*	NA	1	2	3	4	5	6	7
105.	*approach at once?*	NA	1	2	3	4	5	6	7

PLAY

How often during the last week did baby:

106.	*look at pictures in books and/or magazines for 2-5 minutes or longer at a time?*	NA	1	2	3	4	5	6	7
107.	*look at pictures in books and/or magazines for 5 minutes or longer at a times?*	NA	1	2	3	4	5	6	7
108.	*stare at a mobile, cot bumper, or picture for 5 minutes or longer at a time?*	NA	1	2	3	4	5	6	7
109.	*play with one toy or object for 5-10 minutes or longer?*	NA	1	2	3	4	5	6	7
110.	*play with one toy or object for 10 minutes or longer?*	NA	1	2	3	4	5	6	7
111.	*spend time just looking at playthings?*	NA	1	2	3	4	5	6	7
112.	*repeat the same sounds over and over again?*	NA	1	2	3	4	5	6	7
113.	*laugh aloud in play?*	NA	1	2	3	4	5	6	7
114.	*smile or laugh when tickled?*	NA	1	2	3	4	5	6	7
115.	*cry or show distress when tickled?*	NA	1	2	3	4	5	6	7
116.	*repeat the same movement with an object for 2 minutes or longer (e.g. putting a block in a cup, kicking or hitting a mobile)?*	NA	1	2	3	4	5	6	7

1 = never	2 = very rarely	3 = less than half of the time	4 = about half the time	5 = more than half the time	6 = almost always	7 = always

When something baby was playing with had to be removed, how often did baby:

117.	*cry or show distress for a short time?*	NA	1 2 3 4 5 6 7
118.	*cry or show distress for several minutes or longer?*	NA	1 2 3 4 5 6 7
119.	*seem not bothered?*	NA	1 2 3 4 5 6 7

When tossed around playfully, how often did the baby:

120.	*smile?*	NA	1 2 3 4 5 6 7
121.	*laugh?*	NA	1 2 3 4 5 6 7

During a peekaboo game, how often did baby:

122.	*smile?*	NA	1 2 3 4 5 6 7
123.	*laugh?*	NA	1 2 3 4 5 6 7

SOOTHING TECHNIQUES

Have any of the following soothing techniques been tried on baby in the last 2 weeks? If so, how often did the method succeed in soothing baby?

124.	*rocking?*	Did not try	1 2 3 4 5 6 7
125.	*holding?*	Did not try	1 2 3 4 5 6 7
126.	*singing or talking?*	Did not try	1 2 3 4 5 6 7
127.	*walking with the baby?*	Did not try	1 2 3 4 5 6 7
128.	*giving the baby a toy?*	Did not try	1 2 3 4 5 6 7
129.	*showing the baby something to look at?*	Did not try	1 2 3 4 5 6 7
130.	*patting or gently rubbing some part of the baby's body?*	Did not try	1 2 3 4 5 6 7
131.	*offering food or liquid*	Did not try	1 2 3 4 5 6 7
132.	*offering baby his/her security object*	Did not try	1 2 3 4 5 6 7
133.	*changing baby's position*	Did not try	1 2 3 4 5 6 7
134.	*other (please specify)*	Did not try	1 2 3 4 5 6 7

Who completed this questionnaire? **(tick all that apply)**

Baby's mother ☐ *Baby's father* ☐ *Baby's grandparent* ☐ *Nanny* ☐
Childminder ☐ *Nursery* ☐ *Other* **(please tick and specify)** ☐ :

How old is your baby now? ..... *months and* ..... *weeks*

Is the baby's father in paid employment at the moment?

Yes ..... hours per week
No ☐
Student ..... hours per week
Not living with family ☐

Is the baby's mother in paid employment at the moment?

Yes ..... hours per week
No ☐
Student ..... hours per week
Not living with family ☐

Who provides child care when the baby's mother is working outside the home?

Not applicable ☐

	1	2	3	4
Baby's father	*Most of the time*	*Some of the time*	*Occasionally*	*Never*
Baby's grandparent	*Most of the time*	*Some of the time*	*Occasionally*	*Never*
Nanny	*Most of the time*	*Some of the time*	*Occasionally*	*Never*
Childminder	*Most of the time*	*Some of the time*	*Occasionally*	*Never*
Nursery	*Most of the time*	*Some of the time*	*Occasionally*	*Never*

Other **(please tick and specify)** ☐ :

Please write the date you complete this ___ / ___ / ___

If the name or address on the envelope was not correct or incomplete, or if you expect to move house in the near future and know your new address, it would help us if you could write it below:

.....

.....

.....

Telephone No. .....

It would help us in our record keeping if you write your name here .....

Was there anything you intended to go back to and complete? Please check.

Please check the following pages in your Personal Child Health Record:

1. If you have not returned the top pink copy of the weaning diary yet, include it when you return this questionnaire even if you have not filled it in.
2. If you have filled in the finger feeding diary, return the top pink copy with this questionnaire.
3. If you have not filled in the finger feeding diary so far, remember to fill it in when your baby starts feeding his/herself.

Please check that you have filled in the table on page one with any weight records you have of your baby since filling in the last questionnaire.

When you have finished, return the questionnaire in the enclosed envelope even if you were not able to answer all of it.

Thank you very much for your help.

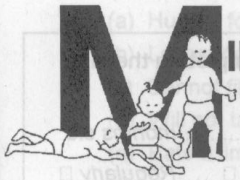
We will be back in touch with you when your baby is twelve months old.

Dr. Kathryn Parkinson, Community Child Health, University of Newcastle upon Tyne, 13 Walker Terrace, Gateshead, NE8 1EB.

Tel. (0191) 4776000

Figure A.6. ***12 Month Questionnaire***

Feeding and growth study: Your baby at twelve months



MILLENNIUM BABY STUDY

ID No:

CONGRATULATIONS ON YOUR CHILD'S FIRST BIRTHDAY! Now your baby is one year we are coming to the end of this part of the study.

Please answer as much of this questionnaire as you feel able to. Any information you give us will be helpful. It will be treated in complete confidence, stored securely, and there will be nothing to identify you on this questionnaire unless you choose to put your name on it.

As before, this questionnaire asks about you and your baby. If, for any reason, your baby is no longer with you, please tick the box below and return the questionnaire to us so we do not trouble you further.

My baby is no longer with me ☐

Weights

Please fill in below all your baby's weights written in your Personal Child Health Record since filling in the last questionnaire at eight months. The weight recording page is normally near the end of your record.

Date	Weight (kg)	Weight (lb/oz)

Date	Weight (kg)	Weight (lb/oz)

Section A: Milk feeding

1. How often does your baby drink milk each day now? ..... *times per day*

2. Which milk is your baby being fed at the moment? **(tick all that apply)**
   Breast ☐ *Formula* ☐ *Cow's milk (doorstep)* ☐ *None* ☐
   Other ☐ ***please specify:***

Answer the questions in this box only if you have stopped breast feeding since completing the last questionnaire.

3. How old was your baby when you last breast fed him/her?
   Less than 8 months ☐ *8-9 months* ☐ *9-10 months* ☐ *10-11 months* ☐
   11-12 months ☐ *Over 12 months* ☐

4. Would you have liked to continue breast feeding for longer? *Yes---Possibly---No*

Section B: Weaning

Answer the questions in this box only if your baby has started solids since completing the last questionnaire (aged eight months)

5. How old was your baby the very first time solid food of any kind was offered? ..... *months old*
6. (a) Since then has your baby had solid foods? *Not at all* ☐ *Occasionally* ☐ *Regularly* ☐
   (b) If regularly, when did your baby first take solids everyday? ..... *months old*
7. I thought my baby started weaning *Too early---At just the right time---Too late*
8. Has it been easy to wean your baby onto solid food?
   Very easy---Easy---All right---Difficult---Very difficult

Section C: General feeding questions

9. Was your baby given any of these foods yesterday? **(tick all that apply)**

	Not at all	*Once*	*More than once*
(a) Home made weaning foods	☐	☐	☐
(b) Tinned/jarred weaning foods	☐	☐	☐
(c) Dried weaning foods	☐	☐	☐
(d) Family foods	☐	☐	☐

10. How many times per day does your baby have solid foods at present? ..... *times*

11. Which of the following statements describes your baby's feeding most accurately? **(tick one only)**
    (a) *Generally still needs to be fully fed* ☐
    (b) *Generally needs to be fed but eats food with fingers* ☐
    (c) *Generally eats with spoon but needs help* ☐
    (d) *Generally eats without help* ☐

♦ **If your baby has started to feed him or herself, have you remembered to fill in the finger feeding diary in your Personal Child Health Record?** ♦

12. When did your baby start reaching out for foods ?
    Not yet ☐ *7-8 months* ☐ *8-9 months* ☐ *9-10 months* ☐
    10-11 months ☐ *11-12 months* ☐ *Over 12 months* ☐

13. When was your baby first given finger foods? (foods children can pick up and feed to themselves)?
    Not yet ☐ *7-8 months* ☐ *8-9 months* ☐ *9-10 months* ☐
    10-11 months ☐ *11-12 months* ☐ *Over 12 months* ☐

14. How often does your baby eat finger foods?
    None ☐ *Once a day* ☐ *2-3 times a day* ☐ *4 or more times a day* ☐

15. At present, how does your baby drink?
    Mainly drinks from feeder cup ☐ *Mainly drinks from bottle* ☐ *Mainly drinks from breast* ☐

16. At present, how long does it take to give your baby a meal?
    Less than 5 mins ☐ *5-15 mins* ☐ *15-25 mins* ☐ *25-35 mins* ☐ *More than 35 mins* ☐

17. At present, how much do the following describe your baby? **Please answer each item**
    - (a) Hungry for foods — *Rarely---Sometimes---Often*
    - (b) Loves food — *Rarely---Sometimes---Often*
    - (c) Cannot fill him/her — *Rarely---Sometimes---Often*
    - (d) Will not take solids — *Rarely---Sometimes---Often*
    - (e) Eats a limited variety of food — *Rarely---Sometimes---Often*
    - (f) Uninterested in food — *Rarely---Sometimes---Often*
    - (g) Prefers drinks to food — *Rarely---Sometimes---Often*
    - (h) Slow feeder — *Rarely---Sometimes---Often*
    - (i) Cries during feeds — *Rarely---Sometimes---Often*

18. How often does your baby do the following when given food? **Please answer each item**
    - (a) Pushes food/spoon away — *Rarely---Sometimes---Often*
    - (b) Turns head away repeatedly — *Rarely---Sometimes---Often*
    - (c) Closes mouth when offered food — *Rarely---Sometimes---Often*
    - (d) Can't chew solid foods — *Rarely---Sometimes---Often*
    - (e) Gags on food — *Rarely---Sometimes---Often*
    - (f) Holds food in mouth — *Rarely---Sometimes---Often*
    - (g) Spits food out — *Rarely---Sometimes---Often*
    - (h) Throws food — *Rarely---Sometimes---Often*
    - (i) Cries/screams during meals — *Rarely---Sometimes---Often*

19. If your baby does not finish a course, or part of a meal, what do you do?
    - (a) Encourage him/her to eat — *Rarely---Sometimes---Often*
    - (b) Make him/her eat the food — *Rarely---Sometimes---Often*
    - (c) Offer something else — *Rarely---Sometimes---Often*

20. If your baby does not finish a course, or part of a meal, what do you do after the meal?
    - (a) Offer the same food again later — *Rarely---Sometimes---Often*
    - (b) Offer something else later — *Rarely---Sometimes---Often*
    - (c) Offer nothing else until the next meal — *Rarely---Sometimes---Often*

21. At present, how is your baby's appetite? *Very good---Good---All right---Poor---Very poor*

22. Overall, is your baby feeding enough? *Yes--Not always---No*

23. At present, are feeding times for you usually:
    Very relaxed---Relaxed---All right---Stressful---Very stressful

24. At present, are feeding times for your baby usually:
    Very relaxed---Relaxed---All right---Stressful---Very stressful---Can't tell

25. At present, is your baby easy to feed? *Very easy---Easy---All right---Difficult---Very difficult*

26. At present, does your baby vomit? *Rarely---Sometimes---Often*

27. At present, is your baby *Very thin---Thin---Average---Chubby---Fat*

28. Is there anything else you would like to say about feeding your baby?

Section D: Your baby's behaviour

29. What is your baby doing for him/herself now? **Please answer each item**
    (a) Sitting without support *Yes---Nearly---No*
    (b) Pulling him/herself up to stand *Yes---Nearly---No*
    (c) Walking around furniture *Yes---Nearly---No*
    (d) Walking without support *Yes---Nearly---No*

30. Is your baby? **Please answer each item**
    (a) Making noises *Yes---No*
    (b) Babbling *Yes---No*
    (c) Making recognisable syllables (such as ma, ba, pa, cu) *Yes---No*
    (d) Saying words with meaning *Yes---No*

31. Do you have any problems with your baby's sleeping? **Please answer each item**
    (a) Won't go to sleep *No problem---Slight problem---Big problem*
    (b) Wakes during night *No problem---Slight problem---Big problem*
    (c) Won't sleep in own cot/bed *No problem---Slight problem---Big problem*

32. Does your baby go to sleep at the same time every night?
    Nearly always---Usually---Some of the time---Rarely---Never

33. In general, how often does your baby wake in the night?
    No □ *Once* □ *Twice* □ *Three times* □ *Four or more times* □

34. What do you usually do if your baby wakes? **Please answer each item**
    (a) Leave him/her alone *Nearly always---Usually---Rarely---Never*
    (b) Go and check but speak to him/her only *Nearly always---Usually---Rarely---Never*
    (c) Pick him/her up and comfort, then leave in cot *Nearly always---Usually---Rarely---Never*
    (d) Take him/her into bed with me *Nearly always---Usually---Rarely---Never*

35. Do you have any problems with your baby crying? **Please answer each item**
    (a) Cries too often *Nearly always---Usually---Rarely---Never*
    (b) Difficult to comfort *Nearly always---Usually---Rarely---Never*
    (c) Cries during night *Nearly always---Usually---Rarely---Never*
    (d) Cries when separated from mother *Nearly always---Usually---Rarely---Never*

36. On average, how often does your baby cry during the day
    Once or twice a day □ *3-4 times a day* □ *Five or more times a day* □

37. On average, how often does your baby pass a stool (poo)?
    More than once a day □ *Everyday* □ *Every two days* □
    Every three to four days □ *Every five to seven days* □ *Less than once a week* □

38. Does your baby have any problems passing stools (doing a poo)? **Please answer each item**
    (a) Difficulty passing stools *No problem---Slight problem---Big problem*
    (b) Hard stools *No problem---Slight problem---Big problem*
    (c) Pain passing stools *No problem---Slight problem---Big problem*
    (d) Rarely passes stools *No problem---Slight problem---Big problem*

Section E: Your baby's illnesses

39. Since the age of eight months, has your baby seen the doctor due to illness, either at home or at the surgery? **(tick one only)**

No ☐ *Once* ☐ *More than once* ☐

40. Since the age of eight months, has your baby had any of the following? **(tick all that apply)**

	No did not have	*Yes but did not see doctor*	*Yes and saw/spoke to a doctor*
(a) Diarrhoea and vomiting	☐	☐	☐
(b) Cough/cold	☐	☐	☐
(c) Ear ache/infection/discharge	☐	☐	☐
(d) Rash	☐	☐	☐
(e) Chest infection/difficulty breathing	☐	☐	☐

(f) Other **(please describe)** .....

.....

41. Since filling in the questionnaire at eight months, has your baby been admitted to hospital? **(tick one only)**

No ☐ *Once* ☐ *More than once* ☐

42. Please describe each admission

Age of baby (months)	*Reason for admission*	*Number of nights in hospital*

Section F: Accidents

43. Has your baby had a serious fall since the last questionnaire? **(tick one only)**

No never	☐ → Please go to Section G
Once	☐ → Please go to Question 44
More than once	☐ → Please go to Question 44

Can you tell us more about that fall? If your baby has had more than one fall, tell us about the most serious one.

44. How old was your baby when the accident happened? **(tick one only)**

Less than 8 months ☐ *8-9 months* ☐ *9-10 months* ☐ *10-11 months* ☐
11-12 months ☐ *More than 12 months* ☐

45. Where did your baby fall from? **(tick one only)**

A bed or sofa ☐ *A table or worktop* ☐ *Someone's arms* ☐
Something else **(please tick box and specify)** ☐:

46. How far did your baby fall? **(tick one only)**

Less than 1 foot (30 cm) ☐ *1-2 feet (½ metre)* ☐ *2-3 feet (1 metre)* ☐
More than 3 feet **(please tick box and specify)** ☐:

47. What sort of surface did your baby fall onto? **(tick one only)**

Padded	☐
Soft (e.g. carpet, grass)	☐
Firm (e.g. wood, vinyl, carpet tiles)	☐
Hard (e.g. concrete, asphalt)	☐

48. Can you tell us in your own words how the fall happened?

49. Was your baby injured at all? **(tick one only)**

No ☐ *Bruising* ☐ *Cut or graze* ☐ *Broken bone or fracture* ☐
Concussion, head injury ☐ *Other* **(please tick box and specify)** ☐:

50. Where was your baby's injury? **(tick one only)**

No injury ☐ *Head or neck* ☐ *Body* ☐ *Arms or legs* ☐

51. (a) Did your baby receive any medical help after the fall? **(tick one only)**

None ☐ *Telephone advice only* ☐ *Attended casualty* ☐ *Saw GP* ☐
Admitted to hospital ☐

(b) If admitted to hospital, how many nights did your baby spend there?
(please write in number of nights) ..... nights spent in hospital

Section G: Mother's eating patterns

Now we want to find out some more about you! For each statement, please circle the response from the options that best describes you.

1 never	2 seldom	3 sometimes	4 often	5 very often	6 not relevant

52.	When you have put on weight do you eat less than you usually do?	1	2	3	4	5	6
53.	Do you try to eat less at mealtimes than you would like to eat?	1	2	3	4	5	
54.	How often do you refuse food or drink offered because you are concerned about your weight?	1	2	3	4	5	
55.	Do you watch exactly what you eat?	1	2	3	4	5	
56.	Do you deliberately eat foods that are slimming?	1	2	3	4	5	
57.	When you have eaten too much, do you eat less than usual the following day?	1	2	3	4	5	6
58.	Do you deliberately eat less in order not to become heavier?	1	2	3	4	5	
59.	How often do you try not to eat between meals because you are watching your weight?	1	2	3	4	5	
60.	How often in the evenings do you try not to eat because you are watching your weight?	1	2	3	4	5	
61.	Do you take your weight into account with what you eat?	1	2	3	4	5	
62.	Do you have a desire to eat when you are irritated?	1	2	3	4	5	6
63.	Do you have a desire to eat when you have nothing to do?	1	2	3	4	5	6
64.	Do you have a desire to eat when you are depressed or discouraged?	1	2	3	4	5	6
65	Do you have a desire to eat when you are feeling lonely?	1	2	3	4	5	6
66.	Do you have a desire to eat when somebody lets you down?	1	2	3	4	5	6
67.	Do you have a desire to eat when you are cross?	1	2	3	4	5	6
68.	Do you have a desire to eat when something unpleasant is about to happen?	1	2	3	4	5	
69.	Do you get the desire to eat when you are anxious, worried or tense?	1	2	3	4	5	
70.	Do you have a desire to eat when things are going against you or have gone wrong?	1	2	3	4	5	
71.	Do you have a desire to eat when you are frightened?	1	2	3	4	5	6
72.	Do you have a desire to eat when you are disappointed?	1	2	3	4	5	6
73.	Do you have a desire to eat when you are emotionally upset?	1	2	3	4	5	6
74.	Do you have a desire to eat when you are bored or restless?	1	2	3	4	5	6
75	If food tastes good to you do you eat more than usual?	1	2	3	4	5	
76.	If food smells and looks good do you eat more than usual?	1	2	3	4	5	

1 never	2 seldom	3 sometimes	4 often	5 very often	6 not relevant

77.	If you see or smell something delicious, do you have a desire to eat it?	1	2	3	4	5	
78.	If you have something delicious to eat do you eat it straight away?	1	2	3	4	5	
79.	If you see others eating do you also want to eat?	1	2	3	4	5	
80.	Do you eat more than usual when you see others eating?	1	2	3	4	5	
81.	When preparing a meal are you inclined to eat something?	1	2	3	4	5	
82.	If you walk past the baker do you have the desire to buy something delicious?	1	2	3	4	5	
83.	If you walk past a snackbar or a café, do you have the desire to buy something delicious?	1	2	3	4	5	
84.	Can you resist eating delicious food?	1	2	3	4	5	

Section H: Mother's own childhood

We would like to find out a little about your own childhood. Like all the information we collect, this will be kept entirely confidential. Please indicate below if you prefer not to complete this section.

I do not wish to complete this section □ → please go to Section I on Page 9

85. Was your childhood happy? *Yes* □ *No* □

86. Do you feel you were loved as a child? *Yes* □ *No* □

87. Would you like to be the same kind of mother that your mother was to you? *Yes* □ *No* □

88. Did you feel your parents were pleased with you? *Yes* □ *No* □

89. Did anyone who took care of you ever hit you hard enough to bruise you? **(tick one only)**

No □ *Once* □ *Twice* □ *Three or four times* □ *Five or more times* □

Section I: Your family's experience of health services

Gateshead Health Trust is interested in how you find some of their services and we are collecting this information on their behalf. Any information you give us will be entirely confidential and will not be passed on to staff looking after you or your baby. The information will be used more generally to improve health services for future families.

Unless otherwise stated, please circle the response for each statement that most applies to how you feel.

90. Which of the following do you and your partner turn to for advice about looking after your baby ?

(a) Your parents	*Never---Sometimes---Often*
(b) Other family and/or friends	*Never---Sometimes---Often*
(c) TV, radio, books, magazines	*Never---Sometimes---Often*
(d) GP	*Never---Sometimes---Often*
(e) Health visitor	*Never---Sometimes---Often*
(f) NHS Direct	*Never---Sometimes---Often*
(g) Other **(please specify):**	

91. Did you visit the Breast Feeding Workshop at the QE Hospital before you had your baby ?
Yes ☐ *No* ☐

92. If yes, indicate how you found it by circling the statement that most applies to you: **Please answer each item**

(a) Informative	*Yes---Probably---Possibly---Not at all*
(b) Helpful	*Yes---Probably---Possibly---Not at all*
(c) Supportive	*Yes---Probably---Possibly---Not at all*
(d) Friendly	*Yes---Probably---Possibly---Not at all*

93. How many times has your baby been seen by your GP ? **(tick one only)**

(a) At the surgery ?	*Never* ☐	*Once* ☐	*Twice* ☐	*3-5 times* ☐	*6-12 times* ☐	*More often* ☐
(b) At home ?	*Never* ☐	*Once* ☐	*Twice* ☐	*3-5 times* ☐	*6-12 times* ☐	*More often* ☐

94. Was it easy to make contact with your GP ? *Yes---Probably---No---Not tried*

95. In general, how did you find contact with your GP ?

(a) It reassured me	*Yes---Probably---Possibly---Not at all*
(b) It worried me	*Yes---Probably---Possibly---Not at all*
(c) It made me feel better	*Yes---Probably---Possibly---Not at all*

96. Do you have enough time to talk to your GP ? *Usually---Sometimes---No*

97. How often have you taken your baby to the baby clinic, either for an innoculation or to be weighed ? **(tick one only)**

Never ☐ *Once* ☐ *Twice* ☐ *3-5 times* ☐ *6-12 times* ☐ *More often* ☐

98. On what day is the clinic you mainly attend ? *Mon / Tues / Wed / Thurs / Fri*

99. Approximately what time does the clinic you mainly attend begin ?

9-10 o'clock ☐ *11-12 o'clock* ☐ *1-2 o'clock* ☐ *3-4 o'clock* ☐

100. In your opinion, is the clinic

(a) Near enough to your home ? *Yes---Probably---No*
(b) Convenient for your home ? *Yes ---Probably---No*
(c) Frequent enough ? *Yes---Probably---No*
(d) At a convenient time of day ? *Yes---Probably---No*

101. If the clinic is not held at a convenient time, when would be a better time ? **(tick one only)**

Morning rather than afternoon ☐
Afternoon rather than morning ☐
Earlier in the morning ☐
Later in the morning ☐
Earlier in the afternoon ☐
Later in the afternoon ☐
Early evening / Saturday ☐

102. In general, how did you find visits to the baby clinic ?

(a) They reassured me *Yes---Probably---Possibly---Not at all*
(b) They worried me *Yes---Probably---Possibly---Not at all*
(c) They made me feel better *Yes---Probably---Possibly---Not at all*

103. Do you have enough time to talk to staff ? *Usually---Sometimes---No*

104. Do you have enough privacy to talk to staff ? *Usually---Sometimes---No*

105. How often have you spoken to your health visitor on the telephone since this baby was born ?

Never ☐ *Once* ☐ *Twice* ☐ *3-5 times* ☐ *6-12 times* ☐ *More often* ☐

106. Was it easy to make contact ?

Very easy---Quite easy---No---Not tried

107. If you need to ring your health visitor do you have the telephone number ?

Yes---Not sure---No

108. How many times has your health visitor visited you at home since this baby was born ?
Never ☐ *Once* ☐ *Twice* ☐ *3-5 times* ☐ *6-12 times* ☐ *More often* ☐

109. The last time your health visitor visited your home, was it because **(tick one only)**
I asked him/her to visit ☐ *He/she chose to visit me* ☐

110. At that visit, did your health visitor **(tick one only)**
Just turn up ☐ *Make an appointment* ☐

111. At that visit did you know why your health visitor visited you ? *Yes---Not really---No*

112. In general, how do you find your health visitor's visits ? **(please circle appropriate response for each item)**

(a) Visits were helpful	*Definitely---Probably---Possibly---Not at all*
(b) Visits made me feel better	*Definitely---Probably---Possibly---Not at all*
(c) Visits were frequent enough	*Definitely---Probably---Possibly---Not at all*
(d) Visits reassured me	*Definitely---Probably---Possibly---Not at all*

113. Where do you prefer to see your health visitor ?
At home / baby clinic / both

114. How much do **each** of the following statements generally describe your health visitor ?
(if you have seen more than one, describe the health visitor you have seen most often)

(a) Kind and supportive	*Definitely---Slightly---Not at all*
(b) I hardly know her	*Definitely---Slightly---Not at all*
(c) Knowledgeable and up-to-date	*Definitely---Slightly---Not at all*
(d) Spends too little time with me	*Definitely---Slightly---Not at all*
(e) Gives sound advice	*Definitely---Slightly---Not at all*
(f) Knows me and my family well	*Definitely---Slightly---Not at all*
(g) Bossy and interfering	*Definitely---Slightly---Not at all*
(h) Friendly and chatty	*Definitely---Slightly---Not at all*
(i) Asks me the right questions about how I feel	*Definitely---Slightly---Not at all*
(j) Appears judgmental or disapproving	*Definitely---Slightly---Not at all*
(k) Listens to me	*Definitely---Slightly---Not at all*
(l) Only visits to check up on me	*Definitely---Slightly---Not at all*
(m) Helps arrange things for me	*Definitely---Slightly---Not at all*
(n) Helps me to get to appointments	*Definitely---Slightly---Not at all*

115. Was there anything you particularly liked about the service you received from your health visitor ?

116. Was there anything you particularly disliked about the service you received from your health visitor ?

117. Have you any suggestion for improving the service s/he offers ?

Who completed this questionnaire? **(tick all that apply)**

Baby's mother ☐ *Baby's father* ☐ *Baby's grandparent* ☐ *Nanny* ☐

Childminder ☐ *Nursery* ☐ *Other* **(please tick and specify)** ☐ :

Please write the date you complete this questionnaire ___ / ___ / ___

How old is your baby now? ..... *months and* ..... *weeks*

Is the baby's mother working or studying outside the home at the moment? *Yes* ☐ *No* ☐

If so, who provides child care when the baby's mother is working outside the home?

Not applicable ☐

Baby's father *Most of the time---Some of the time---Occasionally---Never*

Baby's grandparent *Most of the time---Some of the time---Occasionally---Never*

Nanny *Most of the time---Some of the time---Occasionally---Never*

Childminder *Most of the time---Some of the time---Occasionally---Never*

Nursery *Most of the time---Some of the time---Occasionally---Never*

Other **(please tick and specify)** ☐ :

It would help us in our record keeping if you write your name here .....

Was there anything you intended to go back to and complete? Please check. When you have finished, return the questionnaire in the enclosed envelope even if you were not able to complete all of it.

Please check that you have filled in the table on page one with any weight records you have of your baby since filling in the last questionnaire. Have you returned the top pink copies of the Weaning Diary and Finger Feeding Diary in your Personal Child Health Record? If not, please include them when you return this questionnaire, even if they are not filled in.

Thank you for all your help with this part of the Millennium Baby Study. We hope you have enjoyed it. We will see you at the health check, and if you are willing, we hope to keep in touch until your child goes to school.

If the name or address on the envelope was not correct or incomplete, or if you expect to move house in the near future and know your new address, it would help us if you could write it below:

.....

.....

.....Telephone No. .....

Dr. Kathryn Parkinson, Community Child Health, University of Newcastle upon Tyne, 13 Walker Terrace, Gateshead, NE8 1EB. Tel. (0191) 4776000

Figure A.7. ***30 Month Questionnaire***

Feeding and growth study: Your child at 2½ years

ID No:

Please answer as much of this questionnaire as you feel able to. Any information you give us will be helpful. It will be treated in complete confidence, stored securely, and there will be nothing to identify you on this questionnaire unless you choose to put your name on it.

As before, this questionnaire asks about you and your child. If, for any reason, your child is no longer with you, please tick the box below and return the questionnaire to us so we do not trouble you further.

My child is no longer with me ☐

How to fill in the questionnaire

1. Some questions on the following pages can be answered simply by putting a tick in the box next to the answer that applies to you.

   Example *Yes* ☐ *No* ☐

2. Some questions on the following pages can be answered by circling the response that applies to you.

   Example *Not at all---Occasionally---Frequently*

   If you really feel that you are in-between two of the descriptions, you can indicate this by circling the dotted line.

3. Usually after answering each question you go on to the next one unless a box you have ticked has an arrow next to it with an instruction to go to another question.

   Example *Yes* ☐ → Go to Question 5
   No ☐

Weights

Please fill in up to **three recent clinic weights** since the age of 18 months if you have them. If your child has not been weighed recently, perhaps you could make a special visit to the clinic. If your child is due to have his/her 2½ year check by your health visitor or nursery nurse soon, you might prefer to wait until then before returning the questionnaire. You will remember that the weight page in your Personal Child Health Record is normally near the end.

We would also like a record of your child's height if it has been measured at the clinic or by your health visitor at home.

Date	Weight (kg)	Weight (lb/oz)

Date	Height (cm)	Height (feet)

Section A: General feeding questions

1. Which of the following statements describes your child's feeding most accurately? **(tick one only)**

   Generally still needs to be fully fed by a carer ☐
   Generally needs to be fed but eats food with fingers ☐
   Generally eats with spoon or fork but needs help ☐
   Generally eats without help ☐

2. At present, how long does it take to give your child breakfast?

   Less than 5 mins ☐ *5-15 mins* ☐ *15-25 mins* ☐ *25-35 mins* ☐ *35-45 mins* ☐
   45-60 mins ☐ *More than 60 mins* ☐

3. At present, how long does it take to give your child a midday meal?

   Less than 5 mins ☐ *5-15 mins* ☐ *15-25 mins* ☐ *25-35 mins* ☐ *35-45 mins* ☐
   45-60 mins ☐ *More than 60 mins* ☐

4. At present, how long does it take to give your child an evening meal?

   Less than 5 mins ☐ *5-15 mins* ☐ *15-25 mins* ☐ *25-35 mins* ☐ *35-45 mins* ☐
   45-60 mins ☐ *More than 60 mins* ☐

5. In general, how many meals does your child have each day?

   One ☐ *Two* ☐ *Three* ☐ *Four* ☐ *Five* ☐ *Six or more* ☐

6. In general, how many snacks does your child have each day?

   One ☐ *Two* ☐ *Three* ☐ *Four* ☐ *Five* ☐ *Six or more* ☐

7. At present, how much do the following describe your child? **Please answer each item**

   (a) Hungry for foods — *Rarely---Sometimes---Often*
   (b) Loves food — *Rarely---Sometimes---Often*
   (c) Cannot fill him/her — *Rarely---Sometimes---Often*
   (d) Will not take solids — *Rarely---Sometimes---Often*
   (e) Eats a limited variety of food — *Rarely---Sometimes---Often*
   (f) Uninterested in food — *Rarely---Sometimes---Often*
   (g) Prefers drinks to food — *Rarely---Sometimes---Often*
   (h) Slow feeder — *Rarely---Sometimes---Often*

8. Does your child do any of the following when eating? **Please answer each item**

   (a) Pushes food/spoon away — *Rarely---Sometimes---Often---Only feeds self*
   (b) Can't chew solid foods — *Rarely---Sometimes---Often*
   (c) Gags on food — *Rarely---Sometimes---Often*
   (d) Holds food in mouth — *Rarely---Sometimes---Often*
   (e) Spits food out — *Rarely---Sometimes---Often*
   (f) Throws food — *Rarely---Sometimes---Often*
   (g) Cries/screams during meals — *Rarely---Sometimes---Often*

9. If your child does not finish a course, or part of a meal, what do you do?
   (a) Encourage him/her to eat — *Rarely---Sometimes---Often*
   (b) Make him/her eat the food — *Rarely---Sometimes---Often*
   (c) Offer something else — *Rarely---Sometimes---Often*

10. If your child does not finish a course, or part of a meal, what do you do after the meal?
   (a) Offer the same food again later — *Rarely---Sometimes---Often*
   (b) Offer something else later — *Rarely---Sometimes---Often*
   (c) Offer nothing else until the next meal — *Rarely---Sometimes---Often*

11. Which of the following do you use to encourage him/her to eat?
   (a) Have the TV or video on — *Never---Rarely---Sometimes---Often*
   (b) Play music — *Never---Rarely---Sometimes---Often*
   (c) Play games with food — *Never---Rarely---Sometimes---Often*
   (d) Offer novelty food, e.g. Postman Pat spaghetti shapes — *Never---Rarely---Sometimes---Often*
   (e) Say food will be taken away or given to someone else — *Never---Rarely---Sometimes---Often*
   (f) Offer food reward, e.g. dessert, sweets — *Never---Rarely---Sometimes---Often*
   (g) Offer other reward, e.g. trip to park, watch TV — *Never---Rarely---Sometimes---Often*

12. Do you punish your child for behaving badly during mealtimes in the following ways?
   (a) Sent to bedroom — *Never---Rarely---Sometimes---Often*
   (b) Food taken away — *Never---Rarely---Sometimes---Often*
   (c) Threaten to smack — *Never---Rarely---Sometimes---Often*
   (d) Smack — *Never---Rarely---Sometimes---Often*
   (e) Not give dessert / pudding / sweets — *Never---Rarely---Sometimes---Often*

13. At present, how is your child's appetite? — *Very good---Good---All right---Poor---Very poor*

14. Overall, is your child feeding enough? — *Yes--Not always---No*

15. At present, are feeding times for you usually:
   Very relaxed---Relaxed---All right---Stressful---Very stressful

16. At present, are feeding times for your child usually:
   Very relaxed---Relaxed---All right---Stressful---Very stressful---Can't tell

17. At present, is your child easy to feed? — *Very easy---Easy---All right---Difficult---Very difficult*

18. At present, does your child vomit? — *Rarely---Sometimes---Often*

19. At present, is your child — *Very thin---Thin---Average---Chubby---Fat*

20. Is there anything else you would like to say about feeding your child?

Section B: Food preferences

This section asks about lots of different foods that your child might have tasted. Please indicate whether your child likes or dislikes each type of food by circling one of the numbers according to the scale below.

1 = Dislikes a lot	2 = Dislikes a little	3 = Neither likes nor dislikes	4 = Likes a little	5 = Likes a lot	6 = Never tried

For example, if your child loves spaghetti, circle 5 (for 'likes a lot'), and if you child has never tried the food, circle 6 (for 'never tried'), and so on. **Please answer each item**

21	*White bread*	1	2	3	4	5	6
22.	*Brown or wholemeal bread*	1	2	3	4	5	6
23.	*Boiled rice*	1	2	3	4	5	6
24.	*Spaghetti*	1	2	3	4	5	6
25.	*Crumpets*	1	2	3	4	5	6
26.	*Jam doughnuts*	1	2	3	4	5	6
27.	*Jam tarts*	1	2	3	4	5	6
28.	*Scones*	1	2	3	4	5	6
29.	*Fruit cake*	1	2	3	4	5	6
30.	*Fancy iced cakes*	1	2	3	4	5	6
31.	*Jam sponge cake*	1	2	3	4	5	6
32.	*Milk Chocolate digestives*	1	2	3	4	5	6
33.	*Custard creams*	1	2	3	4	5	6
34.	*Rich Tea biscuits*	1	2	3	4	5	6
35.	*Coco Pops*	1	2	3	4	5	6
36.	*Muesli*	1	2	3	4	5	6
37.	*Frosties*	1	2	3	4	5	6
38.	*Corn Flakes*	1	2	3	4	5	6
39.	*Weetabix*	1	2	3	4	5	6
40.	*Lemon meringue pie*	1	2	3	4	5	6
41.	*Crème caramel pudding*	1	2	3	4	5	6
42.	*Chocolate mousse*	1	2	3	4	5	6
43.	*Rice pudding*	1	2	3	4	5	6
44.	*Jelly*	1	2	3	4	5	6
45.	*Ice cream*	1	2	3	4	5	6
46.	*Choc Ice*	1	2	3	4	5	6
47.	*Milk*	1	2	3	4	5	6
48.	*Yoghurt*	1	2	3	4	5	6
49.	*Cheddar cheese*	1	2	3	4	5	6
50.	*Cheese spread*	1	2	3	4	5	6
51.	*Ricotta*	1	2	3	4	5	6
52.	*Eggs*	1	2	3	4	5	6
53.	*Quiche*	1	2	3	4	5	6
54.	*Butter*	1	2	3	4	5	6
55.	*Margarine (such as Flora)*	1	2	3	4	5	6
56.	*Bacon*	1	2	3	4	5	6
57.	*Minced beef*	1	2	3	4	5	6
58.	*Shepherd's pie*	1	2	3	4	5	6
59.	*Pork chops*	1	2	3	4	5	6
60.	*Lamb chops*	1	2	3	4	5	6
61.	*Roast chicken*	1	2	3	4	5	6
62.	*Partridge*	1	2	3	4	5	6
63.	*Sausages*	1	2	3	4	5	6
64.	*Sausage roll*	1	2	3	4	5	6
65.	*Burger*	1	2	3	4	5	6
66.	*Soup*	1	2	3	4	5	6

67.	*Fish in batter*	1	2	3	4	5	6
68.	*Fish fingers*	1	2	3	4	5	6
69.	*Haddock*	1	2	3	4	5	6
70.	*Tuna*	1	2	3	4	5	6
71.	*Sushi*	1	2	3	4	5	6
72.	*Sugar*	1	2	3	4	5	6
73.	*Jam*	1	2	3	4	5	6
74.	*Milk chocolate*	1	2	3	4	5	6
75.	*Toffee*	1	2	3	4	5	6
76.	*Fruit pastilles*	1	2	3	4	5	6
77.	*Peppermints*	1	2	3	4	5	6
78.	*Crisps*	1	2	3	4	5	6
79.	*Wotsits or Quavers or Monster Munch*	1	2	3	4	5	6
80.	*Chips*	1	2	3	4	5	6
81.	*Boiled potatoes*	1	2	3	4	5	6
82.	*Roast potatoes*	1	2	3	4	5	6
83.	*Yam*	1	2	3	4	5	6
84.	*Carrots*	1	2	3	4	5	6
85.	*Tomatoes*	1	2	3	4	5	6
86.	*Baked beans*	1	2	3	4	5	6
87.	*Peas*	1	2	3	4	5	6
88.	*Lettuce*	1	2	3	4	5	6
89.	*Cucumber*	1	2	3	4	5	6
90.	*Onions*	1	2	3	4	5	6
91.	*Okra*	1	2	3	4	5	6
92.	*Cabbage*	1	2	3	4	5	6
93.	*Gourd*	1	2	3	4	5	6
94.	*Oranges*	1	2	3	4	5	6
95.	*Mangoes*	1	2	3	4	5	6
96.	*Apples*	1	2	3	4	5	6
97.	*Lychees*	1	2	3	4	5	6
98.	*Banana*	1	2	3	4	5	6
99.	*Guava*	1	2	3	4	5	6
100.	*Tinned peaches*	1	2	3	4	5	6
101.	*Tea*	1	2	3	4	5	6
102.	*Coffee*	1	2	3	4	5	6
103.	*Coco Cola or Pepsi*	1	2	3	4	5	6
104.	*Lemonade*	1	2	3	4	5	6
105.	*Orange juice*	1	2	3	4	5	6
106.	*Apple juice*	1	2	3	4	5	6
107.	*Tomato ketchup*	1	2	3	4	5	6
108.	*Mustard*	1	2	3	4	5	6
109.	*Vinegar*	1	2	3	4	5	6
110.	*Mayonnaise*	1	2	3	4	5	6

111. What types of food does your child like or dislike? Foods that are:

a)	*Bright or colourful*	1	2	3	4	5	6
b)	*Slimy*	1	2	3	4	5	6
c)	*Crunchy*	1	2	3	4	5	6
d)	*Chewy*	1	2	3	4	5	6
e)	*Soft or sloppy*	1	2	3	4	5	6
f)	*Messy or sticky*	1	2	3	4	5	6
g)	*Mixed up together (e.g. stews)*	1	2	3	4	5	6
h)	*Strongly favoured (e.g. curry)*	1	2	3	4	5	6
i	*Unfamiliar*	1	2	3	4	5	6

112. Does your child ever refuse to eat a food because

a) it is "damaged" (e.g. a broken biscuit) *Never---Rarely---Sometimes---Often*

b) not a particular brand *Never---Rarely---Sometimes---Often*

Section C: Your child's drinks

113. At present, how does your child drink?

- *Mainly drinks from breast* ☐
- *Mainly drinks from bottle* ☐
- *Mainly drinks from feeder cup* ☐
- *Mainly drinks from cup* ☐
- *Mainly drinks from cup or other with straw* ☐

114. My child drinks continuously throughout the day *Usually---Sometimes---Rarely---Never*

115. Each day, how much of the following does your child drink:

(a) Milk	*None / 1 cup / 2 cups / 3 cups / 4 cups / 5 or more cups*
(b) Fresh fruit juice	*None / 1 cup / 2 cups / 3 cups / 4 cups / 5 or more cups*
(c) Fruit juice, squash, Ribena, etc.	*None / 1 cup / 2 cups / 3 cups / 4 cups / 5 or more cups*
(d) Soft drinks (e.g. Coke, Fanta)	*None / 1 cup / 2 cups / 3 cups / 4 cups / 5 or more cups*
(e) Low calorie drinks (e.g. Diet Coke)	*None / 1 cup / 2 cups / 3 cups / 4 cups / 5 or more cups*
(f) Water	*None / 1 cup / 2 cups / 3 cups / 4 cups / 5 or more cups*
(g) Tea	*None / 1 cup / 2 cups / 3 cups / 4 cups / 5 or more cups*
(h) Coffee	*None / 1 cup / 2 cups / 3 cups / 4 cups / 5 or more cups*
(i) Hot milk drinks	*None / 1 cup / 2 cups / 3 cups / 4 cups / 5 or more cups*

Section D: Your child's behaviour

116. Does your child go to sleep at the same time every night?

Nearly always---Usually---Some of the time---Rarely---Never

117. In general, when do you put your child to bed? ..... *o'clock*

118. In general, what time does your child fall asleep? ..... *o'clock*

119. In general, how often does your child wake in the night?

No □ *Once* □ *Twice* □ *Three times* □ *Four or more times* □

120. What usually happens if your child wakes? **Please answer each item**

(a) Leave him/her alone *Nearly always--Usually--Rarely--Never*
(b) Go and check but speak to him/her only. *Nearly always--Usually--Rarely--Never*
(c) Pick him/her up and comfort, then leave in bed *Nearly always--Usually--Rarely--Never*
(d) Take him/her into bed with me *Nearly always--Usually--Rarely--Never*
(e) Child climbs into my bed on his/her own *Nearly always--Usually--Rarely--Never*

121. Does your child have a daytime nap?

More than once a day □ *Everyday* □ *Every two days* □
Every three to four days □ *Every five to seven days* □ *Less than once a week* □ *Never* □

122. On average, how often does your child cry during the day?

1-2 times a day □ *3-4 times a day* □ *5 or more times a day* □

123. Does your child have temper tantrums?

More than once a day □ *Everyday* □ *Every two days* □
Every three to four days □ *Every five to seven days* □ *Less than once a week* □ *Never* □

124. Does your child ever hold his/her breath when angry/frightened until he/she turns blue?

More than once a day □ *Everyday* □ *Every two days* □
Every three to four days □ *Every five to seven days* □ *Less than once a week* □ *Never* □

125. Is your child using a potty/toilet to pass urine (wee) during the day?

Nearly always--Usually--Rarely--Never

126. Is your child using a potty/toilet to pass a stool (poo) during the day?

Nearly always--Usually--Rarely--Never

127. On average, how often does your child pass a stool (poo)?

More than once a day □ *Everyday* □ *Every two days* □
Every three to four days □ *Every five to seven days* □ *Less than once a week* □

128. Does your child:

(a) Bite *Never---Rarely---Sometimes---Often*
(b) Pinch *Never---Rarely---Sometimes---Often*

Section E: Repetitive behaviour

Young children often repeat the same behaviour over and over again. Some children are more repetitive than others and we need to understand why this is so. Please rate the repetitive behaviours your child has shown **over the last month** and rate the most **usual** way he/she displays this behaviour.

129.	**Does your child:**	***Never*** *or rarely*	***One*** *or more times daily*	***15*** *or more times daily (or once an hour)*	***30*** *or more times daily (or twice an hour)*
a)	Arrange toys or other items in rows or patterns?	☐	☐	☐	☐
b)	Repetitively fiddle with toys or other items? (e.g. spin, twiddle, bang, tap, twist, or flick anything repeatedly?)	☐	☐	☐	☐
c)	Spin him/herself around and around?	☐	☐	☐	☐
d)	Rock backwards and forwards, or side to side, either when sitting or when standing?	☐	☐	☐	☐
e)	Pace or move around repetitively? (e.g. walk to and fro across a room, or around the same path in the garden?)	☐	☐	☐	☐
f)	Make repetitive hand and/or finger movements? (e.g. flap, wave, or flick, his/her hands or fingers repetitively?)	☐	☐	☐	☐

130.	**Does your child:**	**Never** or rarely	**Mild** or **occasional**	**Marked** or **notable**
a)	Have a fascination with specific objects? (e.g. trains, road signs or other things?)	☐	☐	☐
b)	Like to look at objects from particular or unusual angles?	☐	☐	☐
c)	Have a special interest in the smell of people or objects?	☐	☐	☐
d)	Have a special interest in the feel of different surfaces?	☐	☐	☐
e)	Have any special objects he/she likes to carry around? (e.g. a teddy, a blanket, a book, or a stick?)	☐	☐	☐
	Please describe the object:			
	.....			
f)	Collect or hoard items of any sort?	☐	☐	☐
	Please describe what your child collects:			
	.....			

131. **Does your child:**		***Never*** *or rarely*	***Mild*** *or* ***occasional*** *(does not effect others)*	***Marked*** *or* ***notable*** *(effects others on a regular basis)*
a)	Insist on things at home remaining the same? (e.g. furniture staying in the same place, things being kept in certain places, or arranged in certain ways?)	☐	☐	☐
b)	Get upset about minor changes to objects (e.g. flecks of dirt on his clothes, minor scratches on toys)	☐	☐	☐
c)	Insist that aspects of daily routine must remain the same?	☐	☐	☐
d)	Insist on doing things in a certain way or re-doing things until they are "just right"?	☐	☐	☐

132. **Does your child:**		***Never*** *or rarely*	***Mild*** *or* ***occasional*** *(will tolerate alternatives when necessary)*	***Marked*** *or* ***notable*** *(will not tolerate any alternatives)*
a)	Play the same music, game or video, or read the same book repeatedly?	☐	☐	☐
b)	Insist on wearing the same clothes or refuse to wear new clothes?	☐	☐	☐
c)	Insist on eating the same foods, or a very small range of foods, at every meal?	☐	☐	☐

133. What sort of activity will your child choose if they are left to occupy themselves? **(tick one)**

(a) *A range of different and flexible self-chosen activities* ☐

(b) *Some varied and flexible interests but commonly chooses the same activities* ☐

(c) *Almost always chooses from a restricted range of repetitive activities* ☐

Section F: Difficulties with your child

Difficulties with behaviour are extremely common in two year olds. The following section asks about problems you might be experiencing with your child at the moment.

134. Do you see your child as having eating problems at present? *Yes* ☐
Sometimes ☐
No ☐ → Go to Question 137

135. What sort of eating problems does your child have? **Please answer each item**

(a) A poor eater	*Definitely---Maybe---No*
(b) A faddy eater	*Definitely---Maybe---No*
(c) A greedy eater	*Definitely---Maybe---No*
(d) Behaves badly at mealtimes	*Definitely---Maybe---No*

136. Have you ever asked for or received help with your child's eating from **Please answer each item**

(a) GP	*No---Yes, helpful---Yes, not helpful*
(b) Health visitor	*No---Yes, helpful---Yes, not helpful*
(c) Dietitican	*No---Yes, helpful---Yes, not helpful*
(d) Paediatrician	*No---Yes, helpful---Yes, not helpful*
(e) Psychologist	*No---Yes, helpful---Yes, not helpful*

137. Do you have any problems with your child's sleeping? **Please answer each item**

(a) Won't go to sleep	*No problem---Slight problem---Big problem*
(b) Wakes during night	*No problem---Slight problem---Big problem*
(c) Won't sleep in own bed	*No problem---Slight problem---Big problem*

138. Do you have any problems with your child crying? **Please answer each item**

(a) Cries too often	*Nearly always---Usually---Rarely---Never*
(b) Difficult to comfort	*Nearly always---Usually---Rarely---Never*
(c) Cries during night	*Nearly always---Usually---Rarely---Never*
(d) Cries when separated from mother	*Nearly always---Usually---Rarely---Never*

139. Does your child have any problems passing stools (doing a poo)? **Please answer each item**

(a) Difficulty passing stools	*No problem---Slight problem---Big problem*
(b) Hard stools	*No problem---Slight problem---Big problem*
(c) Pain passing stools	*No problem---Slight problem---Big problem*
(d) Rarely passes stools	*No problem---Slight problem---Big problem*

140. Is your child ever: **Please answer each item**

(a) Possessive	*Never---Rarely---Sometimes---Often*
(b) Jealous of brothers and sisters	*Never---Rarely---Sometimes---Often---Not applicable*
(c) Difficult to control	*Never---Rarely---Sometimes---Often*
(d) A poor sleeper	*Never---Rarely---Sometimes---Often*
(e) Very shy, fearful, anxious	*Never---Rarely---Sometimes---Often*

141. Have you ever asked for or received help for any other behaviour **apart** from eating from **Please answer each item**

(a) GP	*No---Yes, helpful---Yes, not helpful*
(b) Health visitor	*No---Yes, helpful---Yes, not helpful*
(c) Paediatrician	*No---Yes, helpful---Yes, not helpful*
(d) Psychologist	*No---Yes, helpful---Yes, not helpful*
(e) Social Worker	*No---Yes, helpful---Yes, not helpful*
(f) Other *(please specify)*	.....

Section G: Your child's illnesses

142. Since the age of one year, has your child seen the doctor due to illness, either at home or at the surgery? **(tick one only)**

No ☐ *Once* ☐ *More than once* ☐

143. Since filling in the questionnaire at one year, has your child been admitted to hospital? **(tick one only)**

No ☐ *Once* ☐ *More than once* ☐

144. Please describe the reasons for each admission and the treatment given.

Who completed this questionnaire? **(tick all that apply)**

Child's mother □ *Child's father* □ *Child's grandparent* □ *Nanny* □
Childminder □ *Nursery* □ *Other* **(please tick and specify)** □ :

Please write the date you completed this questionnaire ___ / ___ / ___

How old is the child now? ..... *years and* ..... *months*

Is the child's mother working or studying outside the home at the moment? *Yes* □ *No* □

If so, who provides child care when the child's mother is working outside the home?

Not applicable	□
Child's father	*Most of the time---Some of the time---Occasionally---Never*
Child's grandparent	*Most of the time---Some of the time---Occasionally---Never*
Nanny	*Most of the time---Some of the time---Occasionally---Never*
Childminder	*Most of the time---Some of the time---Occasionally---Never*
Nursery	*Most of the time---Some of the time---Occasionally---Never*
Other **(please tick and specify)** □ :	

We know that lots of people change their child's GP over the first few years and we need to keep our records up to date. Please write below the child's present GP, and previous GP if appropriate.

Present GP ..... Previous GP.....

It would help us in our record keeping if you write your name here .....

Was there anything you intended to go back to and complete? Please check. When you have finished, return the questionnaire in the enclosed envelope even if you were not able to complete all of it.

Please check that you have filled in the table on page one with up to three recent weights of the child.

Thank you for your help with this part of the Millennium Baby Study.

If the name or address on the envelope was not correct or incomplete, or if you expect to move house in the near future and know your new address, it would help us if you could write it below:

.....

.....

.....Telephone No. .....

Dr. Kathryn Parkinson, Community Child Health, University of Newcastle, 13 Walker Terrace, Gateshead, NE8 1EB. Tel. (0191) 4776000

Appendix B

Section Non-Response

Table B.1. Section Non-Response for Newborn Questionnaire

Newborn Questionnaire	≥ 1 **Qu. answered by respondents to qu're**	**% of Respondents (/1027)**	**% of Recruits (/1029)**
Section A	1024	99.7	99.5
Section B	1016	98.9	98.7
Section C	1022	99.5	99.3

Table B.2. Section Non-Response for 6 Week Questionnaire

6 Week Questionnaire	**≥ 1 Qu. answered by respondents to qu're**	**% of Respondents (/831)**	**% of Recruits (/1029)**
Section A	829	99.8	80.6
Section B	801	96.4	77.8
Section C	831	100.0	80.8
Section D	830	99.9	80.7
Section E	820	98.7	79.7

Table B.3. Section Non-Response for 4 Month Questionnaire

4 Month Questionnaire	**≥ 1 Qu. answered by respondents to qu're**	**% of Respondents (/762)**	**% of Recruits (/1029)**
Section A	754	99.0	73.3
Section B	750	98.4	72.9
Section C	762	100.0	74.1
Section D	752	98.7	73.1
Section E	755	99.1	73.4
Section F	745	97.8	72.4

Table B.4. Section Non-Response for 8 Month Questionnaire

8 Month Questionnaire	**≥ 1 Qu. answered by respondents to qu're**	**% of Respondents (/676)**	**% of Recruits (/1029)**
Section A	669	99.0	65.0
Section B	633	93.6	61.5
Section C	676	100.0	65.7
Section D	667	98.7	64.8
Section E	666	98.5	64.7
Section F	673	99.6	65.4

Table B.5. Section Non-Response for 12 Month Questionnaire

12 Month Questionnaire	**≥ 1 Qu. answered by respondents to qu're**	**% of Respondents (/633)**	**% of Recruits (/1029)**
Section A	626	98.9	60.8
Section B	156	24.6	15.2
Section C	632	99.8	61.4
Section D	632	99.8	61.4
Section E	632	99.8	61.4
Section F	622	98.3	60.4
Section G	630	99.5	61.2
Section H	605	95.6	58.8
Section I	629	99.4	61.1

Table B.6. Section Non-Response for 30 Month Questionnaire

30 Month Questionnaire	≥ 1 **Qu. answered by respondents to qu're**	**% of Respondents (/491)**	**% of Recruits (/1029)**
Section A	490	99.8	47.6
Section B	491	100.0	47.7
Section C	0	0.0	0.0
Section D	490	99.8	47.6
Section E	0	0.0	0.0
Section F	489	99.6	47.5
Section G	489	99.6	47.5

Appendix C

Item Non-Response

Table C.1. Item Non-Response for Newborn Questionnaire

Newborn Questionnaire Questions	Qu. answered by respondents to qu're	% of Respondents (/1027)	% of Recruits (/1029)
1	1019	99.2	99.0
2	1022	99.5	99.3
3	989	96.3	96.1
4	984	95.8	95.6
5	971	94.5	94.4
6	975	94.9	94.8
7	981	95.5	95.3
8	983	95.7	95.5
9(a)	966	94.1	93.9
9(b)	953	92.8	92.6

cont.

Newborn Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/1027)**	**% of Recruits (/1029)**
9(c)	954	92.9	92.7
10(a)	923	89.9	89.7
10(b)	929	90.5	90.3
10(c)	916	89.2	89.0
10(d)	921	89.7	89.5
10(e)	913	88.9	88.7
11	976	95.0	94.8
12	951	92.6	92.4
13	1008	98.1	98.0
14	1020	99.3	99.1
15	1020	99.3	99.1
16	1016	98.9	98.7
17	1021	99.4	99.2
18	1020	99.3	99.1
19	1019	99.2	99.0
20	1020	99.3	99.1
21	1020	99.3	99.1
22	1018	99.1	98.9

Table C.2. Item Non-Response for 6 Week Questionnaire

6 week Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/831)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
1	828	99.6	80.5	
2	822	98.9	79.9	
3	789	94.9	76.7	
4a	752	90.5	73.1	
4b	659	79.3	64.0	
5	822	98.9	79.9	
Stopped BF	334	40.2	32.5	
6	173	20.8	16.8	100.0
7	170	20.5	16.5	98.3
8a	159	19.1	15.5	91.9
8b	161	19.4	15.6	93.1
8c	163	19.6	15.8	94.2
8d	161	19.4	15.6	93.1
8e	158	19.0	15.4	91.3
8f	158	19.0	15.4	91.3
8g	77	9.3	7.5	44.5
9	801	96.4	77.8	
10	21	2.5	2.0	100.0
11a	21	2.5	2.0	100.0

cont.

6 week Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/831)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
11b	8	1.0	0.8	100.0
12a	13	1.6	1.3	61.9
12b	16	1.9	1.6	76.2
12c	18	2.2	1.7	85.7
13	20	2.4	1.9	95.2
14	21	2.5	2.0	100.0
15	19	2.3	1.8	90.5
16a	18	2.2	1.7	85.7
16b	17	2.0	1.7	81.0
16c	18	2.2	1.7	85.7
16d	18	2.2	1.7	85.7
16e	21	2.5	2.0	100.0
17	823	99.0	80.0	
18	826	99.4	80.3	
19	826	99.4	80.3	
20	813	97.8	79.0	
21	818	98.4	79.5	
22	822	98.9	79.9	
23a	820	98.7	79.7	
23b	818	98.4	79.5	
23c	821	98.8	79.8	

cont.

6 week Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/831)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
24a	819	98.6	79.6	
24b	819	98.6	79.6	
24c	812	97.7	78.9	
24d	817	98.3	79.4	
24e	807	97.1	78.4	
25	831	100.0	80.8	
26	818	98.4	79.5	
27	828	99.6	80.5	
Anything else feeding	831	100.0	80.8	
28	818	98.4	79.5	
29a	791	95.2	76.9	
29b	809	97.4	78.6	
29c	784	94.3	76.2	
29d	784	94.3	76.2	
29e	788	94.8	76.6	
29f	781	94.0	75.9	
30	823	99.0	80.0	

cont.

6 week Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/831)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
31a age	72	8.7	7.0	100.0
31a no of nights	69	8.3	6.7	95.8
31a reason	72	8.7	7.0	100.0
31b age	7	0.8	0.7	100.0
31b no of nights	7	0.8	0.7	100.0
31b reason	7	0.8	0.7	100.0
31c age	1	0.1	0.1	14.3
31c no of nights	1	0.1	0.1	14.3
31c reason	1	0.1	0.1	14.3
32	803	96.6	78.0	
33	797	95.9	77.5	
34	797	95.9	77.5	
35	791	95.2	76.9	
35wait	792	95.3	77.0	
36	301	36.2	29.3	80.9
37	320	38.5	31.1	86.0
38	356	42.8	34.6	95.7
39	797	95.9	77.5	
40	796	95.8	77.4	
41	806	97.0	78.3	

cont.

6 week Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/831)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
42	798	96.0	77.6	
43	772	92.9	75.0	
44	767	92.3	74.5	
45	796	95.8	77.4	
46	781	94.0	75.9	
47	742	89.3	72.1	
48	764	91.9	74.2	
49	791	95.2	76.9	
50	808	97.2	78.5	
51	780	93.9	75.8	
52	790	95.1	76.8	
Had bath	814	98.0	79.1	
53	768	92.4	74.6	96.0
54	777	93.5	75.5	97.1
55	759	91.3	73.8	94.9
56	790	95.1	76.8	
57	800	96.3	77.7	
Hair washed	814	98.0	79.1	
58	778	93.6	75.6	97.7
59	774	93.1	75.2	97.2

cont.

6 week Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/831)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
60	807	97.1	78.4	
61	801	96.4	77.8	
62	803	96.6	78.0	
63	802	96.5	77.9	
64	798	96.0	77.6	
65	794	95.5	77.2	
66	803	96.6	78.0	
67	802	96.5	77.9	
68	796	95.8	77.4	
Placed in car seat	809	97.4	78.6	
69	711	85.6	69.1	97.9
70	700	84.2	68.0	96.4
71	718	86.4	69.8	98.9
72	712	85.7	69.2	98.1
Returned from being away	798	96.0	77.6	
73	599	72.1	58.2	96.8
74	805	96.9	78.2	
75	807	97.1	78.4	
76	808	97.2	78.5	
77	801	96.4	77.8	

cont.

6 week Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/831)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
78	793	95.4	77.1	
79	798	96.0	77.6	
80	803	96.6	78.0	
81	802	96.5	77.9	
82	802	96.5	77.9	
83	95	11.4	9.2	
83 score	133	16.0	12.9	

Table C.3. Item Non-Response for 4 Month Questionnaire

4 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/762)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
1	754	99.0	73.3	
2	752	98.7	73.1	
3	747	98.0	72.6	
4a	714	93.7	69.4	
4b	657	86.2	63.8	
Stopped BF	332	43.6	32.3	
5	97	12.7	9.4	91.5
6	100	13.1	9.7	94.3
7a	92	12.1	8.9	86.8
7b	96	12.6	9.3	90.6
7c	94	12.3	9.1	88.7
7d	93	12.2	9.0	87.7
7e	89	11.7	8.6	84.0
7f	92	12.1	8.9	86.8
7g	91	11.9	8.8	85.8
7h	43	5.6	4.2	40.6
8	750	98.4	72.9	
9	676	88.7	65.7	96.2
10a	694	91.1	67.4	98.7

cont.

4 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/762)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
10b	601	78.9	58.4	96.5
11a	477	62.6	46.4	67.9
11b	551	72.3	53.5	78.4
11c	579	76.0	56.3	82.4
12	696	91.3	67.6	99.0
13	697	91.5	67.7	99.1
14	687	90.2	66.8	97.7
15a	652	85.6	63.4	92.7
15b	639	83.9	62.1	90.9
15c	644	84.5	62.6	91.6
15d	679	89.1	66.0	96.6
15e	679	89.1	66.0	96.6
16	698	91.6	67.8	99.3
17a	650	85.3	63.2	92.5
17b	641	84.1	62.3	91.2
17c	651	85.4	63.3	92.6
17d	651	85.4	63.3	92.6
17e	655	86.0	63.7	93.2
18	755	99.1	73.4	
19	757	99.3	73.6	
20	756	99.2	73.5	

cont.

4 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/762)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
21	755	99.1	73.4	
22	752	98.7	73.1	
23	756	99.2	73.5	
24	755	99.1	73.4	
25a	756	99.2	73.5	
25b	755	99.1	73.4	
25c	756	99.2	73.5	
26a	757	99.3	73.6	
26b	756	99.2	73.5	
26c	757	99.3	73.6	
26d	754	99.0	73.3	
26e	754	99.0	73.3	
27	753	98.8	73.2	
28	756	99.2	73.5	
29	758	99.5	73.7	
Else feed	113	14.8	11.0	

cont.

4 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/762)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
30	752	98.7	73.1	
31	20	2.6	1.9	95.2
32	21	2.8	2.0	100.0
32 other	7	0.9	0.7	100.0
33	21	2.8	2.0	100.0
33 other	0	0.0	0.0	0.0
34	21	2.8	2.0	100.0
35	20	2.6	1.9	95.2
36	21	2.8	2.0	100.0
36 other	0	0.0	0.0	0.0
37	18	2.4	1.7	85.7
38a	21	2.8	2.0	100.0
38b	0	0.0	0.0	0.0
39	751	98.6	73.0	
40a	696	91.3	67.6	
40b	730	95.8	70.9	
40c	676	88.7	65.7	
40d	688	90.3	66.9	
40e	689	90.4	67.0	
40f	388	50.9	37.7	
40f other	161	21.1	15.6	

cont.

4 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/762)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
41	749	98.3	72.8	
42a age	22	2.9	2.1	37.9
42a reason	57	7.5	5.5	98.3
42a no. of nights	15	2.0	1.5	25.9
42b	9	1.2	0.9	75.0
Do not wish to complete	56	7.3	5.4	
43 partner	0	0.0	0.0	0.0
43 child	5	0.7	0.5	0.7
43 parent	14	1.8	1.4	2.0
43 other	106	13.9	10.3	15.4
44 arg	205	26.9	19.9	29.8
44 unf	11	1.4	1.1	1.6
44 sep	41	5.4	4.0	6.0
44 div	3	0.4	0.3	0.4
44 dovi	4	0.5	0.4	0.6
45 self	10	1.3	1.0	1.5
45 partner	7	0.9	0.7	1.0
45 child	11	1.4	1.1	1.6
45 parent	38	5.0	3.7	5.5
45 other	62	8.1	6.0	9.0

cont.

4 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/762)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
46 loan	6	0.8	0.6	0.9
46 decrease income	144	18.9	14.0	20.9
46 general money worries	201	26.4	19.5	29.2
47u changed jobs	29	3.8	2.8	4.2
47u left job	59	7.7	5.7	8.6
47u lost job	16	2.1	1.6	2.3
47u job demotion	7	0.9	0.7	1.0
47p changed jobs	73	9.6	7.1	10.6
47p left job	10	1.3	1.0	1.5
47p lost job	37	4.9	3.6	5.4
47p job demotion	5	0.7	0.5	0.7
48 arrested	2	0.3	0.2	0.3
48 victim of crime	39	5.1	3.8	5.7
48 victim of police brutality	0	0.0	0.0	0.0
49 single parent	40	5.2	3.9	5.8
49 child custody	18	2.4	1.7	2.6
50	369	48.4	35.9	53.6
51 car accident	19	2.5	1.8	2.8
51 other major accident	4	0.5	0.4	0.6
52	69	9.1	6.7	10.0
53	276	36.2	26.8	40.1

Table C.4. Item Non-Response for 8 Month Questionnaire

8 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/676)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
1	669	99.0	65.0	
2	660	97.6	64.1	
3	108	16.0	10.5	98.2
4	91	13.5	8.8	82.7
5	565	83.6	54.9	
6	314	46.4	30.5	92.9
7a	330	48.8	32.1	97.6
7b	318	47.0	30.9	98.1
8	331	49.0	32.2	97.9
9	326	48.2	31.7	96.4
10a	306	45.3	29.7	90.5
10b	296	43.8	28.8	87.6
10c	298	44.1	29.0	88.2
10d	325	48.1	31.6	96.2
10e	322	47.6	31.3	95.3
11	331	49.0	32.2	97.9
12a	556	82.2	54.0	
12b	580	85.8	56.4	
12c	454	67.2	44.1	

cont.

8 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/676)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
12d	572	84.6	55.6	
13	661	97.8	64.2	
14	669	99.0	65.0	
15	656	97.0	63.8	
16	653	96.6	63.5	
17	660	97.6	64.1	
18	667	98.7	64.8	
19	669	99.0	65.0	
20a	664	98.2	64.5	
20b	668	98.8	64.9	
20c	658	97.3	63.9	
20d	656	97.0	63.8	
20e	655	96.9	63.7	
20f	658	97.3	63.9	
20g	660	97.6	64.1	
20h	658	97.3	63.9	
20i	657	97.2	63.8	

cont.

8 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/676)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
21a	659	97.5	64.0	
21b	660	97.6	64.1	
21c	660	97.6	64.1	
21d	650	96.2	63.2	
21e	646	95.6	62.8	
21f	654	96.7	63.6	
21g	655	96.9	63.7	
21h	649	96.0	63.1	
21i	647	95.7	62.9	
22a	654	96.7	63.6	
22b	628	92.9	61.0	
22c	649	96.0	63.1	
23a	639	94.5	62.1	
23b	646	95.6	62.8	
23c	643	95.1	62.5	
24	669	99.0	65.0	
25	667	98.7	64.8	
26	667	98.7	64.8	
27	665	98.4	64.6	
28	669	99.0	65.0	

cont.

8 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/676)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
29	662	97.9	64.3	
30	667	98.7	64.8	
31	119	17.6	11.6	
32	664	98.2	64.5	
33	59	8.7	5.7	98.3
34	56	8.3	5.4	93.3
35	58	8.6	5.6	96.7
37	0	0.0	0.0	0.0
36	59	8.7	5.7	98.3
38	60	8.9	5.8	100.0
39	54	8.0	5.2	90.0
40a	60	8.9	5.8	100.0
40b	2	0.3	0.2	100.0
41	650	96.2	63.2	
42a	596	88.2	57.9	
42b	631	93.3	61.3	
42c	575	85.1	55.9	
42d	578	85.5	56.2	
42e	583	86.2	56.7	
42f	262	38.8	25.5	
42 other	129	19.1	12.5	

cont.

8 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/676)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
43	549	81.2	53.4	
45	672	99.4	65.3	
46	664	98.2	64.5	
47	664	98.2	64.5	
48	665	98.4	64.6	
49	662	97.9	64.3	
50	651	96.3	63.3	
51	660	97.6	64.1	
52	662	97.9	64.3	
53	668	98.8	64.9	
54	665	98.4	64.6	
55	659	97.5	64.0	
56	664	98.2	64.5	
57	668	98.8	64.9	
58	666	98.5	64.7	
59	669	99.0	65.0	
60	653	96.6	63.5	
61	659	97.5	64.0	
62	662	97.9	64.3	
63	657	97.2	63.8	

cont.

8 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/676)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
64	661	97.8	64.2	
65	659	97.5	64.0	
66	664	98.2	64.5	
67	664	98.2	64.5	
68	664	98.2	64.5	
69	665	98.4	64.6	
70	666	98.5	64.7	
71	664	98.2	64.5	
72	665	98.4	64.6	
73	662	97.9	64.3	
74	658	97.3	63.9	
75	656	97.0	63.8	
76	657	97.2	63.8	
77	653	96.6	63.5	
78	665	98.4	64.6	
79	663	98.1	64.4	
80	659	97.5	64.0	
81	663	98.1	64.4	
82	662	97.9	64.3	
83	647	95.7	62.9	

cont.

8 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/676)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
84	665	98.4	64.6	
85	661	97.8	64.2	
86	663	98.1	64.4	
87	660	97.6	64.1	
88	660	97.6	64.1	
89	663	98.1	64.4	
90	665	98.4	64.6	
91	663	98.1	64.4	
92	657	97.2	63.8	
93	659	97.5	64.0	
94	660	97.6	64.1	
95	660	97.6	64.1	
96	660	97.6	64.1	
97	667	98.7	64.8	
98	664	98.2	64.5	
99	659	97.5	64.0	
100	660	97.6	64.1	
101	662	97.9	64.3	
102	664	98.2	64.5	
103	666	98.5	64.7	

cont.

8 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/676)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
104	666	98.5	64.7	
105	666	98.5	64.7	
106	665	98.4	64.6	
107	665	98.4	64.6	
108	665	98.4	64.6	
109	663	98.1	64.4	
110	658	97.3	63.9	
111	662	97.9	64.3	
112	661	97.8	64.2	
113	663	98.1	64.4	
114	662	97.9	64.3	
115	648	95.9	63.0	
116	665	98.4	64.6	
117	659	97.5	64.0	
118	653	96.6	63.5	
119	656	97.0	63.8	
120	659	97.5	64.0	
121	658	97.3	63.9	
122	659	97.5	64.0	
123	657	97.2	63.8	

cont.

8 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/676)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
124	657	97.2	63.8	
125	658	97.3	63.9	
126	657	97.2	63.8	
127	655	96.9	63.7	
128	653	96.6	63.5	
129	655	96.9	63.7	
130	657	97.2	63.8	
131	654	96.7	63.6	
132	642	95.0	62.4	
133	656	97.0	63.8	
134	217	32.1	21.1	

Table C.5. Item Non-Response for 12 Month Questionnaire

12 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/633)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
1	610	96.4	59.3	
2 breast	43	6.8	4.2	
2 formula	347	54.8	33.7	
2 cows milk	372	58.8	36.2	
2 none	8	1.3	0.8	
2 other	6	0.9	0.6	
Type of Milk	623	98.4	60.5	
3	94	14.8	9.1	98.9
4	83	13.1	8.1	87.4
5	151	23.9	14.7	96.8
6a	154	24.3	15.0	98.7
6b	144	22.7	14.0	94.7
7	149	23.5	14.5	95.5
8	147	23.2	14.3	94.2
9a	454	71.7	44.1	
9b	466	73.6	45.3	
9c	338	53.4	32.8	
9d	605	95.6	58.8	
10	616	97.3	59.9	

cont.

12 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/633)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
11	629	99.4	61.1	
12	610	96.4	59.3	
13	607	95.9	59.0	
14	627	99.1	60.9	
15	631	99.7	61.3	
16	630	99.5	61.2	
17a	631	99.7	61.3	
17b	631	99.7	61.3	
17c	629	99.4	61.1	
17d	625	98.7	60.7	
17e	629	99.4	61.1	
17f	629	99.4	61.1	
17g	630	99.5	61.2	
17h	628	99.2	61.0	
17i	627	99.1	60.9	

cont.

12 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/633)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
18a	629	99.4	61.1	
18b	629	99.4	61.1	
18c	631	99.7	61.3	
18d	625	98.7	60.7	
18e	622	98.3	60.4	
18f	627	99.1	60.9	
18g	628	99.2	61.0	
18h	626	98.9	60.8	
18i	624	98.6	60.6	
19a	621	98.1	60.3	
19b	592	93.5	57.5	
19c	619	97.8	60.2	
20a	617	97.5	60.0	
20b	621	98.1	60.3	
20c	615	97.2	59.8	
21	632	99.8	61.4	
22	631	99.7	61.3	
23	631	99.7	61.3	
24	628	99.2	61.0	
25	631	99.7	61.3	

cont.

12 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/633)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
26	623	98.4	60.5	
27	632	99.8	61.4	
29a	628	99.2	61.0	
29b	628	99.2	61.0	
29c	628	99.2	61.0	
29d	630	99.5	61.2	
30a	627	99.1	60.9	
30b	627	99.1	60.9	
30c	628	99.2	61.0	
30d	625	98.7	60.7	
31a	625	98.7	60.7	
31b	630	99.5	61.2	
31c	631	99.7	61.3	
32	631	99.7	61.3	
33	627	99.1	60.9	
34a	613	96.8	59.6	
34b	611	96.5	59.4	
34c	608	96.1	59.1	
34d	614	97.0	59.7	

cont.

12 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/633)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
35a	629	99.4	61.1	
35b	628	99.2	61.0	
35c	630	99.5	61.2	
35d	629	99.4	61.1	
36	625	98.7	60.7	
37	631	99.7	61.3	
38a	631	99.7	61.3	
38b	632	99.8	61.4	
38c	632	99.8	61.4	
38d	627	99.1	60.9	
39	621	98.1	60.3	
40a	593	93.7	57.6	
40b	613	96.8	59.6	
40c	594	93.8	57.7	
40d	574	90.7	55.8	
40e	570	90.0	55.4	
40f	625	98.7	60.7	
41	0	0.0	0.0	

cont.

12 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/633)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
43	618	97.6	60.1	
44	63	10.0	6.1	96.9
45	63	10.0	6.1	96.9
46	64	10.1	6.2	98.5
48	0	0.0	0.0	0.0
47	65	10.3	6.3	100.0
49	65	10.3	6.3	100.0
50	61	9.6	5.9	93.8
51a	65	10.3	6.3	100.0
51b	0	0.0	0.0	0.0
52	626	98.9	60.8	
53	626	98.9	60.8	
54	627	99.1	60.9	
55	626	98.9	60.8	
56	626	98.9	60.8	
57	629	99.4	61.1	
58	626	98.9	60.8	
59	620	97.9	60.3	
60	627	99.1	60.9	
61	625	98.7	60.7	

cont.

12 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/633)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
62	625	98.7	60.7	
63	628	99.2	61.0	
64	629	99.4	61.1	
65	629	99.4	61.1	
66	626	98.9	60.8	
67	629	99.4	61.1	
68	627	99.1	60.9	
69	627	99.1	60.9	
70	626	98.9	60.8	
71	627	99.1	60.9	
72	627	99.1	60.9	
73	625	98.7	60.7	
74	627	99.1	60.9	
75	626	98.9	60.8	
76	627	99.1	60.9	
77	625	98.7	60.7	
78	627	99.1	60.9	
79	622	98.3	60.4	
80	622	98.3	60.4	
81	627	99.1	60.9	

cont.

12 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/633)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
82	623	98.4	60.5	
83	626	98.9	60.8	
84	626	98.9	60.8	
85	603	95.3	58.6	99.7
86	603	95.3	58.6	99.7
87	600	94.8	58.3	99.2
88	600	94.8	58.3	99.2
89	599	94.6	58.2	99.0
90a	621	98.1	60.3	
90b	620	97.9	60.3	
90c	612	96.7	59.5	
90d	615	97.2	59.8	
90e	621	98.1	60.3	
90f	601	94.9	58.4	
90g	22	3.5	2.1	
91	618	97.6	60.1	
92a	50	7.9	4.9	98.0
92b	51	8.1	5.0	100.0
92c	51	8.1	5.0	100.0
92d	51	8.1	5.0	100.0

cont.

12 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/633)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
93a	624	98.6	60.6	
93b	491	77.6	47.7	
94	613	96.8	59.6	
95a	601	94.9	58.4	
95b	588	92.9	57.1	
95c	598	94.5	58.1	
96	610	96.4	59.3	
97	615	97.2	59.8	
98	609	96.2	59.2	
99	611	96.5	59.4	
100a	616	97.3	59.9	
100b	610	96.4	59.3	
100c	608	96.1	59.1	
100d	611	96.5	59.4	
101	186	29.4	18.1	
102a	607	95.9	59.0	
102b	600	94.8	58.3	
102c	603	95.3	58.6	
103	614	97.0	59.7	
104	610	96.4	59.3	

cont.

12 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/633)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
105	622	98.3	60.4	
106	555	87.7	53.9	
107	616	97.3	59.9	
108	623	98.4	60.5	
109	619	97.8	60.2	
110	621	98.1	60.3	
111	618	97.6	60.1	
112a	623	98.4	60.5	
112b	622	98.3	60.4	
112c	621	98.1	60.3	
112d	622	98.3	60.4	
113	619	97.8	60.2	
114a	621	98.1	60.3	
114b	615	97.2	59.8	
114c	613	96.8	59.6	
114d	619	97.8	60.2	
114e	615	97.2	59.8	
114f	620	97.9	60.3	
114g	617	97.5	60.0	
114h	620	97.9	60.3	

cont.

12 month Questionnaire Questions	**Qu. answered by respondents to qu're**	**% of Respondents (/633)**	**% of Recruits (/1029)**	**Conditional Response Rate (%)**
114i	618	97.6	60.1	
114j	617	97.5	60.0	
114k	620	97.9	60.3	
114l	611	96.5	59.4	
114m	605	95.6	58.8	
114n	600	94.8	58.3	
115	262	41.4	25.5	
116	119	18.8	11.6	
117	116	18.3	11.3	

Appendix D

Different Analyses Performed

D.1 EM Algorithm

D.1.1 ANOVA for Linear Trend for TI0-12m ∼ 6 Week Appetite Rate

Table D.1. Imputing 12 Month Weight z-scores using Birthweight z-scores

	No. of Cases Included	F-Statistic	P-Value
Analysis 1	678	12.14	0.0005
Analysis 2	749	11.7	0.0007
Analysis 3.1	774	14.16	0.0002
Analysis 3.2	774	10.59	0.0012
Analysis 3.3	774	9.73	0.0019
Analysis 4.1	923	14.37	0.0002
Analysis 4.2	923	10.33	0.0014
Analysis 4.3	923	9.95	0.0017

Notes on Table D.1

Analysis 1 is the complete case analysis. **Analysis 2** is the analysis where only wtz12m is imputed using bwtz. **Analysis 3.1** is the analysis where only 6 week appetite is imputed using 4 month appetite. **Analysis 3.2** is the analysis where only 6 week appetite is imputed using 12 month appetite. **Analysis 3.3** is the analysis where only 6 week appetite is imputed using 4 month, 8 month and 12 month appetite. **Analysis 4.1** is where both wtz12m and 6 week appetite using 4 month appetite are imputed. **Analysis 4.2** is where both wtz12m and 6 week appetite using 12 month appetite are imputed. **Analysis 4.3** is where both wtz12m and 6 week appetite using 4 month, 8 month and 12 month appetite are imputed.

Table D.2. Imputing 12 Month Weight z-scores using 8 Month Weight z-scores

	No. of Cases Included	F-Statistic	P-Value
Analysis 1	678	12.14	0.0005
Analysis 2	749	14.77	0.0001
Analysis 3.1	774	14.16	0.0002
Analysis 3.2	774	10.59	0.0012
Analysis 3.3	774	9.73	0.0019
Analysis 4.1	923	15.78	0.0001
Analysis 4.2	923	10.64	0.0011
Analysis 4.3	923	11.56	0.0007

Notes on Table D.2

Analysis 1 is the complete case analysis. **Analysis 2** is the analysis where only wtz12m is imputed using wtz8m. **Analysis 3.1** is the analysis where only 6

week appetite is imputed using 4 month appetite. **Analysis 3.2** is the analysis where only 6 week appetite is imputed using 12 month appetite. **Analysis 3.3** is the analysis where only 6 week appetite is imputed using 4 month, 8 month and 12 month appetite. **Analysis 4.1** is where both wtz12m and 6 week appetite using 4 month appetite are imputed. **Analysis 4.2** is where both wtz12m and 6 week appetite using 12 month appetite are imputed. **Analysis 4.3** is where both wtz12m and 6 week appetite using 4 month, 8 month and 12 month appetite are imputed.

Table D.3. Imputing 12 Month Weight z-scores using 8 Month, 4 Month and 6 Week Weight z-scores

	No. of Cases Included	F-Statistic	P-Value
Analysis 1	678	12.14	0.0005
Analysis 2	749	16.87	0.0001
Analysis 3.1	774	14.16	0.0002
Analysis 3.2	774	10.59	0.0012
Analysis 3.3	774	9.73	0.0019
Analysis 4.1	923	16.35	0.0001
Analysis 4.2	923	17.23	0.0001
Analysis 4.3	923	14.35	0.0002

Notes on Table D.3

Analysis 1 is the complete case analysis. **Analysis 2** is the analysis where only wtz12m is imputed using wtz8m, wtz4m and wtz6wk. **Analysis 3.1** is the analysis where only 6 week appetite is imputed using 4 month appetite. **Analysis 3.2** is the analysis where only 6 week appetite is imputed using 12 month appetite. **Analysis 3.3** is the analysis where only 6 week appetite is imputed using 4

month, 8 month and 12 month appetite. **Analysis 4.1** is where both wtz12m and 6 week appetite using 4 month appetite are imputed. **Analysis 4.2** is where both wtz12m and 6 week appetite using 12 month appetite are imputed. **Analysis 4.3** is where both wtz12m and 6 week appetite using 4 month, 8 month and 12 month appetite are imputed.

D.1.2 ANOVA for Linear Trend for TI0-12m ∼ 12 Month Appetite Rate

Table D.4. Imputing 12 Month Weight z-scores using Birthweight z-scores

	No. of Cases Included	F-Statistic	P-Value
Analysis 1	569	12.99	0.0003
Analysis 2	578	12.91	0.0004
Analysis 3.1	774	12.29	0.0005
Analysis 3.2	774	7.94	0.005
Analysis 3.3	774	7.07	0.008
Analysis 4.1	923	12.26	0.0005
Analysis 4.2	923	7.96	0.0049
Analysis 4.3	923	6.9	0.0088

Notes on Table D.4

Analysis 1 is the complete case analysis. **Analysis 2** is the analysis where only wtz12m is imputed using bwtz. **Analysis 3.1** is the analysis where only 12 month appetite is imputed using 6 week appetite. **Analysis 3.2** is the analysis where only 12 month appetite is imputed using 8 month appetite. **Analysis 3.3** is the analysis where only 12 month appetite is imputed using 6 week, 4 month and

8 month appetite. **Analysis 4.1** is where both wtz12m and 12 month appetite using 6 week appetite are imputed. **Analysis 4.2** is where both wtz12m and 12 month appetite using 8 month appetite are imputed. **Analysis 4.3** is where both wtz12m and 12 month appetite using 6 week, 4 month and 8 month appetite are imputed.

Table D.5. Imputing 12 Month Weight z-scores using 8 Month Weight z-scores

	No. of Cases Included	F-Statistic	P-Value
Analysis 1	569	12.99	0.0003
Analysis 2	578	11.4	0.001
Analysis 3.1	774	12.29	0.0005
Analysis 3.2	774	7.94	0.005
Analysis 3.3	774	7.07	0.008
Analysis 4.1	923	6.74	0.0096
Analysis 4.2	923	3.7	0.0547
Analysis 4.3	923	7.48	0.0064

Notes on Table D.5

Analysis 1 is the complete case analysis. **Analysis 2** is the analysis where only wtz12m is imputed using wtz8m. **Analysis 3.1** is the analysis where only 12 month appetite is imputed using 6 week appetite. **Analysis 3.2** is the analysis where only 12 month appetite is imputed using 8 month appetite. **Analysis 3.3** is the analysis where only 12 month appetite is imputed using 6 week, 4 month and 8 month appetite. **Analysis 4.1** is where both wtz12m and 12 month appetite using 6 week appetite are imputed. **Analysis 4.2** is where both wtz12m and 12 month appetite using 8 month appetite are imputed. **Analysis 4.3** is where both

wtz12m and 12 month appetite using 6 week, 4 month and 8 month appetite are imputed.

Table D.6. Imputing 12 Month Weight z-scores using 8 Month Weight z-scores

	No. of Cases Included	F-Statistic	P-Value
Analysis 1	569	12.99	0.0003
Analysis 2	578	12.96	0.0003
Analysis 3.1	774	12.29	0.0005
Analysis 3.2	774	7.94	0.005
Analysis 3.3	774	7.07	0.008
Analysis 4.1	923	12.3	0.0005
Analysis 4.2	923	7.73	0.0055
Analysis 4.3	923	6.72	0.0097

Notes on Table D.6

Analysis 1 is the complete case analysis. **Analysis 2** is the analysis where only wtz12m is imputed using wtz8m, wtz4m and wtz6wk. **Analysis 3.1** is the analysis where only 12 month appetite is imputed using 6 week appetite. **Analysis 3.2** is the analysis where only 12 month appetite is imputed using 8 month appetite. **Analysis 3.3** is the analysis where only 12 month appetite is imputed using 6 week, 4 month and 8 month appetite. **Analysis 4.1** is where both wtz12m and 12 month appetite using 6 week appetite are imputed. **Analysis 4.2** is where both wtz12m and 12 month appetite using 8 month appetite are imputed. **Analysis 4.3** is where both wtz12m and 12 month appetite using 6 week, 4 month and 8 month appetite are imputed.

D.2 Single Hot Deck Imputation

D.2.1 ANOVA for Linear Trend for TI0-12m ∼ 6 Week Appetite Rate

Table D.7. Imputing 12 Month Weight z-scores using Birthweight z-scores

	No. of Cases Included	F-Statistic	P-Value
Analysis 1	678	11.55	0.0007
Analysis 2	749	12.99	0.0003
Analysis 3	766	12.04	0.0005
Analysis 4	912	11.17	0.0009

Notes on Table D.7

Analysis 1 is the complete case analysis. **Analysis 2** is the analysis where only wtz12m is imputed using bwtz. **Analysis 3** is the analysis where only 6 week appetite is imputed using 4 month appetite. **Analysis 4** is where both wtz12m and 6 week appetite are imputed.

Table D.8. Imputing 12 Month Weight z-scores using 8 Month Weight z-scores

	No. of Cases Included	F-Statistic	P-Value
Analysis 1	678	11.55	0.0007
Analysis 2	695	12.93	0.0003
Analysis 3	766	12.04	0.0005
Analysis 4	787	13.39	0.0003

Notes on Table D.8

Analysis 1 is the complete case analysis. **Analysis 2** is the analysis where only wtz12m is imputed using wtz8m. **Analysis 3** is the analysis where only 6 week appetite is imputed using 4 month appetite. **Analysis 4** is where both wtz12m and 6 week appetite are imputed.

Table D.9. Imputing 12 Month Weight z-scores using 8 Month, 4 Month and 6 Week Weight z-scores

	No. of Cases Included	F-Statistic	P-Value
Analysis 1	678	11.55	0.0007
Analysis 2	749	12.22	0.0005
Analysis 3	766	12.04	0.0005
Analysis 4	867	11.54	0.0007

Notes on Table D.9

Analysis 1 is the complete case analysis. **Analysis 2** is the analysis where only wtz12m is imputed using wtz8m, wtz4m and wtz6wk. **Analysis 3** is the analysis where only 6 week appetite is imputed using 4 month appetite. **Analysis 4** is where both wtz12m and 6 week appetite are imputed.

D.2.2 ANOVA for Linear Trend for TI0-12m ∼ 12 Month Appetite Rate

Table D.10. Imputing 12 Month Weight z-scores using Birthweight z-scores

	No. of Cases Included	F-Statistic	P-Value
Analysis 1	569	12.99	0.0003
Analysis 2	578	12.92	0.0004
Analysis 3.1	640	15.2	0.0001
Analysis 3.2	714	12.26	0.0005
Analysis 4.1	668	15.19	0.0001
Analysis 4.2	799	9.85	0.0018

Notes on Table D.10

Analysis 1 is the complete case analysis. **Analysis 2** is the analysis where only wtz12m is imputed using bwtz. **Analysis 3.1** is the analysis where only 12 month appetite is imputed using 8 month appetite. **Analysis 3.2** is the analysis where only 12 month appetite is imputed using 8 month, 4 month and 6 week appetite. **Analysis 4.1** is where both wtz12m and 12 month appetite using 8 month appetite are imputed. **Analysis 4.2** is where both wtz12m and 12 month appetite using 8 month, 4 month and 6 week appetite are imputed.

Table D.11. Imputing 12 Month Weight z-scores using 8 Month Weight z-scores

	No. of Cases Included	F-Statistic	P-Value
Analysis 1	569	12.99	0.0003
Analysis 2	575	11.71	0.0007
Analysis 3.1	640	15.2	0.0001
Analysis 3.2	714	12.26	0.0005
Analysis 4.1	661	14.08	0.0002
Analysis 4.2	735	12.15	0.0005

Notes on Table D.11

Analysis 1 is the complete case analysis. **Analysis 2** is the analysis where only wtz12m is imputed using wtz8m. **Analysis 3.1** is the analysis where only 12 month appetite is imputed using 8 month appetite. **Analysis 3.2** is the analysis where only 12 month appetite is imputed using 8 month, 4 month and 6 week appetite. **Analysis 4.1** is where both wtz12m and 12 month appetite using 8 month appetite are imputed. **Analysis 4.2** is where both wtz12m and 12 month appetite using 8 month, 4 month and 6 week appetite are imputed.

Table D.12. Imputing 12 Month Weight z-scores using 8 Month, 4 Month and 6 Week Weight z-scores

	No. of Cases Included	F-Statistic	P-Value
Analysis 1	569	12.99	0.0003
Analysis 2	578	13.76	0.0002
Analysis 3.1	640	15.2	0.0001
Analysis 3.2	714	12.26	0.0005
Analysis 4.1	668	15.66	0.0001
Analysis 4.2	799	11.2	0.0009

Notes on Table D.12

Analysis 1 is the complete case analysis. **Analysis 2** is the analysis where only wtz12m is imputed using wtz8m, wtz4m and wtz6wk. **Analysis 3.1** is the analysis where only 12 month appetite is imputed using 8 month appetite. **Analysis 3.2** is the analysis where only 12 month appetite is imputed using 8 month, 4 month and 6 week appetite. **Analysis 4.1** is where both wtz12m and 12 month appetite using 8 month appetite are imputed. **Analysis 4.2** is where both wtz12m and 12 month appetite using 8 month, 4 month and 6 week appetite are imputed.

D.3 Multiple Hot Deck Imputation

D.3.1 ANOVA for Linear Trend for TI0-12m $\sim$ 6 Week Appetite Rate

Table D.13. Imputing 12 Month Weight z-scores using Birthweight z-scores

	No. of Cases Included	F-Statistic	P-Value
Analysis 1	678	12.14	0.0005
Analysis 2	749	10.24	0.0022
Analysis 3	766	10.42	0.0022
Analysis 4	912	8.73	0.0071

Notes on Table D.13

Analysis 1 is the complete case analysis. **Analysis 2** is the analysis where only wtz12m is imputed using bwtz. **Analysis 3** is the analysis where only 6 week appetite is imputed using 4 month appetite. **Analysis 4** is where both wtz12m and 6 week appetite are imputed.

Table D.14. Imputing 12 Month Weight z-scores using 8 Month Weight z-scores

	No. of Cases Included	F-Statistic	P-Value
Analysis 1	678	12.14	0.0005
Analysis 2	695	13.13	0.0003
Analysis 3	766	10.42	0.0022
Analysis 4	787	11.91	0.0011

Notes on Table D.14

Analysis 1 is the complete case analysis. **Analysis 2** is the analysis where only wtz12m is imputed using wtz8m. **Analysis 3** is the analysis where only 6 week appetite is imputed using 4 month appetite. **Analysis 4** is where both wtz12m and 6 week appetite are imputed.

Table D.15. Imputing 12 Month Weight z-scores using 8 Month, 4 Month and 6 Week Weight z-scores

	No. of Cases Included	F-Statistic	P-Value
Analysis 1	678	12.14	0.0005
Analysis 2	749	13.31	0.0004
Analysis 3	766	10.42	0.0022
Analysis 4	867	12.17	0.001

Notes on Table D.15

Analysis 1 is the complete case analysis. **Analysis 2** is the analysis where only wtz12m is imputed using wtz8m, wtz4m and wtz6wk. **Analysis 3** is the analysis where only 6 week appetite is imputed using 4 month appetite. **Analysis 4** is where both wtz12m and 6 week appetite are imputed.

D.3.2 ANOVA for Linear Trend for TI0-12m ∼ 12 Month Appetite Rate

Table D.16. Imputing 12 Month Weight z-scores using Birthweight z-scores

	No. of Cases Included	F-Statistic	P-Value
Analysis 1	569	12.99	0.0003
Analysis 2	578	12.41	0.0006
Analysis 3.1	640	12.45	0.0008
Analysis 3.2	714	14.91	0.0002
Analysis 4.1	668	11.25	0.0018
Analysis 4.2	799	10.48	0.0034

Notes on Table D.16

Analysis 1 is the complete case analysis. **Analysis 2** is the analysis where only wtz12m is imputed using bwtz. **Analysis 3.1** is the analysis where only 12 month appetite is imputed using 8 month appetite. **Analysis 3.2** is the analysis where only 12 month appetite is imputed using 8 month, 4 month and 6 week appetite. **Analysis 4.1** is where both wtz12m and 12 month appetite using 8 month appetite are imputed. **Analysis 4.2** is where both wtz12m and 12 month appetite using 8 month, 4 month and 6 week appetite are imputed.

Table D.17. Imputing 12 Month Weight z-scores using 8 Month Weight z-scores

	No. of Cases Included	F-Statistic	P-Value
Analysis 1	569	12.99	0.0003
Analysis 2	575	11.86	0.0006
Analysis 3.1	640	12.45	0.0008
Analysis 3.2	714	14.91	0.0002
Analysis 4.1	661	11.4	0.0016
Analysis 4.2	735	14.41	0.0003

Notes on Table D.17

Analysis 1 is the complete case analysis. **Analysis 2** is the analysis where only wtz12m is imputed using wtz8m. **Analysis 3.1** is the analysis where only 12 month appetite is imputed using 8 month appetite. **Analysis 3.2** is the analysis where only 12 month appetite is imputed using 8 month, 4 month and 6 week appetite. **Analysis 4.1** is where both wtz12m and 12 month appetite using 8 month appetite are imputed. **Analysis 4.2** is where both wtz12m and 12 month appetite using 8 month, 4 month and 6 week appetite are imputed.

Table D.18. Imputing 12 Month Weight z-scores using 8 Month, 4 Month and 6 Week Weight z-scores

	No. of Cases Included	F-Statistic	P-Value
Analysis 1	569	12.99	0.0003
Analysis 2	578	12.78	0.0004
Analysis 3.1	640	12.45	0.0008
Analysis 3.2	714	14.91	0.0002
Analysis 4.1	668	11.81	0.0011
Analysis 4.2	799	14.65	0.0021

Notes on Table D.18

Analysis 1 is the complete case analysis. **Analysis 2** is the analysis where only wtz12m is imputed using wtz8m, wtz4m and wtz6wk. **Analysis 3.1** is the analysis where only 12 month appetite is imputed using 8 month appetite. **Analysis 3.2** is the analysis where only 12 month appetite is imputed using 8 month, 4 month and 6 week appetite. **Analysis 4.1** is where both wtz12m and 12 month appetite using 8 month appetite are imputed. **Analysis 4.2** is where both wtz12m and 12 month appetite using 8 month, 4 month and 6 week appetite are imputed.

Bibliography

Altman, D. G. (1991). *Practical Statistics for Medical Research*. Chapman and Hall.

Carpenter, J. and M. Kenward (2005, March). Chained equations. http://missingdata.lshtm.ac.uk/mi_web/node14.html.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological) 39*(1), 1–38.

Diggle, P., D. Farewell, and R. Henderson (2007). Analysis of longitudinal data with drop-out: Objectives, assumptions and a proposal. *Journal of the Royal Statistical Society. Series C (Applied Statistics) 56*, 499–550.

Diggle, P. J., P. Heagerty, K.-Y. Liang, and S. L. Zeger (2002). *Analysis of Longitudinal Data* (2nd ed.). Oxford University Press Inc.

Freeman, J., T. Cole, S. Chinn, P. Jones, E. White, and M. Preece (1995). Cross sectional stature and weight reference curves for the uk, 1990. *Archives of Disease in Childhood 73*(1), 17–24.

Harrell, F. E. (2001). *Regression Modelling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. Springer Science + Business Media, Inc.

Harrell, F. E. and with contributions from many other users (2007). *Hmisc: Harrell Miscellaneous*. R package version 3.4-2.

Little, R. J. A. (1995). Modelling the drop-out mechanism in repeated measures studies. *Journal of the American Statistical Association 90*(431), 1112–1121.

Little, R. J. A. and D. B. Rubin (2002). *Statistical Analysis with Missing Data* (2nd ed.). John Wiley and Sons, Inc.

McLachlan, G. J. and T. Krishnan (1997). *The EM Algorithm and Extensions*. John Wiley and Sons Ltd.

Meng, X.-L. and D. B. Rubin (1991). Using the em to obtain asymptotic variance-covariance matrices: The sem algorithm. *Journal of the American Statistical Association 86*(416), 899–909.

Molenberghs, G. and M. G. Kenward (2007). *Missing Data in Clinical Studies*. John Wiley and Sons Ltd.

Ported to R by Alvaro A. Novo. Original by Joseph L. Schafer, jls@stat.psu.edu (2002). *norm: Analysis of multivariate normal datasets with missing values*. R package version 1.0-9.

Ported to R by Ted Harding and Fernando Tusell. Original by Joseph L. Schafer (2004). *cat: Analysis of categorical-variable datasets with missing values*. R package version 0.0-6.

R Development Core Team (2010). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

Rao, J. N. K. and J. Shao (1992). Jackknife variance estimation with survey data under hot deck imputation. *Biometrika 79*(4), 811–822.

Rubin, D. B. (1976). Inference and missing data. *Biometrika 63*(3), 581–592.

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. John Wiley and Sons, Inc.

Rubin, D. B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association 91*(434), 473–489.

Schafer, J. L. (1997). *Analysis of Incomplete Multivariate Data*. Chapman and Hall.

Wright, C. M., K. N. Parkinson, and R. F. Drewett (2006a). How does maternal and child feeding behviour relate to weight gain and failure to thrive? data from a prospective birth cohort. *Pediatrics 117*(4), 1262–1269.

Wright, C. M., K. N. Parkinson, and R. F. Drewett (2006b). The influence of maternal socioeconomic and emotional factors on infant weight gain and weight faltering (failure to thrive): Data from a prospective birth cohort. *Archives of Disease in Childhood 91*(4), 312–317.