Using longitudinal measurements to identify undernutrition – a statistical investigation

Fraser Tough

A Dissertation Submitted to the

University of Glasgow

for the degree of

Doctor of Philosophy

School of Medicine, Dentistry and Nursing

December 2016

Abstract

Understanding the ways in which practitioners can identify and manage undernutrition is important within developing world countries. There is still much uncertainty when it comes to understanding which measures of undernutrition are the most effective predictors of adverse outcomes. This thesis explores how children grow and applies statistical methodology to three longitudinal growth datasets with frequent measurements in the first two years, seeking new insights into how measures of undernutrition can be used to predict future adverse outcomes. The three datasets are diverse - from Malawi, South Africa and Pakistan, the latter of which contains 4 subsets of different socioeconomic groups. The large number of children within the sets made it possible to test several different hypotheses.

Assessing growth of children within the developing world compared to the WHO standard

Growth charts (or reference charts) are charts which allow practitioners to compare a given infant's anthropometric measurements with a reference population. We developed growth charts from the available datasets using Generalised Additive Models for Location Scale and Shape (GAMLSS), a method which allows users to flexibly model distributions of measurements over time.

The reference charts we developed describe the growth of samples of children, many of whom will not have grown at a healthy rate. It is preferable to compare children with healthy infants from a composite external standard. The World Health Organization (WHO) growth standard was developed from a variety of populations from across the globe which describes the growth of a 'healthy' population. This suggests an aspirational model, as opposed to a reference, which describes how a sample of children actually grow. In this thesis GAMLSS was used to determine whether real populations of pre-school children from the developing world fit this international standard. We found that relatively affluent populations fit the standard well, or even outperform it, while more deprived populations fall away to varying degrees, then mainly track parallel to the WHO mean beyond 6 months. This suggests that after the first 6 months children from the developing world have rates of weight gain roughly on par with the standard, although the children are much lighter.

Assessing weight gain while conditioning on initial weight

Plotting measurements on growth charts identifies those whose weight Z score or centile is falling relative to the reference. However, children initially at the extremes tend to regress toward the mean. Conditional weight gain (CWG) takes this expected movement into account, but can only be used within the population in which the child originates, due to certain statistical assumptions. We developed a generalised measure of CWG for use with the WHO external standard. This measure requires the correlation between pairs of groups of measurements at different time points, as the amount of regression to the mean is synonymous with this correlation.

If data are not available at these time points, they can be interpolated by firstly computing correlations between all available data, then modelling the resulting matrix. We found that these correlation matrices are heterogeneous within the developing world. Therefore, constructing a generalised correlation model was not possible. This makes the use of the new generalised measure of CWG impractical without access to correlation models computed from local data. However, the measure may be useful within the developed world, where correlation matrices may be less variable.

Pathways through nutritional states

The analysis then explored the ways that children move between different nutritional states, defined as healthy, thin (wasted) and/or short (stunted), over 3-6 month (m), 6-9m and 9-12m timeframes, and the probability these states will lead to death. We used stochastic models to explore the probability of moving state conditional on previous state, exploring the pathways children take through different states over time.

1.) **What states are likely to lead to death?**

Within all timeframes, children who were wasted as well as stunted were more likely to die than wasted children, who were in turn more likely to die than stunted children. Furthermore, as children age, the conditional risk of death in the next time period decreases. However, relative to healthy children, all children were less vulnerable within the middle period (6-9m) regardless of state.

2.) **Does wasting lead to stunting?**

Children who were wasted were at significantly higher risk than healthy children of later wasting, or becoming stunted as well as wasted, over all timeframes.

However, wasting alone significantly increased the risk of later stunting only in the 3-6m timeframe.

3.) **Does wasting usually precede stunting?**

Across the 3-6-9m timeframes children were much more likely to move from either healthy → healthy → stunted, or healthy → stunted → stunted, than from healthy → wasted → stunted. This indicates children are more likely to move directly into a stunted state than from healthy to stunted via wasted.

Is growth or size the best predictor of mortality?

Change in weight (growth) has been shown to be a predictor of mortality in populations of children, but it is not clear if this measure is more predictive than the latest weight (size). Using weighted Cox proportional hazards models, we determined which of these measures is the most valuable predictor of mortality for the majority of children within each individual dataset, conducting analyses using variable levels of weightings for children at the extremes. We included weight-for-age and height-for-age as predictors within our models to determine what combination of predictors best predict mortality. In all unweighted analyses, size was the best predictor of time until death. However, as the weighting increased, growth entered as the best predictor in populations with low rates of undernutrition. In contrast, size always remained the strongest predictor within populations with high rates of undernutrition, since in these populations, such a high proportion of children fall away from within the centre of the normal range, making growth pattern non-discriminating.

Summary

This programme of work applied statistical techniques to three diverse longitudinal datasets, gaining insights into how children grow between different socio-economic backgrounds. We investigated measures of size and measures of growth, utilising methods that control for the inevitable fact that healthy children at the centre of the population distribution tend to dominate analyses. Furthermore, these methods were both multidimensional and time dependant, providing us with a useful framework to assess child growth while controlling for influential factors. The results should improve understanding of both the aetiology of undernutrition and its clinical management.

Acknowledgements

First and foremost I would like to thank Professors John McColl and Charlotte Wright for their guidance throughout both the analysis and write up stages of this thesis. Charlotte and John were always helpful and supportive, which made this work a rewarding learning experience. On a personal level, I could have not asked for two more approachable, friendly people. I thank them both for their continued time, effort and patience. I also thank Professor Tim Cole, whose expertise in the field has paved the way for much of this work, and who has supported much of the analysis throughout.

Without the data made available, this thesis this work would have not been possible. I would therefore like to thank data providers Dr. Ruth Bland, Professor Per Ashorn and Professor Shakila Zaman. Financial support from the Medical Research Council is gratefully acknowledged.

Study at the University of Glasgow has been a large part of my life thus far. Special thanks go to the staff at the School of Mathematics and Statistics who have made my years at the university an unforgettable experience. Without the support of the university it would not have been possible for me to further my experience through postgraduate degrees in statistics.

Finally I would like to thank friends and family, especially my parents, whose direction, patience and support have helped me more than I can express in words.

Contents

List of figures

List of tables

List of abbreviations

AIC	Akaike Information Criterion
AIDS	Acquired immune deficiency syndrome
AM	Additive Model
ARI	Acute Respiratory Infection
BCCG	Box-Cox Cole Green
BCPE	Box-Cox power exponential
BIC	Bayesian Information Criterion
BMI	Body mass index
CDC	Centers for Disease Control and Prevention
D	Dead
DRC	Democratic Republic of the Congo
FP	Fractional polynomial
FTT	Failure to thrive
GAM	Generalised Additive Model
GAMLSS	Generalised Additive Models for Location Scale and Shape
GCV	Generalised Cross-Validation
GLM	Generalised Linear Model
GLRT	Generalised Likelihood Ratio Test
H	Healthy
HAZ	Height-for-age Z score
HIV	Human immunodeficiency virus
HR	Hazard Ratio
LBW	Low birth weight
LCSS	Lungwena Child Survival Study
LLS	Lahore Longitudinal Study
MAM	Moderate Acute Malnutrition
MGRS	Multicentre Growth Reference Study
MSE	Mean square error
MSS	Maximum sum of sensitivity and specificity
MUAC	Mid-Upper Arm Circumference
NCHS	National Center for Health Statistics
ND	Normalised distance
PDF	Probability distribution function
RMS	Root Mean Square
RR	Relative Risk
RTTM	Regression to the mean
S	Stunted
SAM	Severe Acute Malnutrition
SD	Standard deviation
U	Underweight
UNICEF	United Nations Children's Emergency Fund
VTS	Vertical Transmission Study
W	Wasted
WS	Wasted-stunted
WAZ	Weight-for-age Z score
WHO	World Health Organization
WHZ	Weight-for-height Z score

Chapter 1
Introduction
1.1 Preamble

The primary aim of this thesis is to utilise current statistical methods available in the scientific literature, applying them to longitudinal growth datasets in novel ways, allowing us to contribute new insights to the understanding of how children grow by quantifying which measures of undernutrition are useful predictors of future outcomes.

Single measurements of attained size and the change in size between two time points (growth) are frequently used in paediatrics and are often good proxy measures of both the environment children live in and their current state of health. These measurements are usually compared with reference charts developed from populations of children which closely match in terms of sex and age. Previous research has identified that the most recent measurements of size and measures of growth over a recent time period relative to a reference population can both be used to predict future survival. However, clinicians generally assume that measures of growth over time are superior to the most recent measurement, size, as they supply more information about the child's nutritional history. This thesis contributes to the existing body of knowledge by investigating which of these measures is actually the best predictor of future outcomes.

While continuous measures of size and growth have been shown to be predictors of mortality, categorical indicators of undernutrition (also known as nutritional states) such as wasting (thin/low weight-for-height) and/or stunting (short/low height-for-age) relative to the reference population have also been shown to be predictive of mortality.

Furthermore, research suggests that children tend to move from certain states to others over time. For example, wasting is thought to precede stunting and stunting is thought to be irreversible. We investigate the extent to which these presumed transitions are true. Conditional weight gain has been proposed as a more robust alternative for growth velocity as it accounts for the fact that children initially at the extremes are expected to regress toward the centre of the distribution over time. However, the assessment of weight gain depends on the use of a suitable, well matched reference, developed from healthy children. The current expression for conditional weight gain only allows the comparison of an individual child's measurements with the distribution of measurements in the population from which the child originates. Preferably, the comparison should be made with children from an idealised, healthy population of children. We extend the statistical methodology so that children can be compared with 'external' reference distributions, such as the World Health Organization (WHO) standard, creating a generalised measure of conditional weight gain.

This introductory chapter reviews literature which has previously investigated how anthropometric measurements can be used to predict outcomes in the future. Reviewing this work provides an initial understanding of how children in the developing world grow and allows us to develop hypotheses based on previous research. We then test these hypotheses in later chapters by utilising statistical methods which have not yet been applied in the field. The literature and hypotheses can be found in Section 1.3. However, we first present a brief review on how healthy children grow and review undernutrition's effects on children. This provides some background in growth before reviewing previous literature. We also introduce growth monitoring used for the diagnosis of undernutrition, the WHO standard and provide a rationale as to why conditional weight gain should be adapted for use with this standard.

1.2 Review of undernutrition

1.2.1 How do healthy children gain weight and grow?

The first year of a child's life is the fastest they will ever grow. In the first few days after birth an infant loses around 5-10% of their overall bodyweight, but after around 5-7 days they start to gain weight and grow quickly (Wright & Parkinson 2004). From 4-6 months, an infant's weight should be around double its birth weight (Krieger 1970). During months 6-12, weight gain is not as rapid as in earlier life and growth in height slows considerably.

The expected height by age 1 year should be around 1.5 times birth height and the expected weight should be around 3 times birth weight. However, during months 12-24 a child will only gain around 1kg. From age 2, children's weight will increase at a steady pace until puberty (Kliegman 2012), generally around 3.2kg per year and around 6.3cm per year from years 6 to 12 (Beaumont Health System 2015).

When children reach puberty they go through a spurt in growth although not all children reach puberty at the same age. It is hormones released during puberty that control growth and spark physical changes such as breast development in girls and testicular enlargement in boys. In puberty, height velocities (the rate of growth) increase, usually from around 9 years old for girls and 11 years old for boys. Peak height velocity occurs at a mean age of 13.5 years for boys and 11.5 years for girls, where the magnitude of the growth spurt is negatively correlated with age. This spurt in height accounts for around 20% of final adult height - an average of 25.5cm of females and 27cm in males (Neinstein 2002).

Weight gain in puberty accounts for around 50% of an adult's ideal body weight, but the onset of accelerated weight gain within puberty and peak weight velocity are highly variable. Weight velocity at its peak can range from 4.6kg/yr to 10.6kg/yr for girls, and 5.7kg/yr to 13.2kg/yr for boys (Neinstein 2002).

1.2.2 Introduction to undernutrition

Undernutrition is the insufficient or unbalanced consumption or absorption of nutrients (Black et al. 2008). Undernutrition is most prevalent in low and middle income countries, resulting in increases in mortality and overall disease burden. It is one of the main underlying causes of illness and death in African women and children, contributing to more than half the deaths among under five year olds on the continent (Black et al. 2008) and one third of all deaths worldwide (Relman 2013). The estimated proportion of those undernourished around the world between 2014 and 2016 was 10.9%

Undernutrition is one type of malnutrition, where the word malnutrition is used to cover undernutrition, overnutrition and specific deficiencies. The body requires essential nutrients including vitamins and minerals, as well as the energy to cover energy expenditure throughout a normal day. Children also need energy to maintain a healthy immune system and, in the case of children, the energy to grow. If the body does not have sufficient energy it cannot do this. Undernutrition occurs when people do not eat

enough, when the body fails to assimilate nutrients properly, or when there is a greatly increased need for energy (Burgess, 2008). Conditions that may greatly increase the need for energy include injury such as burns, high fevers, infections, pregnancy and surgery. Micronutrient deficiency is a sub category of malnutrition and it occurs when the body lacks one or more micronutrients such as iron or vitamins (Black 2008). Micronutrient deficiency is usually considered a different disorder. However, when energy is deficient, minerals and vitamins are also likely to be deficient.

Some immediate causes of undernutrition are poor diet and disease, or the consumption of meals that may contain substandard nutritional content. The lack of optimal breastfeeding, along with illness such as diarrhoea, pneumonia, malaria and acquired immune deficiency syndrome (AIDS) are major causes of undernutrition. In infancy, underlying causes are family food insecurity, unhygienic living conditions, inadequate health services and poor daily care. Poverty, lack of information and war are also major causes of undernutrition (Burgess & Danga 2008).

As a major cause of death and illness in the third world, it is important to understand both the effects that undernutrition can have on the human body and the diverse effects it can have on the development of children. It is also important to understand the steps being taken to diagnose and prevent the problem.

1.2.3 Effects and causes of undernutrition

In unborn and new born babies undernutrition of the mother may affect the unborn child, although the foetus is relatively protected as it adjusts its metabolism and production of hormones (Barker 2001). An example of this can be found in The Dutch Famine Birth Cohort study which followed mothers who were pregnant during the Dutch winter famine in 1944. Children of mothers who were pregnant during the famine were smaller on average. Furthermore, their children's children were also smaller than average (Painter et al. 2008). Undernutrition may also predispose new-borns to chronic diseases such as diabetes, cardiovascular disease and obesity in later life. However, a child can be small for their age at birth due to medications or toxins, such as alcohol taken by the mother that limits growth during pregnancy (Hanson et al. 1978), but genetics is also a factor. A mother's short stature and low weight before pregnancy has also been shown to be associated with low birth weight (<2500g) (Burgess & Danga 2008) .

For children up to 6 months old, breastfeeding problems, improper formula preparation, inadequate number of feedings and birth defects that affect the child's ability to eat or digest normally can all have an influence on growth. Infants who are not exclusively breastfed may achieve the same level of growth (Kramer et al. 2007), but there is an increase in the risk of infectious morbidity (Stuebe 2009). For those who use formula milk instead of breastmilk there is also a risk of using contaminated water, which combined with the loss of the immune protection conferred by breast milk, can make the baby ill. There may also be an increase in the amount of energy needed, such as in congenital heart disease (Krieger 1970).

Undernutrition is less common in those aged less than 6 months than other ages, unless a baby is not adequately breastfed or the mother is human immunodeficiency virus (HIV) positive.

It has been suggested that poor foetal growth, or stunting in the first 2 years of life, may lead to irreversible damage in later life, including shorter adult height, lower grades in school, lower income in later life and fewer offspring (Victora et al. 2008). In children, there are both short and long term potential complications. Malnutrition can compromise the immune system making them more prone to infectious diseases, in part due to poor nutritional status. This can lead to an even greater risk of weight loss and an increased risk of diarrhoea episodes. A Columbian study estimated that children who experienced between 2 and 8 episodes of diarrhoea within the first 3 years of life had between 2.5 and 10cm less body height than those who did not experience any episodes (Lutter et al. 1988). In the long term, malnutrition can limit bone growth. Children who are severely malnourished have poor developmental levels compared to matched controls or siblings even after they recover (Grantham-McGregor 1995).

1.2.4 Growth monitoring for the diagnosis of undernutrition

Growth monitoring involves regular measurement of children's size, and can be used to diagnose undernutrition. Where medical services are available, health staff will usually perform a physical examination for diagnosis. As part of the physical examination, health staff will measure weight, height, mid-upper arm circumference (MUAC) and from the height and weight measurements calculate body mass index (BMI) or weight-for-height. Health staff then plot these measurements on growth charts which are made up from measures of healthy children and serve as a standard for how children should grow. The

difference between growth 'standards' and 'references' are that standards are made up from a healthy population of children which serve as a standard to how children should grow, whereas growth references simply describe the growth of a sample of individuals.
In the past a number of charts have been used which summarise anthropometric measurements by age. The Tanner-Whitehouse (1966) charts cover weight, height, velocity, head circumference and skin folds from birth to 19 years of age (Tanner et al. 1966; Tanner & Whitehouse 1975). Gairdner-Pearson published charts in 1971 from 28 weeks gestation to 2 years. These charts cover weight, height and head circumference (Gairdner & Pearson 1971). The National Center for Health Statistics (NCHS) is an American reference covering height, weight and head circumference (Hamill et al. 1979). The UK 1990 charts included weight, height, BMI, head circumference and stages of puberty (Freeman et al. 1995). Measurements were from 23 weeks to 20 years (Wright et al. 2002). Note that these charts are all references rather than standards.
The two main charts which are currently used around the world are the World Health Organization (WHO) 2006 standard and the Center for Disease Control (CDC) reference, which superseded the NCHS reference (de Onis et al. 2007). The WHO standard is used within this thesis to compare children's growth, described at length in Section 1.2.6. Measurements obtained by the health staff can be compared with any of the above references or standards so that the centile in which they lie on can be determined.
In developing world countries, health workers collect measurements and monitor growth to detect and intervene in children where there is evidence of malnutrition. This is called growth monitoring (Garner et al. 2000). If a child's weight or height is average, it will follow the median (50^{th} centile) curve of the reference chart. If weight or height gain pattern is slower than average then this qualifies as growth faltering. In this case, an intervention may be needed.
Growth monitoring is considered an essential element of healthcare. In a survey of 178 Ministries of Health, 88% reported that they monitor child growth (de Onis, Wijnhoven, et al. 2004). The United Nations Children's Fund (UNICEF) recommends monthly growth monitoring for all children up to the age of 18 months in developing world countries and is usually considered part of development programmes aimed at improving nutrition (Panpanich & Garner 1999).
Growth charts are intended to provide a diagnostic tool for surveillance of individual children's health and nutrition, allowing effective action to be taken in the event of

growth faltering. They can also be used to teach families, mothers and health workers how weight and height can be affected by illness and poor nutrition, and this can promote awareness. Furthermore, it allows regular contact with primary healthcare services. If growth faltering is detected early and families and health workers are made visually aware of the abnormal height or weight gain, then families and health workers can intervene given that they are offered the motivation and information to do so (Ashworth et al. 2008). Further benefits of growth monitoring include community mobilisation and targeting of supplementary feeding. It also allows health workers to provide information on the prevalence of various indicators of undernutrition. Collecting information can otherwise be time consuming as special studies need to be designed and additional information needs to be collected on a non-routine basis.

There are many factors that determine whether growth monitoring is successful or not. Measurements must be taken with good quality equipment by trained staff who can take reliable measurements (Hamill et al. 1979). Poor growth may be reflected by a single attained extreme measurement, or can be indicated through downward centile crossing by plotting serial measurements (Olsen et al. 2007). Serial measurements in growth are highly correlated and children will usually tend to track along the chart. If a child crosses a number of centiles, there may be cause for concern. However, some centile crossing is actually expected due to “regression to the mean” (RTTM). RTTM is a statistical phenomenon where if individuals or groups of individuals who when measured lie at the extremes, will on average be closer to the population mean when they are measured again (Stigler 1997). This means the further children are initially from the centre of the population distribution, the more they are affected on average. Therefore, measures of growth should ideally take this expected movement into account (see Section 1.2.8 for further information).

Ashworth (2008) presents results from early studies, reporting how the prevalence of infant mortality in developing world countries was reduced through growth monitoring. A study in 1966 in Imesia, Nigeria, showed that infant mortality was 57 per 1000 in the intervention group (growth monitoring) compared with 91 per 1000 in the control village (no growth monitoring). For years 1-4, the intervention group mortality rate was 18 per 1000 and in the control group it was 51 per 1000. Furthermore, children within the intervention group were 0.3-0.4kg heavier and 1.5-3cm taller on average (Cunningham 1978; Gwatkin et al. 1980; Ashworth et al. 2008). Another example from Ashworth (2008)

illustrates the effectiveness of growth monitoring in Hanover, Jamaica. A child nutrition program comprising of growth monitoring and nutritional education showed that the 1-48 month mortality was 15.4 per 1000 in the years 1972-1973. However, after the introduction of growth monitoring and nutritional education this dropped to 7.4 per 1000 in years 1974-1975. Furthermore, the prevalence of underweight children dropped from 11% in July 1973 to 6% in July 1975 which can be attributed to the intervention of growth monitoring and nutritional education in July 1974 (Alderman et al. 1978; Ashworth et al. 2008). However, it must be noted that Roberfroid et al. (2005) concluded that there is too little scientific evidence to support the benefits of growth monitoring but found it may help with public nutrition in specific situations.

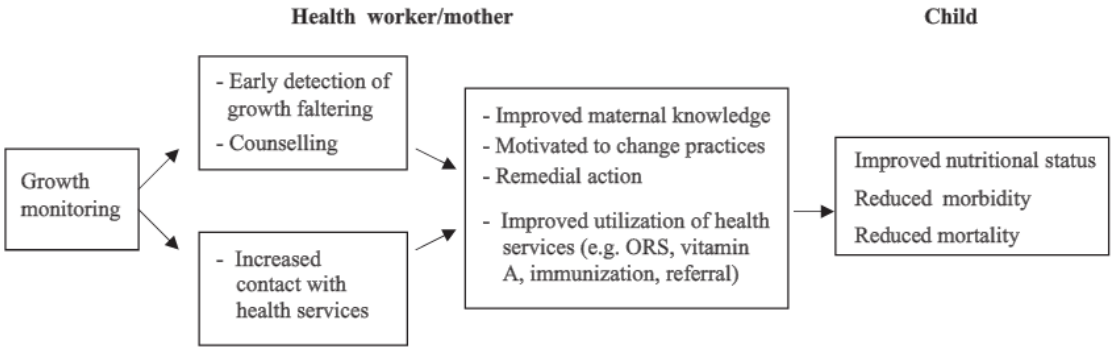

Figure 1.1: Framework for analysing the effectiveness of growth monitoring programmes (Ashworth, Shrimpton, & Jamil, 2008)

1.2.5 How are growth charts developed?

Thinking about the collection of data allows us to think about the construction of growth charts in the simplest form. Given that we would like to calculate centiles for our data so that we can determine where a given child lies relative to the population, we need to collect enough data at certain points along the x axis so that centiles can be calculated. Imagine, for example, that we wish to collect data at 5 months old. We can then calculate centiles for our data at that age. If we simply do not have enough information at 5 months, or measurements are taken slightly before or after this age (perhaps participants turn up late for their scheduled measurement date), then it makes calculations more complicated since we either do not have enough information, or the parameters for the distribution at this age may be slightly wrongly specified (since the population mean will not be at exactly 5 months) so the next stage in the construction of our chart would be to "join up the dots" between samples, producing a chart which has linear sections.

However, the problem with this approach is that it does not make physiological sense that a chart would have linear sections, and the closest effort to build a smooth chart would require vast amounts of data to calculate centiles at a vast number of ages. The concept of developing a smooth chart allows us to get over this hurdle. Smooth centile curves are more pleasing to the eye, but there is also justification that physiologically small changes in the covariate are likely to lead to continuous changes in the measurement, so that centiles ought to change smoothly (Cole & Green 1992).

Growth chart modelling techniques, such as the LMS method, can be used to create growth reference charts with smooth centile curves. The LMS method is arguably the most popular method proposed by Cole (1988) and Cole & Green (1992). This method models the population distribution, taking into account: skewness (L), central tendency (M), and dispersion (S). The parameters are estimated and then smoothed. Any desired centile can then be calculated from the smoothed L, M, and S parameters, and children's measurements can be compared with the chart. The centile that a child lies on can be converted to a Z score. Z scores are transformed measurements that express how many population standard deviations (SD) the measurement is from the population mean (assuming the distribution is a symmetric Gaussian). Z scores are normally distributed with zero mean and unit SD if a well matched reference is used (see Section 2.4 for more details on Z scores).

This methodology can be used to create a number of different kinds of charts. For example: weight-for-age, height-for-age, weight-for-height and BMI-for-age.

It is of interest to use this method to create growth reference charts for the developing world datasets made available for this research. This allows us to gain insights into how children within the sets grow in terms of median growth and variability.

1.2.6 World Health Organization standard

Internal references are those which are developed from the distribution of measurements in which the child originates, which describe how those children grow at a certain time and place. However, there are problems with using 'internal' references to compare child growth. It is healthy children, ideally from an 'external' population of healthy infants that should be used for comparison.

In 1993, the World Health Organization (WHO) reviewed the use of commonly used references and concluded that the NCHS reference did not adequately describe child

growth. There was a need to develop a chart which was not solely based on a sample of measurements at a time and place, and that could be used to compare growth on a universal level.

The WHO 2006 standard is a composite growth standard which includes children from a number of different backgrounds - both developed and developing world countries. The WHO Multicentre Growth Reference Study (MGRS) was a community based multi-country project. The MGRS study was conducted between July 1997 and December 2003 and was a population based study which took place in: Davis, California, USA; Muscat, Oman; Oslo, Norway; and Pelotas, Brazil as well as selected wealthy neighbourhoods of Accra, Ghana and South Delhi, India (de Onis & WHO Multicentre Growth Reference Study Group 2006). Further information on each of the sites can be found on the MGRS website. The process of selecting the study sites lasted 24 months and involved evaluation of eligibility criteria for study subpopulations based on the study protocol. The inclusion criteria were as follows: no known health or environmental constraints; mothers willing to follow MGRS feeding recommendations for at least 4 months; introduction of complementary foods by 6 months; no maternal smoking before or after delivery; single term birth and no significant morbidity. The criteria were defined to help select healthy children from comparable affluent backgrounds across sites (de Onis, Garza, et al. 2004).

The MGRS consisted of a longitudinal study up to 2 years and a cross sectional study from months 18 up to 71. Children were enrolled at birth and were visited a total of 21 times on weeks 1, 2, 4 and 6, monthly to 12 months then bi-monthly up to 2 years. The maximum delay or advance of measurements allowed by the protocol was 10% of the child's age, but in practice, teams worked with more restricted tolerable delay or advance targets (0, 1, 2, 4 and 5 days for visits at weeks 1, 2, 4 and 6 and 2 months respectively). Data was collected on anthropometric measurements, socio-economic, demographic and environmental factors.

The cross sectional design was adopted to avoid the time and cost constraints involved in conducting a longitudinal study to that age range. Furthermore, growth is more linear after 2 years of age. Using 18 months as the lower age limit for the cross sectional study allows an overlap of 6 months with the longitudinal study which provides information on the transition from supine length to standing height and facilitates the joining of the two datasets (de Onis, Garza, et al. 2004).

The field staff collecting anthropometric data had to be at least secondary school educated, be motivated, had to be able to write legibility, speak the local language and had to be able to interact appropriately with the high economic status families that were targeted for the study. Measurement procedures and training were prepared by MGRS coordinating centre in Geneva, Switzerland (de Onis et al. 2006). Anthropometrists were trained at each site by an experienced anthropometrist who followed the procedures outlined by the MGRS coordinating centre. An important aim was for the anthropometrists to measure accurately, without bias. To achieve this, measurements had to be obtained that were on average equal to the values measured by the expert. All study sites used the same measuring equipment.

In total there were 8,440 children studied across the 6 sites. Originally 1743 children were enrolled in the longitudinal sample. Six children were excluded for morbidities affecting growth leaving a final sample of 1,737 children. This was important since the sample was designed to describe healthy growth. Of the 1,737 children, 882 complied fully with the MGRS protocol and completed the follow up period of 24 months. The other 855 contributed only their birth records as they failed to comply with the criteria or dropped out.

The WHO charts were developed from the dataset by de Onis et al. (2006). The original method chosen to model the data was a method called LMSP, which models four distributional parameters: skewness (L) (shape parameter, transformation to symmetry); median (M) (location); coefficient of variation (S) (scale, or variability) and kurtosis (P) (shape parameter, the “peakedness”). However, in the end, the LMS method was ultimately chosen. The LMSP method is a more flexible generalisation of the LMS method however it requires a more complicated Z score calculation, simplifying to the LMS method if no adjustment for kurtosis is necessary. The curves that describe the growth of the population over time can be used to determine how other datasets fit the standard. As the WHO charts represent optimal rather than average growth and are to be used in a wide range of settings, variations in fit of individual samples are to be expected. Roelants (2013), Hui et al. (2008) and Wright et al. (2008) have compared children’s growth from the developed world with the WHO standard. Roelants et al. (2011) compared Belgian and Norwegian children with the reference and found that Belgian children’s weights were roughly on par with the standard at birth, but this average rised to 0.5SD by age 2.

Norwegian children tracked around 0.5SD above the WHO median from birth to age 2 (and beyond to age 5).

Hui et al. (2008) found both boys and girls from Hong Kong were found to gain weight at the same rate as the WHO standard (0SD) up until 3 years, although they were shorter by around -0.3SD. This may be down to the MGRS inclusion criteria as children in the MGRS study were from six countries: Brazil, Ghana, India, Norway, Oman and USA, none of which are in East Asia. Children from the UK in the Gateshead Millennium Baby Study were 0.17SD above the WHO median at birth, while children in the Avon Longitudinal Study of Parents and Children were 0.34SD above the standard at birth. This increased to around 0.67SD by 12 months and growth paralleled the WHO standard at this distance until age 2 and beyond (Wright et al. 2008).

All of these studies focus on identifying how children from the developed world fit the idealised reference by modelling the distributions relative to the standard. On the other hand, although many studies have investigated the proportions of children above and below certain threshold values in the developing world, little work has assessed how children from the developing world fit by modelling the relative distribution. One of the aims of this thesis is to determine how children from the developing world fit the WHO 2006 standard by modelling the relative distributions the diverse populations made available for study.

1.2.7 Indicators of undernutrition

There are various indicators currently used to diagnose undernutrition at either individual or population level. An infant's measurements are usually compared with the reference charts described in Section 1.2.4. For surveillance purposes, if their measurement is below a certain threshold value, then the child is categorised into one of a discrete number of "nutritional states" identified by the indicator. Table 1.1 displays the different available measures.

The WHO Expert Committee (1995) defined terms to describe failings in weight and height: low height-for-age (stunting), low weight-for-height (wasting) and low weight-for-age (undernutrition). Other indicators of undernutrition are also available, such as MUAC and BMI (Jamison et al. 2006).

Indicator	Description
Stunting (low height-for-age)	Children are categorised by height/length measurements, usually based on height-for-age Z score (HAZ). Children who are stunted are not as tall as their healthy counterparts (categorised by a threshold value) who have the same physical characteristics such as similar age and sex. Moderately stunted: -3 $\leq$ HAZ < -2 Severely stunted: HAZ < -3
Underweight (low weight-for-age)	Children are categorised by weight measurements, usually based on weight-for-age Z score (WAZ). Analysis similar to that above. Moderately underweight: -3 $\leq$ WAZ < -2 Severely underweight: WAZ < -3
Wasting (low weight-for-height)	Children are categorised by height and weight measurements, usually based on weight-for-height Z score (WHZ). Moderately wasted: -3 $\leq$ WHZ < -2 Severely wasted: WHZ < -3
Mid-Upper Arm Circumference	Distance around the mid-upper arm, measured at the mid-point between the tip of the shoulder and the tip of the elbow. Moderate Acute Malnutrition (MAM): 115mm $\geq$ MUAC $<$ 125mm Severe Acute Malnutrition (SAM): MUAC<115mm
Body Mass Index	Computed as: $$BMI = \frac{(weight\ in\ kilograms)}{(height\ in\ meters)^2}$$ BMI categories used to define underweight, normal weight, overweight and obese.

Table 1.1: Indicators of undernutrition

The definition of wasting, defined by the WHO Expert Committee (1995), is less than 2SD below the median weight-for-height of the reference population (WHO standard). Hamill et al. (1979) states that age independence which is an assumption of weight-for-height standards "is not quite true", as children's weight and heights increase at different rates at different ages. During the first year of life, children of a certain height tend to weigh more the older they are (Van Wieringen 1978). Therefore, weight-for-height standards are slightly biased (Cole et al. 1981).

However, the effect of age can be included within the assessment of weight-for-height. McLaren and Read (1972) compared the ratio weight/height to the weight-for-

age/height-for-age ratio of the NCHS reference, obtaining an estimate of the ratio as a fraction of the expected ratio. This introduced age, as children were compared with median weight-for-age/median height-for-age of the NCHS standard. However, this measure was criticised as it did not accurately describe the rate of weight gain relative to height gain over the time frame, assuming that weight and height increased at the same rate. This rate was quantified using linear regression by Cole et al. (1981) in children aged 0-12 months. Cole regressed log weight-for-age on log height-for-age in children within the same age group, quantifying the relationship between the two measures. Cole obtained $\log WfA = \log(\alpha) + p.\log(HfA)$ which is equal to $WfA = \alpha.HfA^p$.
Cole (1979) and Cole et al. (1981) found that the regression slope for pre-school children is roughly equal to 2, and α is equal to 1. This means that W/H^2 can be compared with the reference WfA/HfA^2, resulting in a better index than W/H. This leads to the name W/H^2-for-age since it takes age into account via the weight-for-age and height-for-age reference ratio. Rao & Singh (1970) and Babu & Chuttani (1979) suggested W/H^2 as an age independent index for weight-for-height in children. However Van Wieringen (1978) found that the ratio falls between 1 and 6 years, indicating that W/H^2 by itself is unreliable unless standardised for age (Cole et al. 1981).
Therefore, Cole (1985) states that W/H^2 should be used for pre-school children. However, W/H^3 may be a more suitable measure for older children going through puberty since the weight gain relative to height gain increases with age. The power p decreases to 2 after puberty (Cole 1986).
For these reasons, we use BMI-for-age Z score below -2SD relative to the WHO standard to define wasting within our latter analyses. This measure takes into account the relationship between weight, height and age, which weight-for-height references fail to do. When referring to wasting in reference to other published literature, we still refer to the WHO's definition of WHZ < -2. It is important to note that one advantage of weight-for-height standards are that they can be used when children's ages are unknown.

1.2.8 Assessing weight gain over time

The standards described in Sections 1.2.5, 1.2.6 and 1.2.7 are usually referred to as 'cross sectional standards' which describe the distribution of measurements as functions of covariates, usually age and sex. Single measurements compared with cross sectional standards allow paediatricians to assess relative attained size. Serial measurements allow

for a more accurate assessment of a child's growth rate, which describe change in size relative to the reference over time.

The space between the major centile lines are called channels. There is cause for concern if a child crosses more than two channels (Shields et al. 2012).

When assessing growth curves (by joining several measurements of attained size by a line) relative to the reference population, it is normal for children to cross one or two centiles in infancy and puberty, most likely towards the 50th centile due to RTTM (Canadian Paediatric Society 2004). However, measurements showing unexpected downward trend from a previously established rate of growth may be a sign of growth faltering.

On the other hand, if the child sits at a higher centile at time point 2 than they did at time point 1, then they have gained weight at a faster rate than expected. But, how can this 'faster' rate be measured and compared to other children's weight gain?

Growth increments are the difference in size between two time points. For example, for weight measurements at t_1 and t_2 the growth increment is $W_{t_2} - W_{t_1}$. Growth velocity is the difference divided by the length of the interval, expressed as rate of growth per unit, $(W_{t_2} - W_{t_1})/\Delta t$, where $\Delta t = t_2 - t_1$. Velocities can be compared with velocity references, such as the WHO 2009 velocity standard, allowing comparisons with an idealised population of healthy children. As stated however, serial measurements are expected to regress towards the mean, which these measures do not take into account. Conditional measures such as those by Cameron et al. (1980) and Berkey et al. (1983) estimate future size using a linear regression model conditional on size at a previous time point, taking RTTM into account while working on the raw measurement scale. Similar work has been carried out by Cole (1994), Cole (1995) and Cole (1997), aiming to provide a simple method that would allow practitioners to assess centile change between two time points while adjusting for RTTM on the SD scale. This method is known as "conditional weight gain". Conditional weight gain is a valid alternative for growth increment or growth velocity (Roelants 2013), using the properties of Z scores to express the change relative to the population in which the child originates. There are advantages of working on the SD scale over the raw measurement scale (Cole 1995). The Z scores follow a normal distribution with zero mean and unit variance and can therefore be expressed as a centile, if an appropriate growth reference is used.

The method provides an unbiased estimate of the change in weight since RTTM is taken into account. In this context the amount of RTTM is synonymous with correlation (Cole 1995). Therefore, a way to adjust for RTTM when working on the SD scale is to quantify the correlation between groups of measurements at t_1 and t_2. The expected position of a measurement on the growth chart is therefore given by $E(Z_2|Z_1) = rZ_1$. The difference between the true measurement and the expected value can be converted into a Z score itself by dividing by the standard deviation, $\sqrt{1-r^2}$. This measure can then be expressed as a centile (Cole 1995). All that is needed is a model which can interpolate the correlation between measurements on the SD scale.

As noted, a problem with this method is that it requires a well matched reference so that the mean of the scores is 0 and the standard deviation is 1. To guarantee these conditions, only an 'internal' reference population can be used. However, there is a downside to using internal reference populations as they describe how children grow at a certain time and place. Ideally, references such as the WHO standard should be used. An aspect of the primary aim of this thesis is to derive an expression which can be applied to external references, and to investigate which models most accurately represent correlation structures. The new measure will be a generalised version of conditional weight gain, hence the name generalised conditional weight gain.

1.3 Using anthropometry to predict future outcomes

We have described undernutrition and its effects, and now move onto a more detailed review of literature which has investigated the drivers of undernutrition and how the use of anthropometry can predict future adverse outcomes. The original study aim for this thesis was to investigate what level of weight gain was associated with adverse outcomes and whether weight gain was any more predictive than recent size. Later in the programme of work we added an investigation of the relationships between 'nutritional states' as defined in Section 1.2.7. This investigation was added since the review of literature surrounding the original study aim highlighted that, while measures of both size and growth are predictors of adverse outcomes, nutritional states are also predictors. Furthermore, relationships are presumed to exist between them. However, these transitions are poorly understood.

Existing research in these areas is reviewed in detail in this section and hypotheses based on the research are developed at the end. We investigated the general literature in the

area then categorised the results into sub-sections. The review is split into three parts: papers which have investigated measures which have been shown to predict mortality, papers which have investigated whether wasting leads to stunting, and papers which have identified whether stunting is reversible.

Both backward and forward searching was used to find literature, mainly on internet search engines such as Google Scholar. Many of the papers found were observational, but a few were trials. Using search terms like 'weight', 'morbidity' and 'mortality' came up with millions of results due to the number of papers which come under these umbrella terms. Adding terms like 'velocity' decreased the number to hundreds of thousands. Finding relevant literature, therefore, proved hard due to the number of irrelevant results. Combinations of the words: early, infant, life, developing, world, child, children, mortality, morbidity, gain, weight, height velocity, growth, birth, validity, predict, predictive, survival, wasting, stunting, undernutrition, pathways, transition, probabilities, relationship, irreversible, nutritional, state, reversible, severe, moderate, mild, malnutrition and size were used within the search criteria as well as their synonyms, but still yielded fairly sparse literature. The lack of relevant material emphasises the need for further research in this area.

1.3.1 How well do anthropometric measures predict mortality?

This section is split into sub-sections assessing: size/indicators of undernutrition as predictors of mortality, growth as a predictor of mortality, and a section reviewing papers which have evaluated size/indicators of undernutrition vs. growth. Components of papers which review size vs. growth also review measures of size and growth independently in the first two sub-sections. The reason for this is that the original study question was to determine whether size or growth is the best predictor of mortality. Exploring literature which has investigated each of these areas provides an understanding of what is currently known as to how each measure can be used as a predictor, which leads us to develop hypotheses based on the research, which we will then test.

1.3.1.1 Size/indicators of undernutrition

Weights and heights relative to reference populations have been shown to be predictive of mortality. Fawzi et al. (1997), using multivariate logistic regression, found an inverse relationship between height-for-age and mortality within a cohort of 28753 Sudanese children, indicating the higher a child's height-for-age, the lower the risk of death. Bairagi

& Chowdhury (1994) investigated whether height-for-age, weight-for-age and weight-for-height were predictors of 2 year mortality in a cohort of 1,976 12-23 month old children from Matlab, Bangladesh. Bairagi used tabulation (χ^2 tests) and logistic regression and found weight-for-age, height-for-age and arm circumference were significant predictors in univariate logistic regression models. Work assessing weight and height relative to reference populations was also carried out by Pelletier et al. (1995), who studied the proportion of child deaths related to malnutrition, defined as low weight-for-age in 53 developing world countries around the world. Pelletier found that 56% of child deaths were attributable to malnutrition's potentiating effects and 83% of these were attributable to mild to moderate malnutrition (60-80% of the median weight-for-age of the reference population) opposed to severe malnutrition, indicating the impact of malnutrition on child mortality.

Schroeder & Brown (1994) pooled results from 5 previously published studies (using tests of heterogeneity to determine if datasets could be pooled, and pooling relative risks (RR) over the studies) and found that young children (6-60 months) with mild to moderate malnutrition (defined as <80% weight-for-age) had 2.2 times the risk of dying during the follow up period compared with their better nourished counterparts (>80% weight-for-age) and severely malnourished children (<60% weight-for-age) were 6.8 times more likely to die.

Nutritional status, defined using the definitions in Section 1.2.7, has also been shown to predict mortality. Olofin et al. (2013) used Cox proportional hazards models to estimate hazard ratios (HRs) in a pooled analysis involving children aged 1-59 weeks in 10 prospective studies in Africa, Asia and South America. They found that those who were wasted were at a higher risk of mortality than those who were underweight or stunted, and as weight and height Z scores (relative to the WHO standard) decreased, the risk of mortality increased (using most recent measurements to categorise children into each nutritional state).

In the same pooled set as Olofin's, McDonald et al. (2013) used Cox proportional hazards models to determine whether a combination of anthropometric deficits increased the risk of death. They found wasting (W), underweight (U) and stunting (S) were all associated with an increased risk of death. Wasting was worse than underweight, which was in turn worse than stunting, with pooled HRs of 2.30, 2.49 and 1.47 respectively compared with children with no anthropometric deficits. Furthermore, there was a cumulative effect

where children had more than one anthropometric deficit. McDonald found HRs of 3.4 for S+U, 4.7 for W+U and 12.3 for W+S+U (again, using their most recent measurements to categorise children).

Vesel et al. (2010) also carried out work on the predictive validity of nutritional status using sensitivity and specificity to determine whether stunting, wasting and underweight were good predictors of mortality from 6 weeks to 6 months, and from 6 months to 12 months in Ghana, India and Peru. Sensitivity is the proportion correctly predicted of those who died, while specificity is the proportion correctly predicted of those who did not die. Using the WHO standard, in Indian infants from 6 weeks to 6 months, underweight was the best predictor, with sensitivity of 70.2% and specificity of 85.8%. There was no indicator available in Ghana or Peru <6m. From 6-12 months, underweight had the highest sensitivity and specificity in Ghana (37.0% and 82.2%, respectively) and Peru (33.3% and 97.9% respectively), while wasting was the best predictor in India (sensitivity: 54.6%; specificity: 85.5%).

Van Den Broeck et al. (1993) studied the influence of nutritional status on child mortality in rural Zaire (children aged 0-72 months), setting out to determine the association between mild to moderate undernutrition and mortality in short term (3 month) and long term (3-30 month) periods where risk of death was attributed to nutritional status at the initial time point. For short term mortality, children who survived were included every successive 3 month period. Children were categorised into WAZ, HAZ, WHZ and MUAC-for-age SD categories < -4, < -3, < -2 and the reference category of $\geq$ -2. Those in category < -4 had RRs for short term mortality of 176, 143, 159 and 15 for MUAC-for-age Z, WHZ, WAZ and HAZ. This reduced to 20, 40, 15 and 2 for those in category < -3 and 3, 3, 2 and 1.2 in category < -2. For long term mortality, those in category < -4 had RRs of 28, NA, 47 and 14 for MUAC-for-age Z, WHZ, WAZ and HAZ. This reduced to 3, 2, 10 and 3 for those in category < -3 and 2, 2, 6 and 1 in category < -2.

The last four papers to be reviewed have compared size as well as growth as predictors of mortality. Therefore, they appear in the next two sections as well as this one, reviewing the relevant components of the papers within each section. They are reviewed in chronological order. Three of them utilise datasets from Bangladesh, and one from the Democratic Republic of Congo (DRC).

Bairagi et al. (1985) investigated whether weight-for-age %, height-for-age % and weight-for-height % (NCHS reference) were significant predictors of one year mortality (children

aged 0-4 years) in rural Bangladesh. To compare indicators between living and dead children, Bairagi used normalised distances (ND) (a measure which allows comparisons between different indicators) and maximum sum of sensitivity and specificity (MSS). Weight-for-age % (ND = 0.71, MSS = 144) was the best predictor, followed by height-for-age % (ND = 0.72, MSS = 126), then weight-for-height % (ND = 0.38, MSS = 116).

Bairagi et al. (1993) also conducted research in Bangladesh to investigate whether weight-for-age % (NCHS reference) is a significant predictor of mortality. Bairagi used NDs and t-tests to compare the weights of living and deceased children for significance. The weights of 1,876 children aged 6-36 months were measured at time point t_1. 1,900 children were measured 3 months later, t_2. 1,664 children had measurements at both time points. Of the 1,664 children, 27 died during the 12 month follow up period from the last measurement t_2. The results showed that those who were still alive at the end of the study had a weight-for-age percentage of 70.5%, while those who died had a weight-for-age of 61.48% prior to death $(p < 0.05)$.

Again, in Bangladesh, Briend & Bari (1989) investigated how weight-for-age % (NCHS reference) affected mortality. Briend used child months for analysis (a child was a survivor if alive at the next month) with children aged less than 36 months, measured each month for up to two years. MSS was used and significant differences were assessed using logistic regression and relative risks (RR). For weight-for-age <60% of the reference, the MSS was 161.5 and the odds of death were 14.7 using logistic regression ($p < 0.05$). The RR was 14.5 (95% CI: 9.8 – 21.3) compared with those >60%.

More recently, O'Neil et al. (2012) studied child mortality in rural Africa, predicting mortality by nutritional status. 2,402 children between 0 and 24 months in the rural areas of the DRC with high malnutrition and mortality rates were studied. First measurements were taken between the ages 0 and 2 years old and every three months thereafter (up to 24 months). 57 children died during follow up (median follow up time of 15 months). O'Neill split up anthropometric indices into three categories once they were transformed to Z scores: > -2, $-3 < Z < -2$ and $Z < -3$. O'Neill carried out a univariate age adjusted analysis to determine HRs for those within each category of: weight-for-height, BMI-for-age and MUAC-for-age. Baseline anthropometric indices were specified as the second measurement taken (the second measurement was used since the first was used along with the second to assess growth as a predictor).

Age adjusted HRs and HRs without age adjustment were calculated. In the age adjusted univariate analysis, those with $-3 < \text{HAZ} < -2$ were no more likely to die than those with $\text{HAZ} > -2$ and those with $\text{HAZ} < -3$ were 2.3 times more likely than those with $\text{HAZ} > -2$. BMI-for-age Z was an equivalent predictor of WHZ. Those with $-3 < \text{BMI} - \text{for} - \text{age} < -2$ were 3 times more likely to die than those with $\text{BMI} - \text{for} - \text{age} > -2$ and those with $\text{BMI} - \text{for} - \text{age} < -3$ were 8.6 times more likely. Those with MUAC-for-age $Z < -3$ were 3.1 times more likely to die than those in with MUAC-for-age > -2. This review of literature has found that weight-for-age, height-for-age, weight-for-height, arm circumference and nutritional status (including combinations of stunting, wasting and underweight) are all associated with mortality. The lower the Z score of a given measure, the higher the risk of death. Generally, wasting is a better predictor of mortality than stunting and children with more than one anthropometric deficit are at a higher risk of death.

1.3.1.2 Growth/weight gain

Various measures of growth have been found to be valuable predictors of mortality. Bairagi et al. (1985), Briend & Bari (1989), Bairagi et al. (1993) and O'Neill et al. (2012), reviewed in the previous section, also investigated the predictive validity of growth velocity as well as size.

Bairagi et al. (1985) assessed the best interval to compute velocity over, by comparing velocity's predictive validity over 2, 4 and 6 month intervals of children under 10. Bairagi found 4 month velocity (g/m) was a better predictor than 2 or 6 month velocity. However, while 4 month velocity was a better predictor than 2 and 6 month velocity, none of the measures were significant predictors.

Briend & Bari (1989) investigated whether weight change in the previous month and weight change in the previous 3 months were significant predictors of mortality. For weight change in the previous month, MSS was 142 and t-tests showed significant differences between weight change of the living and deceased (taking into account various diseases). The RR for mortality was 5.8 (95% CI: 3.6 – 9.0) if children experienced one month weight loss. For weight change in the previous 3 months, the RR was 11.1 (95% CI: 7.3–16.7). Research by Bairagi et al. (1993) agreed with Briend's work, finding velocity to be a significant predictor. Bairagi et al. (1993) investigated whether weight velocity (g/month) was a significant predictor of mortality. Children were measured aged

6-36 months, then 3 months later. Bairagi also defined a 3 month weight velocity index as the ratio of the weight velocity of the child, to the weight velocity in the 50th centile of the NCHS standard, for a child of the same age and sex, multiplied by 100. Bairagi found that both weight velocity and the index were significant predictors, although weight velocity was better (using ND and MSS). Children in the living group had a weight velocity of 179g/month whereas children in the deceased group had a weight velocity of 78g/month. O'Neill et al. (2012) also investigated whether weight gain was a significant predictor of mortality in the DRC. O'Neill used 3 month weight velocity-for-age Z score (initial measurement between 0 and 24 months), relative to the WHO standard. O'Neill used a Cox proportional hazards model, categorising velocity-for-age Z score into one of three groups. In the univariate age adjusted analysis, those with weight velocity-for-age Z score between -2 and -3 were no more likely to die than those with weight velocity-for-age Z score > -2 and those with weight velocity-for-age Z score < -3 were 5.3 times more likely to die ($p < 0.01$) than those with weight velocity-for-age Z score > -2.

Finally, in a developed world setting, sudden unexpected infant deaths are defined as deaths in infants less than 1 year of age that occur suddenly and unexpectedly, and whose cause of death is not immediately obvious prior to investigation (Centers for Disease Control and Prevention, Sudden Infant Death, 2013). Blair et al. (2000) investigated patterns of infant growth that might influence the risk of Sudden Infant Death Syndrome (SIDS). The study used the British 1990 growth reference (Cole et al. 1998) to measure growth of children who died due to SIDS. Using logistic regression, Blair found weight gain Z score from birth to the final weight observation, and birth to 6 weeks, were significantly less among the SIDS infants after adjustment for potential confounders (including birthweight). The mean change in weight Z score was -0.38SD for those who died from SIDS versus the controls with a mean change in Z score of 0.22SD.

To summarise, these papers suggest that growth is a significant predictor of mortality. Blair et al. (2000) found conditional weight gain Z score was significantly associated with SIDS. O'Neill et al. (2012) found that 3 month weight velocity Z score (WHO standard) (initial weight between 0 and 24 months) was a significant predictor. Bairagi et al. (1993) found both 3 month weight velocity and a weight velocity index were both significant predictors (initial weight between 3 and 36m). However, weight velocity was better than the index. Briend & Bari (1989) found that weight change in the past 3 months was a

better predictor than weight change in the past month. However, Bairagi et al. (1985) did not find that velocity was a significant predictor in their study.

1.3.1.3 Size vs. growth

Only four papers could be found which made comparisons between size and growth as predictors of mortality: Bairagi et al. (1985), Briend & Bari (1989), Bairagi et al. (1993) and O'Neill et al. (2012), as reviewed in the previous two sections.

Bairagi et al. (1985) compared weight-for-age %, height-for-age %, weight-for-height % (as continuous measures) and 2, 4 and 6 month weight velocity (NCHS reference). Four month velocity was the best interval to compute velocity over. However, 4 month weight velocity (ND = 0.38, MSS = 116) was not as a good a predictor as weight-for-age (ND = 0.71, MSS = 144), height-for-age (ND = 0.72, MSS = 126) or weight-for-height (ND = 0.65, MSS = 124). In fact, weight velocity was not significant, whereas all other measures were.

Briend and Bari (1989) agree with Bairagi et al. (1985) who also found that size is a better predictor than growth. Briend used MSS and Z_{d_a} tests (for predictors which may be on different scales) to compare the normalised distances of those who survived and died between weight-for-age (ND = 1.618) and: weight change in the previous month (ND = 0.799) (g), weight change in the previous 3 months (ND = 1.304) (g), weight-for-age change in previous month (ND = 0.683) (%) and weight-for-age change in previous 3 months (ND = 1.147) (%). The result of the Z_{d_a} test coupled with the ND and MSS indicated weight-for-age % was a significantly better predictor than the rest.

Briend found that weight-for-age % had a larger ND (1.618) and a larger MSS (161.5) than the rest. Weight change in the previous 3 months was the next best predictor with an ND of 1.3 and MSS of 154. Briend also used a logistic regression model including both weight-for-age <60% as a categorical variable and weight loss in the previous month. Weight-for-age <60% had an odds ratio of 12 (95% CI: 7.2 – 20.2) and weight loss in the previous month had an odds ratio of 4.3 (95% CI: 2.5 – 7.2).

Bairagi et al. (1993) compared weight-for-age % (NCHS reference) and weight velocity (g/m) (and the velocity index defined in 1.3.1.2) of those who died and survived over one year with MSS and normalised distances. Weight-for-age % had the largest normalised distance (0.76) and the largest MSS (133), followed by weight velocity (ND = 0.50, MSS = 128), then the index (ND = 0.43, MSS = 124).

Results from Bairagi et al. (1993), Bairagi et al. (1985) and Briend & Bari (1989) all agree that size is a better predictor than growth. On the other hand, O'Neill et al. (2012) found weight gain to be a better predictor than size. O'Neill used a multivariate age-adjusted Cox proportional hazards model including HAZ, BMI-for-age, MUAC-for-age and 3 month weight velocity-for-age Z score, calculated between 0 and 24 months to predict mortality. In the multivariable adjusted analysis, only those in BMI-for-age category 3 and velocity-for-age category 3 were significant. Those in BMI-for-age category 3 were 3 times more likely to die than category 1 ($p = 0.04$). Those in weight velocity-for-age group 3 were 3.6 times more likely to die than those in category 1 ($p < 0.01$). O'Neill's analyses suggest both measures are predictive (although only for Z < -3) of mortality. However, from O'Neill's results, 3 month weight velocity Z score is more predictive than recent BMI-for-age Z score.

To summarise, these papers suggest that measures of both size and growth are significant predictors of mortality. Bairagi et al. (1985), Briend & Bari (1989) and Bairagi et al. (1993) agree that size was a better predictor than growth whereas O'Neill et al. (2012) found growth was a better predictor than size.

However, O'Neill used BMI-for-age Z score to represent size, and 3 month growth velocity Z score to represent growth within her models (relative to the WHO standard), whereas Bairagi et al. (1985), Bairagi et al. (1993) and Briend used weight-for-age % (relative to the NCHS reference) to represent size and the weight velocity to represent growth. O'Neill acknowledged her results are in contrast with Briend's, and speculated that the difference may be down to the use of the WHO weight increment standards. These standards are 1 month increments from birth until 12 months and 2-6 month increments from birth to 24 months. O'Neill suggested that growth monitoring programs should take the WHO 2009 velocity standards into consideration when defining critical values of low weight velocity.

Furthermore, children in O'Neill's analysis were 0-24 months old at their first visit, whereas they were 0-48 months in Bairagi's 1985 analysis, 3-36 months in Bairagi's 1993 analysis and 0-36 months in Briend's. This may have influenced results. Briend et al. noted that both weight gain and mortality was higher in younger children. A confounding effect may, therefore, explain the poor performance of velocity. However, age was included within a logistic regression model which did not change the relationship between velocity and mortality.

1.3.2 How do wasting and stunting relate?

As both stunting and wasting are known consequences of undernutrition it is to be expected that there will be a relationship between wasting and stunting, and studies have generally found that there is. Victora (1992) calculated correlations and regressed the prevalence of wasting on stunting in different regions around the globe. Correlations were 0.7 in Asia, 0.65 in the Eastern Mediterranean, 0.3 in Latin America and 0.04 in Africa. Gorstein et al. (1994) also found little correlation between wasting and stunting in Africa ($r = 0.1$). Stunting was significantly associated with wasting in Asia and the Eastern Mediterranean.

Recent work by Richard et al. (2012) found wasting is significantly related to low height-for-age later in life. Richard analysed 8 cohort studies using correlation coefficients, mixed effects models (linear model with random effects term) and logistic regression. Richard found that wasting at 6-11 or 12-17 months was associated with decreased HAZ at 18-24 months, although children who were wasted at 0-5 months did not suffer any long term growth deficits compared to children who were not wasted.

Walker & Golden (1988) studied children treated for severe malnutrition at the Tropical Metabolism Research Unit, University Hospital of the West Indies (mean age of 12.6 years). Children were treated as soon as they entered the hospital. Of those who had linear catch up growth, two thirds started to grow only after reaching 85% of the NCHS median weight-for-height. Those who had linear catch up were initially more stunted than the group as a whole ($p < 0.0001$).

Studies by Doherty et al. (2001), Costello (1989) and Walker et al. (1996) found that low initial WHZ predicted subsequent low growth in height. Doherty et al. (2001) studied children aged 6 to 36 month who were followed for 90 days in Bangladesh. Doherty used regression analysis to identify predictors of linear growth. Maternal height and baseline WHZ were both significant predictors of change in HAZ over 90 days. More evidence supporting the hypothesis that children gain weight at the expense of height can be found in Costello (1989), who analysed data taken from 441 children aged 0-6 years in Nepal using linear regression, then again 6 months later. Costello found a positive relationship between initial WHZ and height velocity over the 6 months, and a negative relationship with weight velocity, indicating thin children seemed to catch up weight at the expense of height. Costello argued that stunting is caused largely by a reduced growth velocity during the nutrition dependent infantile phase of growth.

Walker et al. (1996) also found a positive relationship between initial weight-for-height and linear growth. Walker studied the relationship between wasting and linear growth in a group of stunted children aged 9-24 months (measured every 6 months for a 2 year period) in Kingston, Jamaica, using linear regression. Stunting was defined as < -2SD below the NCHS reference median. The results showed that initial weight-for-height predicted linear growth in the following 6 month interval in the stunted children, and the change in weight-for-height in the preceding interval was a better predictor of linear growth in the next interval than attained weight-for-height at the beginning of the interval.

In summary, associations have been found between measures which categorise children thin and short. Victora (1992) found that a correlation between wasted and stunted states exists in some regions around the world. Richard et al. (2012) found that early wasting was associated with later HAZ. In fact, Doherty et al. (2001), Costello (1989) and Walker et al. (1996) all found that low initial WHZ predicted subsequent low growth in height, while Costello (1989) found low initial WHZ predicted subsequent higher growth in weight. Walker & Golden (1988) agreed that catch up would only occur once a certain WHZ was reached.

These papers indicate that initial wasting results in subsequent low linear growth, suggesting that wasting precedes stunting. Various papers have discussed this presumed relationship, such as Briend et al. (2015). Briend states, “The relationship between wasting and stunting is poorly understood. Although a coherent plausible picture emerges from our current knowledge, there remain evidence gaps and mechanisms that require confirmation”. Briend, however, also states “children with severe wasting are often stunted, suggesting that wasting and stunting have a common cause or that one form of malnutrition can contribute to the other”. Richard et al. (2012) notes “the longitudinal relationship between stunting and wasting in children is poorly characterized” but suggests “instances of wasting or poor weight gain may precede linear growth retardation”. An Emergency Nutrition Network (ENN) report by Khara & Dolan (2014) provides further discussion of the relationship between wasting and stunting, quoting Black et al. (2013) “severe infectious disease can lead to wasting, which may have longer-term consequences on linear growth” and Brown et al. (1982) “periods of lowest linear growth follow periods of lowest weight acquisition”. Ashworth & Millward (1976) state

that children recovering from severe protein energy malnutrition catch up in weight-for-height before catch up in linear growth.

1.3.3 Is stunting irreversible?

Generally, research indicates that children who are stunted tend to stay stunted, and that stunting is largely irreversible. Vella et al. (1994) studied child stunting in Uganda. Vella used logistic regression to predict stunting 2 years after children entered the study and found prior stunting, no matter what age, was a significant predictor of stunting in the future. Mothers education was the only predictor of recovery from stunting (notably, income was not significant).

Earlier work by MacWilliam & Dean (1965) followed up children who were discharged from hospital in Uganda and found that they moved from 87% to 90% of the international standard after 36 months but 6-10 years later they were still stunted, with little evidence of catch up to the international standard.

Beaton (1990) states that children who are stunted and remain in the same environment are not likely to recover. In fact, Mjönes (1987) studied children who changed environment and the results agree with Beaton, finding that children who change environment do recover. Mjönes compared Turkish school children born in Sweden with children matched for sex, age and physical environment with children born in Turkey but who moved to Sweden. The emigrants were shorter than the Swedish born counterparts on arrival but soon caught up. Martorell (1992) studied a Guatemalan cohort and found that in early childhood (three year olds), those randomised to the atole group (which had protein supplementation) had around half as many severely stunted ($Z < -3$) children compared to the fresco (control) group after intervention, suggesting that catch up growth is possible with intervention.

Another study that investigated a change of environment was Schumacher et al. (1987). Schumacher showed that a study of 835 refugee children moving to the United States were almost all below the 25th NCHS centile for height before arrival, but quickly caught up to with many growth velocities around the 90th centile. Schumacher compared the number of children with HAZ, WAZ and WHZ scores categorised into positive and negative change across Chinese, Southeast Asian, Hispanic and Filipinos in their first year in the US. Significant positive change was found in HAZ scores for all areas, WAZ in all except Filipino boys and WHZ in Chinese and Southeast Asian boys. Martorell et al. (1994) states that the

potential for catch up growth increases as maturation is delayed and the growth period is prolonged. Golden (1994) agrees, stating “If puberty is delayed and/or growth continues into the early or mid-twenties, then an acceptable final adult height is achieved”. However, maturation delay in developing countries is only enough to compensate for a small fraction of growth retardation. Evidence for catch up growth can be found in Adair (1999) who states her results “suggest that there is a large potential for catch-up growth in the preadolescent years”. Of 63% of children within a Filipino sample who were stunted at age 2, 30% were no longer stunted by age 8.5 and 32.5% were no longer stunted by age 12. On the other hand, Adair found that those not stunted at age 2 were 3.97 times more likely to not be stunted at age 8.5 (95% CI: 2.93 – 5.36) compared with those stunted at age 2. Adair also found that mean HAZ scores for those stunted at age 2 was significantly lower for those stunted at birth than those who were not and this significantly reduced the likelihood of catch up in later life.

In summary, research suggests that stunting is largely irreversible, unless a child changes environment, where it is more likely that they will exhibit catch-up growth. Potential for catch up growth among stunted children is generally thought to be limited after 24 months (Adair 1999). Children in the developing world are on average shorter than those from more affluent societies (Golden 1994), born slightly below the international reference standard for height and weight. Victora et al. (2010) found children from low to middle income countries quickly fall behind within the first 2 years of life in terms of both HAZ and WAZ, relative to the WHO standard. Victora also assessed different regions, finding children from South Asia had the lowest WAZ, HAZ and WHZ scores on average, followed by sub-Saharan Africa.

1.3.4 Hypotheses

From Section 1.3.1, previous research suggests measures of both size and growth predict mortality, but papers that have investigated which measure is better have published conflicting results. However, the majority suggest size is a better predictor than velocity. From Section 1.3.2, papers suggest that those who are wasted are at a higher risk of mortality than those who are stunted, and those who are wasted as well as stunted (wasted-stunted) are at a higher risk of mortality than those who are just wasted or stunted alone. Research also suggests that wasting leads to stunting and that children who become stunted will most likely stay stunted

Hypotheses are presented based on this summary of literature, which we aim to test within the thesis.

Hypothesis 1: those who are both wasted and stunted are at a higher risk of death than those who are wasted, who are in turn at a higher risk than those who are stunted

Hypothesis 2: wasting tends to precede stunting

Hypothesis 3: children who become stunted are likely to stay stunted

Hypothesis 4: recent size is a more useful predictor of later mortality than growth over a recent time period

Our main aims for this thesis are therefore to identify the trends that children take through nutritional states (hypotheses 1-3) and to identify whether growth over a recent time period adds value to size when predicting mortality (hypothesis 4).

These aims, as well as others which focus on data preparation and modelling, are outlined in the next section.

1.4 Aims and objectives

This section outlines the aims and objectives for this thesis. Firstly, the aims focus on the available data: getting the data into a common format, summarising and modelling the sets. The aims then focus on testing the hypotheses developed in Section 1.3.4.

1.4.1 Data formatting and modelling

This first section of aims focus on data formatting, summarising and modelling.

1. *Import the three developing world data sets into a common format, model and summarise them.*

Reformatting the datasets into a common format prepares us for all further analyses within the thesis. After reformatting the datasets, it is of interest to model weight measurements as functions of age and sex to understand how children from different datasets grow. This provides insight into various aspects of the available datasets such as median growth and variability within the population of interest over time. Furthermore, these models can be used to create growth reference charts for each of the populations. However, the growth references that we will develop only describe how children grow at a certain time and place, which are not ideal to use as reference populations as they do

not describe optimal growth. Ideally, it is with a composite population of healthy children from a number of locations and settings that infants should be compared.

2. *Describe how real populations of pre-school children from the developing world fit the 2006 WHO growth standard.*

The WHO state that all children grow the same under optimal conditions. The WHO 2006 standard describes optimum growth as children who make up the chart grew up in ideal conditions. Work has been carried out in the past which has investigated how children from the developed world grow relative to the standard. However, little is known about how children from the developing world fit. It is of interest to assess the fit of the WHO standard with the populations made available within our research, allowing us to assess how children from the developing world fit this idealised reference. It is expected that populations of children from poorer backgrounds will sit a number of standard deviations below the standard, while populations from more developed backgrounds will fit well or even outperform the standard. Furthermore, we expect patterns to emerge in which children fall away or catch up with the standard at certain ages.
The growth references we develop and the WHO standard can be used to assess growth with age by plotting serial measurements on the charts, determining whether their weight measurements consistently track along a centile. If a child crosses centile lines downwards, it may be a sign of growth faltering. However, to compare growth in different children, other measures must be used. Measures such as weight velocity and conditional weight gain (introduced in Section 1.2.8) can be used to compare growth with both the reference population and other children.

3. *Describe what impact using an external reference has on some common measures of weight gain and investigate how to modify the measures in order to allow for this.*

Conditional weight gain Z scores have zero mean and unit standard deviation if measurements are converted with a suitable, well matched distance reference. To guarantee these conditions only an 'internal' reference can be used, like the ones we will develop in addressing aim 1. However, it is with an external reference that children should be compared, such as the WHO standard. We aim to develop a new measure of growth for use with an external reference - generalised conditional weight gain.

1.4.2 Testing hypotheses

The remainder of the aims focus on testing the hypotheses developed in this chapter within Section 1.3.4. These hypotheses highlight the presumed associations between anthropometry and adverse outcomes.

4. *Explore the pathways children take through nutritional states over time*

Hypotheses 1 – 3 suggest that relationships exist between various nutritional states:

Hypothesis 1: those who are both wasted and stunted are at a higher risk of death than those who are wasted, who are in turn at a higher risk than those who are stunted

Hypothesis 2: wasting tends to precede stunting

Hypothesis 3: children who become stunted are likely to stay stunted

We test these hypotheses by applying statistical methodology to longitudinal growth datasets in new ways, assessing how children move to nutritional states conditional on previous state.

If children between datasets share similar tendencies of moving between nutritional states, these datasets can potentially be combined. Pooling data together is appropriate when the datasets are homogeneous with respect to the parameters of interest.

5. *Identify the extent to which the data can be merged into a larger dataset*

By developing statistical tests for use on longitudinal datasets, we can determine if sets can be merged. Merging sets gives the analysis used to test hypothesis 4 more power as the sample sizes increase.

6. *Determine whether it is size or growth which best predicts mortality*

Past research suggests both size and growth are significant predictors of mortality. However, papers which have investigated which of these measures is the best predictor have published conflicting results. We aim to address the following hypothesis using a new weighted approach:

Hypothesis 4: recent size is a more useful predictor of later mortality than growth over a recent time period

Standard analyses are swamped by children within the centre of the population distribution. We aim to use a weighted analysis to gain new insights by effectively over-

representing children within the extremes of the distribution, allowing us to determine which predictors are useful for children within the tails of the population distribution.

1.5 Chapter conclusions

Within this chapter we introduced undernutrition and its effects, growth monitoring, the WHO standard and measures of growth. We also reviewed literature which has investigated the use of anthropometric measurements in predicting adverse future outcomes.

Six aims were developed in Section 1.4. These aims focus on getting the data into a common format so that modelling and summarising can take place, identifying how well these datasets fit the WHO standard, adapting the expression for conditional weight gain so that an external reference can be used and finally understanding how anthropometric measurements and indices can be used to predict future outcomes. Hypotheses based on previous research were developed in light of the available literature. It is our aim to identify the extent to which these are true using statistical methods that have not yet been applied in the field. The statistical methods used to tackle these aims are presented within Chapter 2. The datasets that these methods are applied to are introduced within Chapter 3. Statistical analyses are carried out in Chapters 4 – 6.

Chapter 2
Statistical methodology
2.1 Introduction
Chapter 1 introduced undernutrition, growth reference charts, measures of growth and reviewed literature which has investigated how anthropometric measurements can be used to predict future outcomes. Hypotheses were developed based on this literature and aims were established at the end of the chapter. The aims first focus on introducing, summarising and modelling the data. Latter aims focus on testing the hypotheses developed in Section 1.3.4 which highlight the presumed associations between measures of anthropometry and adverse outcomes.

This chapter presents the software and statistical models which are used to address the aims highlighted in the previous chapter. A summary of the proposed methods are found in Section 2.3. Further details of these models are found within the body of the chapter and further discussion of the methods can be found at the end of the corresponding chapters where they are applied.

2.2 Software
All statistical aspects of the thesis were handled using the statistical programming language, R. R is a software environment which is widely used for data management, software development, graphics and data analysis (R Development Core Team 2013). Many of the statistical models applied within this thesis are available in R's core software but an extensive number of libraries (also called packages) are also available. Packages extend the capabilities of R with user-created algorithms and models for specialist

techniques, which can be loaded into the software. Any additional packages used to apply the statistical models outlined in this chapter will be specified.

2.3 Methods summary

This section briefly outlines methods that will be utilised throughout the thesis to address each of the aims outlined in Chapter 1. Refer to each of the sections specified within the summary for a full description of each method.

The methodology described in the summary firstly applies to aims focusing on data formatting and modelling then moves on to methods which address the hypotheses developed in Chapter 1.

2.3.1 Data formatting and modelling

Firstly, datasets will be reformatted into a common format. By reformatting the datasets into a common format, we prepare for all further analyses within the thesis, creating a data dictionary.

1. *Import the three developing world data sets into a common format, model and summarise them.*

After reformatting the data, growth chart modelling can take place. This allows us to gain insights into each of the datasets and allows us to create growth reference charts. A commonly used method used to create growth charts is the LMS method (Cole & Green 1992). This method allows users to model distributions smoothly over time and if necessary, sex. Children's measurements can then be compared with the reference distribution, determining which centile measurements lie on. Methodology for the LMS method is presented in Section 2.5.

In R however, this method is commonly applied via Generalised Additive Models for Location Scale and Shape (GAMLSS) (Akantziliotou et al. 2002). These models are semi-parametric regression models where the central tendency of the response variable can be modelled as a function of the explanatory variables. The framework can be extended to model not only the central tendency, but also dispersion and two shape parameters e.g. skewness and kurtosis.

A special case of the distribution is the three parameter Box-Cox Cole and Green (BCCG), also known as the Box-Cox normal distribution (modelling location, variability and skewness). Assuming this distribution in the GAMLSS framework is the equivalent of the

LMS method. The theory of application of the LMS method within the GAMLSS framework can be found in Section 2.6. The $gamlss$ package in R (Rigby & Stasinopoulos 2005) contains all functions required for application of the modelling framework.

The 'internal' reference charts we will develop via GAMLSS only describe the growth of children at a certain time and place and, therefore, care is needed when using these references. The WHO advise universal use of the WHO 2006 standards, which were developed from healthy children from a number of locations around the world. Work has been carried out by Roelants (2013), Hui et al. (2008) and Wright et al. (2008), who determined that children from the developed world either grow on par with the standard or outperform it. However, little is known about how children from the developing world fit the standard. Describing how real populations of pre-school children from the developing world fit the WHO 2006 growth standard is our second aim.

2. *Describe how real populations of pre-school children from the developing world fit the 2006 WHO growth standard.*

We expect the average rate of growth to differ between our cohorts and the WHO standard. To quantify the difference in growth rates, the WHO LMS model will be used to convert children's weight measurements for each of our datasets to Z scores, resulting in a transformed series. Assuming a normal distribution, we model the mean and standard deviation of this series, which describes measurements relative to the standard. These parameters, $\mu_{Z_t}$and σ_{Z_t}, can be smoothly modelled using GAMLSS. The estimates can then be plotted to assess relative growth.

As discussed, both the internal references which we will develop and the WHO standard can be used to assess child growth over time (growth velocity) by plotting serial measurements on the charts. If children's measurements cross centiles downwards, this implies poor growth. Using charts to assess growth velocity however does not allow comparisons with other children. Other measures such as growth velocity or conditional weight gain allow comparisons with both the reference distribution and other infants. Our third aim is:

3. *Describe what impact using an external reference has on some common measures of weight gain and investigate how to modify the measures in order to allow for this.*

Conditional weight gain was introduced as a method which allows relative comparison of growth with both the reference population and other children expressed as a Z score,

conditional on the first measurement. Z scores are distributed as $Z_X \sim N(0,1)$ if a well matched reference is used to convert measurements, e.g. using an 'internal' reference population. However, as discussed, there are problems in using internal reference populations to assess growth. Using the WHO standard as an external reference is a better alternative but the assumptions underlying conditional weight gain do not hold (Z scores are not distributed as $Z_X \sim N(0,1)$). To address this, the expression for conditional weight gain needs to be extended to include $\mu_{Z_t} \neq 0$ and $\sigma_{Z_t} \neq 1$. Methodology in Section 2.7 investigates conditional weight gain and subsequently explores development of a generalised measure allowing the use of an external reference. The generalised measure also takes RTTM into account. In fact, the amount of RTTM is synonymous with correlation when working on the SD scale (Cole 1995), therefore, a way to quantify the amount of RTTM is to quantify this correlation. Correlations can be calculated between groups of measurements at scheduled measurement ages, but statistical models are needed to interpolate correlations between these ages. Cole (1995) adopted fractional polynomials (FP) (see Royston & Altman (1994)), a flexible alternative to a regular polynomial to model the correlation matrix of the Cambridge Infant Growth study (Whitehead et al. 1989). Argyle (2002) developed the Argyle model, a model based on linear regression used to interpolate correlations in a dataset from Newcastle, England. We aim to apply both of these models to our datasets to assess their performance as well as two others, our own FP models and Generalised Additive Models (GAM). FP models can be applied using the mfp package in R. The function used to apply fractional polynomials uses a selection algorithm to choose the best fitting model, described in more detail in Section 2.8.3. GAMs (Wood 2006) are non-parametric alternatives which allow the user to specify the degree of smoothness. These can be applied using the $mgcv$ package in R. Further details on the models and packages are in Section 2.8.1.

2.3.2 Testing hypotheses

After addressing aims which focus on data formatting and modelling, the primary objective of the thesis can be addressed – determining how anthropometry can be used to predict adverse future outcomes. The review in Chapter 1 examined research which has previously investigated this topic. Hypotheses were developed based on the evidence. Aim 4 focuses on testing hypotheses 1 to 3.

4. Explore the pathways children take through nutritional states over time

We adopt a new approach to assess the presumed associations outlined in Section 1.3.4, using stochastic models. The approach allows modelling of a discrete sample space, takes an element of time into consideration and models the conditional probability of moving from one state to another.

The definitions in Section 1.2.7 can be used to categorise children into nutritional states. We use these definitions to categorise children into dead (D), wasted-stunted (WS), wasted (W), stunted (S) and healthy (H) states at various time points. We use a BMI threshold to define wasted, the rationale for which was explained in Section 1.2.7. The probabilities of moving from each state to all others can be represented in a 5 by 5 matrix. Therefore, we assess the probability of moving state conditional on previous state. The methodology required for this presented is in Section 2.10.

5. Identify the extent to which the data can be merged into a larger dataset

We propose that if the underlying stochastic model for each dataset is the same, then the datasets can be pooled. This is because the transition probabilities of moving from one state to another will not be significantly different from one another. A Generalised Likelihood Ratio Test (GLRT) is appropriate for use under these circumstances. The development of this test is presented in Section 2.10.1. Pooling datasets provides more power to the analysis used for aim 6.

6. Determine whether it is size or growth which best predicts mortality

Research papers which have investigated whether recent size or growth over a recent time period is the best predictor of time until death have published conflicting results. This aim focuses on determining which measure is more valuable using a new approach. Cox proportional hazards models are a standard approach which can be used to predict mortality in the future. However, standard models are dominated by children within the centre of the population distribution. By using a weighted approach we can effectively over represent children within the tails of the distribution, allowing us to determine which predictors are most valuable for children who are outwith the normal range. Cox proportional hazards models can be applied using the $coxph$ package in R. Weights in Cox proportional hazards models are simply replicates of the data. Therefore, a weighting function is required. The weighting function is presented in Chapter 6. However, survival analysis and Cox proportional hazards models are presented in Section 2.11.

2.4 Z scores

The concept of Z scores (or SD scores) underpins some key methods used throughout this thesis. A Z score provides an estimate of where measurement X lies relative to its population mean, μ_X, in terms of the population SD, σ_X. The measurement X is, therefore, expressed as the number of SDs from the population mean.

Let:

$$Z_X = \frac{X - \mu_X}{\sigma_X} \tag{2.1}$$

If the raw measurement, X, is normally distributed and estimates of the mean, μ_Z, and SD, σ_Z, are from the same population from which X derives, then we can assume that, approximately, $Z_X \sim N(0, 1)$. Z_X can be converted to a centile using the distribution function of the standard normal distribution if the variance is known, or the t-distribution if it is estimated. This provides an estimate of the proportion of the population below or above X.

If estimates of the mean and SD do not derive from the same population as X, then $\mu_Z \neq 0$ and $\sigma_Z \neq 1$. These will, therefore, need to be estimated. In this case Z_X is t-distributed.

The situation is usually more complicated than this, since the raw measurement is not normally distributed in the population, Z has to be further adjusted to eliminate skewness and kurtosis and hence produce a measure that is approximately normal (see Section 2.5).

2.5 The LMS method

The LMS method allows users to create growth reference charts, modelling the entire distribution taking into account the Box-Cox power (L), median (M), and coefficient of variation (S). The LMS method involves transforming y at age t by the Box-Cox transformation such that:

$$Z = \frac{\left(\frac{y}{\mu}\right)^{\lambda} - 1}{\lambda\sigma}, \qquad \lambda \neq 0 \tag{2.2}$$

or

$$Z = \frac{\log\left(\frac{y}{\mu}\right)}{\sigma}, \qquad \lambda = 0 \tag{2.3}$$

where Z has distribution $N(0,1)$. The aim is to make the distribution symmetric, so any skewness in y is removed by a suitable choice of λ, and therefore, the mean of Z corresponds to μ, the median of y.

We assume that λ, μ and σ vary with t. These can be read off the smooth curves $L(t), M(t)$ and $S(t)$, so we can say:

$$Z(t) = \frac{\left[\frac{y(t)}{M(t)}\right]^{L(t)} - 1}{L(t)S(t)}, \qquad L(t) \neq 0 \tag{2.4}$$

or

$$Z(t) = \frac{\log\left(\frac{y(t)}{M(t)}\right)}{S(t)}, \qquad L(t) = 0 \tag{2.5}$$

If we rearrange these, we can show that the centile 100α of y at t is given by:

$$C_{100\alpha}(t) = M(t)(1 + L(t)S(t)Z_\alpha)^{1/L(t)}, \qquad L(t) \neq 0 \tag{2.6}$$

or

$$C_{100\alpha}(t) = M(t)\exp(S(t)Z_\alpha), \qquad L(t) = 0 \tag{2.7}$$

where Z_α is the normal equivalent deviate of size α (Cole & Green 1992).

2.6 Applying the LMS method

Generalised Additive Models for Location Scale and Shape are semi-parametric models which allow modelling of up to four distributional parameters. This model, applied with the package $gamlss()$ in the statistical programming R, allows us to apply the LMS method as well as its extensions (e.g. distributions that take kurtosis into account). As this is a semi-parametric method which uses properties of both parametric and non-parametric models, the terms parametric and non-parametric are firstly introduced within Sections 2.6.1 and 2.6.2 before presenting GAMLSS in Section 2.6.3.

2.6.1 Parametric models

A linear model is a parametric model which predicts response variable y, by explanatory variables $x_1, \dots, x_n$. Linear models assume the identity link function $g(.)$, which provides a 1:1 relationship between the mean μ and the linear predictors such that $g(\mu) = X\beta$.

This model is in the form:

$$E(y) = \mu = g^{-1}(X\beta) = g^{-1}(\beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n) \tag{2.8}$$

Linear models assume that the response variable y is normally distributed. They also require independence of observations and constant variance.

In fact, linear regression is a special case of a Generalised Linear Model (GLM). GLMs are in the same form as (2.8) but GLMs allow the link function $g(.)$ to take any form which transforms the expectation of the response $E(y) \equiv \mu$ to the linear predictor $\beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n$. For example, we may want to use the logit transformation in the case of a binomial response i.e. $g(\mu) = \ln\left(\frac{\mu}{1-\mu}\right)$, constraining μ between 0 and 1. GLMs are also parametric in that the response variable comes from any distribution within the exponential family, for example, binomial, Poisson or gamma. Furthermore, GLMs relax the normality and homoscedasticity assumptions. However, there is substantial precision loss if the distribution is specified incorrectly.

2.6.2 Non-parametric models

Additive Models (AMs) are non-parametric models which replace the terms $\beta_1 x_1, \ldots, \beta_n x_n$ of a linear regression model with smooth functions $h_1(x_1), \ldots, h_n(x_n)$. The model is in the form:

$$E(y) = \mu = \beta_0 + h_1(x_1) + \cdots + h_n(x_n) \tag{2.9}$$

Generalised Additive Models (GAMs) combine elements of GLMs and AMs, allowing a link function $g(.)$ such that:

$$E(y) = \mu = g^{-1}\big(\beta_0 + h_1(x_1) + \cdots + h_n(x_n)\big) \tag{2.10}$$

These unspecified non-parametric functions provide a smooth fit with relaxed assumptions about the relationship between the response and the predictor.

GAM models tend to fit data better than GLMs, however, they also tend to over fit and are computationally expensive. GAMs can include both linear and additive terms and are therefore in the realm of semi-parametric models. Furthermore, the relationships between the response and predictors are harder to interpret.

2.6.3 Generalised Additive Models for Location Scale and Shape

GAMLSS models are semi parametric regression models which were introduced by Akantziliotou, Rigby & Stasinopoulos (2002) and Rigby & Stasinopoulos (2005). GAMLSS models are parametric in that they require a distributional assumption for the response and semi parametric in the sense that modelling the parameters of the distribution may involve using non-parametric smoothing functions.

GAMLSS assume independence for observations y with probability density function (PDF) $f(y|\mu, \sigma, \lambda, \tau)$ conditional on the four distributional parameters μ (location), σ (scale), λ (shape) and τ (shape).

The model formulation is as follows. We have monotonic link functions $g_k(.)$ where $k = 1, 2, 3, 4$, relating the distribution parameters to explanatory variables by:

$$\mu = g_1^{-1}\left(X_1\beta_1 + \sum_{j=1}^{J_1} h_{j1}(x_{j1})\right)$$

$$\sigma = g_2^{-1}\left(X_2\beta_2 + \sum_{j=1}^{J_1} h_{j2}(x_{j2})\right)$$

$$\lambda = g_3^{-1}\left(X_3\beta_3 + \sum_{j=1}^{J_1} h_{j3}(x_{j3})\right)$$

$$\tau = g_4^{-1}\left(X_4\beta_4 + \sum_{j=1}^{J_4} h_{j4}(x_{j4})\right)$$

where $X_k\beta_k$ are linear terms and $\sum_{j=1}^{J_k} h_{jk}(x_{jk})$ are additive terms in the model.

One particular case of the GAMLSS distribution is the BCCG or Box-Cox normal distribution. When assuming this distribution, it is the equivalent of the LMS method. This distribution is only a three parameter distribution, using the identity link function for the median μ (although the log link is often better as children cannot have negative size) and the Box-Cox power λ. The log link is used for the coefficient of variation σ.

The untruncated BCCG function $f(y|\mu, \sigma, \lambda)$ can be defined as,

$$f\big(y\big|M(t), S(t), L(t)\big) = \frac{1}{S(t)\sqrt{2\pi}} \cdot \frac{Y^{(L(t)-1)}}{M(t)^{L(t)}} \cdot \exp\left(-\frac{Z^2}{2}\right) \qquad (2.11)$$

where if $L(t) \neq 0$ then $Z = \left(\left[\frac{y}{M(t)}\right]^{L(t)} - 1\right)/L(t)S(t)$, else $Z = \log\left(\frac{y}{M(t)}\right)/S(t)$, for $y > 0, M(t) > 0, S(t) > 0$ and $L(t)$ has no limits. The truncated version adjusts (2.11) for the condition $y > 0$. If $Z \sim N(0,1)$ then $Y \sim BCCG\big(M(t), S(t), L(t)\big)$.

The log likelihood function is (apart from the constant),

$$l = l(L, M, S) = \sum_{i=1}^{n} (L(t_i) \log\left(\frac{y_i}{M(t_i)}\right) - \log\big(S(t_i)\big) - \frac{1}{2}Z_i^2) \qquad (2.12)$$

Estimates of L M and S can be found by maximising the penalised likelihood function,

$$l - \frac{1}{2}\alpha_\lambda \int [L''(t)]^2 dt \; - \frac{1}{2}\alpha_\mu \int [M''(t)]^2 dt \; - \frac{1}{2}\alpha_\sigma \int [S''(t)]^2 dt \qquad (2.13)$$

where $\alpha_\lambda, \alpha_\mu, \alpha_\sigma$ are the smoothing parameters. The integrals provide roughness penalties according to the squared second derivatives of the L, M and S curves.

Maximising this function will balance fit and smoothness of the curves. See Rigby & Stasinopoulos (2005) for further information on how this penalised likelihood function is derived.

Other distributions are available within the GAMLSS framework. For example, the Box-Cox power exponential (BCPE) distribution (which also takes kurtosis into account, which was considered by the WHO. The WHO also considered considered the Box-Cox t, Johnson's SU and the modulus-exponential normal distribution (de Onis & WHO Multicentre Growth Reference Study Group 2006).

The centile curve modelling procedure carried out in this thesis demonstrates the use of GAMLSS and how it is applied to real life data, allowing development of internal references or growth standards.

2.7 Weight gain Z score

Growth increments are the absolute difference in size between two time points. Growth velocity is the difference divided by the length of the interval. Weight gain can be expressed as a Z score, which is a valid alternative for growth increment or growth velocity (Roelants 2013), using the properties of Z scores to express the change relative to the population in which the child originates from, which can be expressed as a centile. For example, imagine a child measured at time t_1 who lies on the 50th ($Z = 0$) centile, which is the average weight for a child of that age. The child is then measured again at time t_2 and has gained weight but now lies on the 40th centile ($Z = -0.25$). The child has gained weight but has dropped 10 centiles, meaning that the child now weighs more than they did at t_1, but less than average for those aged t_2.

By calculating a change in Z score between two time points and dividing by the standard deviation of the change, a Z score for the change in Z scores can be assessed. The resulting value allows us to determine what proportion of the population are likely to gain at least as much weight as the child in question. By conditioning on initial weight, the fact that light infants tend to grow faster than heavy infants can be taken into account (Cole 1995).

Current published methodology allows us to obtain conditional weight gain Z scores while using an internal reference. However, there are problems in using internal references to assess growth as they are developed from a sample of children and may be biased. Ideally, the WHO standard (or an idealised reference) should be used to assess

growth as children who make up the standard are considered healthy. However, the assumptions for conditional weight gain break down while using external references or standards. In this section we introduce unconditional weight gain, conditional weight gain and then derive a generalised measure for use with growth standards.

2.7.1 Unconditional weight gain

Unconditional weight gain is a measure of weight gain that is not adjusted for (i.e. conditioned on) the initial weight.

The method (Cole 1997) consists of converting two successive weights to Z scores. The Z scores are calculated while referring to an 'internal' reference population.

Let $X = \Delta Z = Z_2 - Z_1$. To convert X to a SD score we use the following transformation, as described in (2.1):

$$Z_X = \frac{X - \mu_X}{\sigma_X}$$

where $\mu_X = E(Z_2 - Z_1)$ and $\sigma_X = \sqrt{Var(Z_2 - Z_1)}$. The mean and standard deviation are:

$$\begin{aligned} \mu_X &= E(Z_2 - Z_1) \\ &= E(Z_2) - E(Z_1) \\ &= \mu_{Z_2} - \mu_{Z_1} \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} \sigma_X &= \sqrt{Var(Z_2 - Z_1)} \\ &= \sqrt{\sigma_{Z_2}^2 + \sigma_{Z_2}^2 - 2\sigma_{Z_1 Z_2}} \\ &= 2 - 2r \\ &= 2(1 - r) \end{aligned}$$

where $r = cor(Z_1, Z_2)$.

(Note since $\sigma_{Z_1} = \sigma_{Z_2} = 1$ and $r = \sigma_{Z_1 Z_2}/(\sigma_{Z_1}\sigma_{Z_2})$ then $r = \sigma_{Z_1 Z_2}$)

To convert X to a Z score we use Equation (2.1):

$$\begin{aligned} Z_X &= \frac{\mathrm{X} - 0}{\sqrt{2(1 - r)}} \\ &= \frac{\mathrm{X}}{\sqrt{2(1 - r)}} \\ &= \frac{Z_2 - Z_1}{\sqrt{2(1 - r)}} \end{aligned} \qquad (2.14)$$

2.7.2 Conditional weight gain

Unconditional weight gain ignores the correlation between weight gain and initial weight. Therefore any 'change in centile' would be attributed to the change in weight of that child, assuming that no natural phenomenon would influence that change. However, this is not the case due to RTTM.

The problem with unconditional weight gain is that if a child moves centile towards the population median, the change in Z score is completely attributable to their change in weight. In fact their movement towards the population median may be expected, due to RTTM. Conditional weight gain accounts for this expected movement, whereas unconditional weight gain would explain this expected movement as crossing centiles. The amount of expected tracking is actually synonymous with the amount of correlation between Z_1 and Z_2 such that the expected value of Z_2 is $r.Z_1$ (Cole 1995) (see proof below). In the unconditional weight gain case perfect tracking would require that $Z_2 = Z_1$, which would require the correlation between Z_2 and $Z_1$to equal 1, which is impossible.

To derive conditional weight gain we use Equation (2.1), letting $X = Z_2 - E(Z_2|Z_1)$, $\mu_X = E(Z_2 - E(Z_2|Z_1))$ and $\sigma_X = \sqrt{Var(Z_2 - E(Z_2|Z_1))}$.

Firstly, start with the linear regression model:

$$E(Z_2|Z_1) = \alpha + \beta Z_1 \tag{2.15}$$

which implies that

$$E(Z_2) = E\{E(Z_2|Z_1)\} = \alpha + \beta E(Z_1)$$

Similarly,

$$E(Z_1 Z_2) = E\{Z_1 E(Z_2|Z_1)\} = E\{\beta Z_1^2\} = \beta E(Z_1^2)$$

Therefore,

$$\begin{aligned} \sigma_{Z_1 Z_2} &= E(Z_1 Z_2) - E(Z_1)E(Z_2) \\ &= \beta E(Z_1^2) - \beta[E(Z_1)]^2 \\ &= \beta \sigma_{Z_1}^2 \\ r &= \frac{\sigma_{Z_1 Z_2}}{\sigma_{Z_1}\sigma_{Z_2}} = \beta \frac{\sigma_{Z_1}}{\sigma_{Z_2}} \end{aligned} \tag{2.16}$$

however since $\sigma_{Z_2} = \sigma_{Z_1} = 1$, then $r = \beta$.

Following on from Equation (2.15), using results from Equation (2.16),

$$E(Z_2|Z_1) = \alpha + \beta Z_1 = 0 + rZ_1 = r.Z_1$$

since $E(Z_2) = E(Z_1) = 0$, therefore $\alpha = 0$.

Therefore $X = Z_2 - E(Z_2|Z_1) = Z_2 - rZ_1$.

The mean and standard deviation are:

$$\begin{aligned}\mu_X &= E(Z_2 - r.Z_1)\\ &= \mu_{Z_2} - r\mu_{Z_1}\\ &= 0\end{aligned}$$

and

$$\begin{aligned}\sigma_X &= \sqrt{Var(Z_2 - rZ_1)}\\ &= \sqrt{\sigma_{Z_2}^2 + r^2\sigma_{Z_2}^2 - 2.r.\sigma_{Z_1Z_2}}\\ &= \sqrt{1 + r^2 - 2r^2}\\ &= \sqrt{1 - r^2}\end{aligned}$$

So,

$$\begin{aligned}Z_X &= \frac{X - \mu_X}{\sigma_X}\\ &= \frac{Z_2 - rZ_1 - 0}{\sqrt{(1 - r^2)}}\\ &= \frac{Z_2 - rZ_1}{\sqrt{(1 - r^2)}}\end{aligned} \qquad (2.17)$$

The expression above takes into account RTTM and therefore the Z score for the change in Z scores takes into account the expected amount of movement towards the median, conditional on Z_1.

The predicted value of Z_2, $E(Z_2|Z_1)$, is assumed to be normally distributed. The correlation coefficient is always less than or equal to one which means that $E(Z_2|Z_1) = rZ_1$ will always be expected to be closer to the median than Z_1.

2.7.3 What is an 'external' reference?

When using an internal reference, the measurement of interest can be compared with its own parent population which allows us to compare all children with those who are living under the same socio-economic conditions. When using an external reference we compare children's weights with a reference or standard which the child does not originate from. An expression for conditional weight gain while using an external reference has not previously been published.

2.7.4 Generalised conditional weight gain

When using an external reference, conditional weight gain is not as simple an expression as in Section 2.7.2 as many of the assumptions for Z scores no longer hold. The mean of

the Z scores will not necessarily equal zero and the standard deviations may not equal one. This means that the ratio of Z scores will not equal one, therefore $\beta \neq r$ and $\alpha \neq 0$. Equation (2.15) still holds. However, $\sigma_{Z_2}/\sigma_{Z_1} \neq 1$. Let $\sigma_{Z_2}/\sigma_{Z_1} = k$, therefore,

$$\beta = r\frac{\sigma_{Z_2}}{\sigma_{Z_1}} = rk$$

Furthermore, since $E(Z_2) \neq E(Z_1) \neq 0$, we yield,

$$E(Z_2|Z_1) = \alpha + \beta Z_1 = \alpha + rkZ_1 \tag{2.18}$$

To derive an expression which generalises conditional weight gain, we use Equation (2.1) while referring (2.18). Let:

$$X = Z_2 - E(Z_2|Z_1) = Z_2 - \alpha + rkZ_1,$$

$$\mu_X = E(Z_2 - E(Z_2|Z_1)) = E(Z_2 - \alpha + rkZ_1)$$

$$\sigma_X = \sqrt{Var(Z_2 - E(Z_2|Z_1))} = \sqrt{Var(Z_2 - \alpha - rkZ_1)}$$

The mean and standard deviation are:

$$\begin{aligned}\mu_X &= E(Z_2 - (\alpha + \beta Z_1)) \\ &= E(Z_2 - \alpha - rkZ_1) \\ &= \mu_{Z_2} - \alpha - rk\mu_{Z_1}\end{aligned}$$

and,

$$\begin{aligned}\sigma_X &= \sqrt{Var(Z_2 - \alpha - rkZ_1)} \\ &= \sqrt{\sigma_{Z_2}^2 + \left(r\frac{\sigma_{Z_2}}{\sigma_{Z_1}}\right)^2 \sigma_{Z_1}^2 - 2r\frac{\sigma_{Z_2}}{\sigma_{Z_1}} r.\sigma_{Z_1}.\sigma_{Z_2}} \\ &= \sqrt{\sigma_{Z_2}^2 + r^2\frac{\sigma_{Z_2}^2}{\sigma_{Z_1}^2}.\sigma_{Z_1}^2 - 2r^2\sigma_{Z_2}^2.\frac{\sigma_{Z_1}}{\sigma_{Z_1}}} \\ &= \sqrt{\sigma_{Z_2}^2 + r.\sigma_{Z_2}^2 - 2r^2\sigma_{Z_2}^2} \\ &= \sqrt{\sigma_{Z_2}^2 - r^2\sigma_{Z_2}^2} = \sqrt{\sigma_{Z_2}^2(1-r^2)} = \sigma_{Z_2}\sqrt{(1-r^2)}\end{aligned}$$

So,

$$\begin{aligned}Z_X &= \frac{X - \mu_X}{\sigma_X} \\ &= \frac{Z_2 - \alpha - rkZ_1 - \mu_{Z_2} + \alpha + r.k.\mu_{Z_1}}{\sigma_X}\end{aligned}$$

$$Z_X = \frac{Z_2 - \mu_{Z_2} - r\frac{\sigma_{Z_2}}{\sigma_{Z_1}}(Z_1 - \mu_{Z_1})}{\sigma_{Z_2}\sqrt{(1-r^2)}} \tag{2.19}$$

In the event that the SDs for Z_2 and Z_1 are equal (i.e. $\sigma_{Z_1} = \sigma_{Z_2}$), the slope parameter $\beta = rk = r\frac{\sigma_{Z_2}}{\sigma_{Z_1}}$ is simply equal to the correlation between the two SD scores, r. Letting $\mu_{Z_1} = \mu_{Z_2} = 0$ and $\sigma_{Z_1} = \sigma_{Z_2} = 1$, we obtain;

$$Z_X = \frac{Z_2 - rZ_1}{\sqrt{1 - r^2}}$$

which is equal to the expression for internal conditional weight gain. Equation (2.19) is the result of aim 3 of this thesis.

2.8 Modelling correlation

As discussed, the correlation coefficient, r, is needed to adjust for RTTM. However, to interpolate the correlation between scheduled measurement dates, where no data are available, we need to build a statistical model.

While obtaining correlation coefficients, one problem is that the sample size to calculate each coefficient for each pair of time points may differ and as a result the variation for each correlation may differ. The confidence bands surrounding each correlation will therefore vary in size. As the number of pairs of measurements used to calculate each correlation coefficient increases, the variance decreases, giving a better estimate of the true correlation.

With correlations for $n > 500$, correlation variance can be expressed as:

$$Var(r) = \frac{(1 - r^2)^2}{n}$$

It is unwise to use this expression if $n < 500$, as r tends to normality too slowly (Kendall & Stuart 1979). In our case it is better to use Fisher's Z transformation of the correlation coefficient (Fisher 1921). The reason for using this transformation is because as the true population parameter $|\rho|$ gets closer to 1, the variance of r decreases. The Fisher transformation is an approximate variance stabilising transformation, meaning the variance of the transformed correlation is approximately constant for all values of the population correlation coefficient.

Let $Z(r) = \phi(r)$ to avoid any confusion with Z scores. We transform each correlation for pair t_1, t_2 with:

$$\phi(r) = \frac{1}{2}\log\left(\frac{1 + r}{1 - r}\right) \tag{2.20}$$

The associated variance with each correlation coefficient becomes:

$$Var\big(\phi(r)\big) = \frac{1}{n - 3}$$

Once the transformed data are modelled, the predicted values can be converted back using:

$$r = \frac{\exp\big(2\phi(r)\big) - 1}{\exp\big(2\phi(r)\big) + 1} \tag{2.21}$$

Fisher's transformation will be used to transform correlations for three of the four models within this thesis. The other transformation is a log transform, used by the 'Argyle' model. An approximate confidence interval for a correlation coefficient is:

$$CI_{\phi_{r(1-\alpha)}} = \phi_r \pm Z_{\frac{\alpha}{2}}\sqrt{\frac{1}{n-3}}$$

where ϕ_r is the Fishers transformation of r (Argyle 2002). Therefore, the larger n is, the smaller the interval will be.

2.8.1 Generalised Additive Models

Bivariate splines are one of the modelling approaches used within this thesis. A popular approach which we can use to fit bivariate splines are GAMs. GAM models were developed by Hastie & Tibishirani (1986) to combine the properties of both Generalised Linear Models (GLMs) with Additive Models (AMs). GAMs can be applied with the $mgcv$ package in R, written by Simon Wood. Wood (2006) illustrates how GAMs can be represented using penalised regression splines.

The model's structure may take the form, for example:

$$g(\mu) = \beta_0 + h_1(x_1) + \cdots + h_n(x_n) \tag{2.22}$$

where

$$E(Y) = \mu \text{ and } Y \sim exponential\ family\ distribution$$

Equation (2.22) can also include linear terms. For each smooth function h_i, we must set the maximum number of basis functions, k, which determines the flexibility of our smooth. If k is too small we will over smooth, and if k is too large it will result in an over complicated model. Effectively, setting k equal to the number of datapoints will track the data perfectly, but ideally we would like our spline estimate of the data to be as close as possible with the minimum amount of basis functions.

Our response in this distinct case is Fishers Z transform ($\phi(r_{t_i,t_j})$) of the original correlation coefficients of the groups of Z scores of weight measurements between time points.

Our link function $g(.)$ is the identity link, which is simply a 1:1 function where $g\left(\phi(r_{t_i,t_j})\right) = \phi(r_{t_i,t_j})$. Our GAM will take the form:

$$E\left(\phi\left(r_{t_i,t_j}\right)\right) = \beta + h(t_i, t_j)$$

where we have intercept β. Row i of the vector of time points t_i and t_j are the vector time points and h is our two dimensional smooth for t_i and t_j.

One issue with the use of a spline based approach is the smoothness of our model. As discussed, we wish to choose k such that it captures most of the variation in our response, however, we do not wish to over fit as this can be computationally expensive. Choosing k is arbitrary, however, two criteria which are available to aid us in choosing our smoothing parameter is Generalised Cross Validation (GCV) and Akaike Information Criterion (AIC) (Refer to Section 2.8.3).

GCV was first introduced by Craven & Wahba (1979). This criterion allows us to select the smoothness parameter by minimising the value of:

$$GCV = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i - \hat{y}_i}{1 - \frac{1}{n}tr(H)}\right)^2$$

where

$$H = X(X^T X)^{-1}X^T$$

In this case H is the hat matrix where X is our matrix of explanatory variables (and contains the basis of the smooths) and $\hat{y}_i$ is the estimate of y_i from fitting all the data (Wood 2006).

For each correlation matrix we will assess the fit of a number of GAM models based on both the GCV and AIC, with the intention of quantifying the most amount of variation but by keeping k low. These models will firstly be fitted to the datasets and the best fitting models will be used within our cross validation approach.

Wood (2006) states "the exact choice of k is not critical, it should be large enough that you are reasonably sure of having enough degrees of freedom to represent the underlying 'truth' reasonably well, but small enough to maintain reasonable computational efficiency".

2.8.2 Argyle model

A very simple model to capture the trend and variation within correlation matrices was developed by Argyle (2002). Some of Argyle's early work started by regressing different

functions of age and intervals between ages on the correlation coefficient. This led Argyle to a very simple model form. Argyle (2002) stated that as a result of her work, a plausible model for correlation could be a function of the form:

$$r_{t_i t_j} = \rho^{f(t_i, t_j)}$$

where r is the actual correlation in row i and column j of a correlation matrix. ρ is an unknown constant and $f(t_i, t_j)$ is a function of t_i and t_j where $t_i < t_j$. However, taking the log of each side yields:

$$log\left(r_{t_i t_j}\right) = f(t_i, t_j)\log(\rho)$$

such that the logarithm of the correlation between time points i and j could be modelled as a function of t_i and t_j. This requires the correlation to be positive.

By using linear regression, Argyle assessed different models by regressing functions of t_i and t_j on $\log\left(r_{t_i t_j}\right)$. Transformations such as the square root, the difference and the reciprocals of t_i and t_j were assessed however Argyle arrived at the form:

$$log\left(r_{t_i t_j}\right) = Alog(t_i + 1) + Blog(t_j + 1)$$

where the addition of 1 to t_i and t_j to avoid the complications of logging 0 where birth weights are included within the correlation matrix.

This form of the model fitted well. However, Argyle decided to use a constant, τ, instead of simply adding 1. Adding 9 (if units are months) seemed like a reasonable approach since length of gestation is approximately 9 months, so adding this constant would result in gestational age. This model was not as good a fit as the previous model, and was subsequently abandoned.

$$log\left(r_{t_i t_j}\right) = Alog(t_i + \tau) + Blog(t_j + \tau)$$

Argyle (2008) developed the model further, by letting $A = \lambda$ and $B = -\lambda$ such that the model's form becomes:

$$log\left(r_{t_i t_j}\right) = \lambda log\left(\frac{t_i + \tau}{t_j + \tau}\right)$$

$$= log\left(\frac{t_i + \tau}{t_j + \tau}\right)^{\lambda}$$

The final form of the model can be expressed by taking the exponential of each side of the formula, which then yields:

$$r_{t_i t_j} = \left(\frac{t_i + \tau}{t_j + \tau}\right)^{\lambda} \tag{2.23}$$

for $i < j$.

Argyle optimised τ by grid search, plotting the deviance against τ. To explore how changing the constants in her models would affect the fit, Argyle (2002) let the constants vary between ranges of values in steps of 0.5. The resulting deviance (which is proportional to the sum of squares) and coefficients were plotted against the constant term.

This process will be repeated for our analysis. A linear model in the form:

$$E\left(log\left(r_{t_i t_j}\right)\right) = \lambda\, log\left(\frac{t_i + \tau}{t_j + \tau}\right)$$

will be applied while varying τ over a range of values. The constant τ which results in the minimum deviance can then be taken with the associated coefficient (λ) of $log\left(\frac{t_i+\tau}{t_j+\tau}\right)$.

2.8.3 Fractional polynomials

A fractional polynomial is a flexible alternative to a regular polynomial. Low order polynomials are known to not always fit data well, while high order polynomials fit better but still fit badly at the extremes of X (Royston & Altman 1994). Furthermore, as polynomials do not have asymptotes they do not fit data well where limiting behaviour is expected.

A fractional polynomial of degree m is defined as:

$$\Phi_m = (X; \xi, p) = \xi_0 + \sum_{j=1}^{m} \xi_j X^{(p_j)} \tag{2.24}$$

where $\xi = (\xi_0, \xi_1, \dots, \xi_m)$ are real valued coefficients, X is a positive covariate and $p = (p_1, p_2, \dots, p_m)$ is a real valued vector of powers where $p_1 < p_2 < \cdots < p_m$. The bracket notation signifies the Box-Tidwell transformation where,

$$X^{(p_j)} = \begin{cases} X^{p_j} & if\ p_j \neq 0 \\ \ln X & if\ p_j = 0 \end{cases}$$

A stepwise procedure is used for covariate selection and model fitting when several covariates are available.

In the statistical program R, the package mfp is used to implement fractional polynomials, using a form of backward elimination for model selection.

mfp processes each predictor in turn, arranging predictors in order of increasing p-value so that important variables are modelled first. The best fitting fractional polynomial function for each predictor is determined using a selection algorithm (while assuming all other variables are linear). The functional form (but not the estimated coefficients) for each predictor is kept and the process is repeated for other predictors. The fractional polynomial functions are reassessed by re-cycling each predictor in the algorithm until the overall model does not change (convergence is usually achieved within 4 cycles). To determine which function is used for each predictor, mpf starts with the most complex fractional polynomial model available and attempts to simplify it by reducing the degrees of freedom. A selection algorithm called the "closed test procedure" inspires the selection algorithm for mfp in which the p-value is maintained over a series of tests at a pre-specified nominal value such as 0.05. The algorithm with a maximum of 4 degrees of freedom for a single continuous parameter is as follows:

1. Inclusion: test the fractional polynomial in x for possible omission of x (4 df test). If x is significant, continue, otherwise drop x from the model.
2. Non-linearity: test the fractional polynomial in x against a straight line in x (3 df test). If significant, continue, otherwise the chosen model is a straight line.
3. Simplification: test the fractional polynomial with $m = 2$ against the best fractional polynomial with $m = 1$. If significant, choose $m = 2$, otherwise choose $m = 1$ (Ambler & Royston 2001).

To determine whether a predictor is significant, a p-value is calculated based of a difference of deviances having either a χ^2 or F distribution depending on the regression in use. The procedure uses a pre-transformation scaling to avoid numerical problems.

2.8.4 Cole model

Cole (1995) adopted the FP approach to model the correlation structure for the Cambridge infant study. Cole (1995) used the mean $\bar{t} = \frac{t_j + t_i}{2}$ and the difference $\Delta t = t_j - t_i$ between time points where $t_j > t_i$ as explanatory variables in the model. Cole's model ended up with 6 terms, shown below:

$$E\left(\phi\left(r(t_i, t_j)\right)\right) = \beta_0 + \beta_1 \log(\bar{t}) + \beta_2 \log(\Delta t) + \beta_3 \left(\frac{1}{\Delta t}\right) + \beta_4 \log(\bar{t}) \log(\Delta t) + \beta_5 [\log(\bar{t})^2] \quad (2.25)$$

This model performed very well on the Cambridge data with an $R^2(adj.)$ value of 0.99 (Argyle, 2002).

As well as applying our own FP models, we will also apply Cole's model to our datasets and compare the results with the model chosen by the selection algorithm in Section 2.8.3. This will allow us to assess the performance of a model built for a different dataset.

2.9 Assessing model fit and comparisons

To determine which of the 4 correlation models in Section 2.8 fit the data best, various model selection criteria and cross validation methods have been selected.

2.9.1 Cross validation

The Mean Square Error (MSE) is a statistic used to assess model fit and is used to compare models. The MSE is defined as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(\widehat{Y}_i - Y_i\right)^2 \tag{2.26}$$

To calculate the mean square error, the procedure below was followed for all correlation matrices:

1.) Remove upper triangular part of the matrix and diagonal correlations.
2.) Randomly select ¼ of the lower triangular correlations
3.) Apply the models to the remaining ¾ correlations
4.) Predict the correlations removed in step 2, back-transforming as necessary
5.) Calculate the MSE for each set of predictions

This process will be followed 10 times and the average MSE will be taken along with minimum and maximum values.

The Root Mean Square (RMS) of the errors can be calculated by taking the square root of the MSE. The average RMS can be calculated by taking the average of the square root of the ten MSE values.

2.9.2 R-Squared

The coefficient of determination, or more commonly, R^2, is an indicator of how well a statistical model fits its data. The coefficient of determination can be calculated as,

$$R^2 = \frac{\sum_{i=1}^{n}(\widehat{y_i} - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

where $\widehat{y_i}$ is the estimate from our fitted model, y_i are the observed values and $\bar{y}$ is the mean of our observed data calculated as $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$. This value sits between 0 and 1. Values close to 0 indicate a poor fit, values close to 1 indicate a good fit.

However, this expression does not take into account the number of variables included in the model. To adjust for this, we must account for the number of explanatory variables by modifying the expression, as shown here:

$$R^2(adj.) = \left(R^2 - \frac{k}{n-1}\right)\left(\frac{n-1}{n-k-1}\right) \tag{2.27}$$

Since this expression takes the number of explanatory variables k into account as well as the sample size n, we can adjust our R^2 so that terms that add no value decrease R^2.

2.9.3 Akaike Information Criterion

Another statistical tool which can be used to help with model selection is the AIC (Akaike 1974). The AIC is used to measure the quality of a model as a trade-off between the goodness of fit of a model and the complexity of the model. The AIC can be defined as:

$$AIC = 2k - 2log(L)$$

where k in this case is the number of parameters in the model and L is the maximised log likelihood function for the model.

The preferred model from a selection of models is the one with the minimum AIC value. Hence, the AIC does not test a null hypothesis for whether model parameters are significant, it simply indicates which is best model from a selection of models. If all models fitted were a poor fit, the AIC would simply select the best of the selection given. For models that predict on different scales, the AICs are not directly comparable. If however, we multiply the likelihood by the Jacobian of the transformation, then results can be compared directly as mentioned by Box & Cox (1964). As mentioned previously, two transformations are used within our modelling sections, the log and Fisher's transformation.

For log transformations, $f(y) = log(y)$ and the Jacobian becomes $J = \prod \frac{1}{y_i}$. We multiply the likelihood L by J or we add $log(J)$ to $log(L)$.

For the Fisher transformation, $f(y) = \frac{1}{2}\log\left(\frac{1+y}{1-y}\right)$, so J becomes $J = \prod \frac{1}{1-y^2}$. As with the log transformation, we can multiply L by J, or add $log(J)$ to $log(L)$.

Therefore, to calculate the $AIC(adj.)$, we can use the expression:

$$AIC(adj.) = 2k - 2.(log(L) + log(J)) \quad (2.28)$$

2.9.4 Bayesian Information Criterion

The Bayesian Information Criterion (BIC), or Schwarz Bayesian Information Criterion (SBC/SBIC) is another tool which can be used for model selection. The BIC was developed by Gideon E. Schwarz, documented in his 1978 paper "Estimating the Dimension of a Model" (Schwarz 1978). Schwarz set out to present an alternative approach to the AIC, which penalises the maximum log-likelihood in the following way:

$$BIC = log(n).k - 2\,log(L)$$

where L is the maximum log-likeihood, k is the number of parameters and n is the number of observations. The difference between the AIC and BIC is that the BIC is that k is weighted by $\log(n)$ rather than 2.

As with the AIC, models which are on different scales are not directly comparable in terms of BIC. However, the same adjustment with the AIC can be made to the BIC where we multiply the likelihood by the Jacobian, J, or we add the log of the Jacobian to the log likelihood. The expression therefore becomes:

$$BIC(adj.) = log(n).k - 2\,(log(L) + log(J)) \quad (2.29)$$

where for $f(y) = \log(y), J = \prod \frac{1}{y_i}$ and for $f(y) = \frac{1}{2}\log(\frac{1+y}{1-y}),\ J = \prod \frac{1}{1-y^2}$.

The BIC penalises more heavily than AIC. The reason for this is that where n exceeds 7, $log(n) > 2$.

2.10 Transition matrices

Hypotheses 1-3 state presumed transitions between various nutritional states. By calculating the conditional probability of children moving from one state to others, we can quantify the likely pathways children take through time.

The probability of moving between nutritional states from one time point to the next can be presented within a probability transition matrix. Let i be the state at t_1 (row i) and j be the state at t_2 (column j), then the probability of moving from state i to j is $P(j|i) = \frac{n_{ij}}{n_i} = P_{ij}$ where n_{ij} is the number of individuals moving from state i to j between t_1 and t_2 and $n_i = \sum_{j=1}^{k} n_{ij}$ is the number of individuals in state i at t_1.

A probability transition matrix can be therefore defined as,

$$P = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1k} \\ P_{21} & P_{22} & \dots & P_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ P_{k1} & P_{k2} & \dots & P_{kk} \end{bmatrix}$$

where $i, j = 1, \dots, k$ and $\sum_{j=1}^{k} P_{ij} = 1$.

The RR can be calculated as:

$$RR = \frac{P_1}{P_2}$$

where $P_1 = \frac{a}{c}$ and $P_2 = \frac{b}{d}$. Significance can be assessed using the interval:

$$\left(e^{\log(RR) - 1.96\sqrt{\left(\frac{1}{a}+\frac{1}{b}\right)-\left(\frac{1}{c}+\frac{1}{d}\right)}},\ e^{\log(RR) + 1.96\sqrt{\left(\frac{1}{a}+\frac{1}{b}\right)-\left(\frac{1}{c}+\frac{1}{d}\right)}} \right) \tag{2.30}$$

If this interval contains 1, we assume $P_1 = P_2$ (Sistrom & Garvan 2004).

2.10.1 Generalised Likelihood Ratio Test

Before making inference based on the stochastic models, we can test whether our probability transition matrices are significantly different from one another. If they are not, datasets can be pooled. Doing so provides more power to our analyses.

Consider two transition matrices A and B, each with k rows and k columns. Each transition within matrix A is calculated by taking the cell total n_{ij} over the row total n_i such that $p_{ij} = \frac{n_{ij}}{n_i}$ and each transition for matrix B is calculated by taking the cell total m_{ij} over the row total m_i such that $q_{ij} = \frac{m_{ij}}{m_i}$. A GLRT can be carried out to test whether the frequencies used to calculate the transitions for matrices A and B follow the same multinomial distribution. Assume each row i of matrix A and matrix B are multinomial distributed such that $n_{ij} \sim MN(p_{i1}, p_{i2}, p_{i3}, p_{i4}, p_{i5})$ and $m_{ij} \sim MN(q_{i1}, q_{i2}, q_{i3}, q_{i4}, q_{i5})$. Row i of matrices A and B follow the distributions,

$p(n_i|p_i) = \binom{n_i}{n_{i1}, \dots n_{ik}} p_{i1}^{n_{i1}} \dots p_{ik}^{n_{ik}}$ and $p(m_i|q_i) = \binom{m_i}{m_{i1}, \dots m_{ik}} q_{i1}^{m_{i1}} \dots q_{ik}^{m_{ik}}$

and follow joint distribution

$$p(n_i, m_i|p_i, q_i) = \binom{n_i}{n_{i1}, \dots n_{ik}} p_{i1}^{n_{i1}} \dots p_{ik}^{n_{ik}} \cdot \binom{m_i}{m_{i1}, \dots m_{ik}} q_{i1}^{m_{i1}} \dots q_{ik}^{m_{ik}}$$

and have likelihood

$$L(p_i, q_i|n_i, m_i) = \binom{n_i}{n_{i1}, \dots n_{ik}} p_{i1}^{n_{i1}} \dots p_{ik}^{n_{ik}} \cdot \binom{m_i}{m_{i1}, \dots m_{ik}} q_{i1}^{m_{i1}} \dots q_{ik}^{m_{ik}}$$

and log likelihood (constants A and B)

$$l(p_i, q_i|n_i, m_i) = logA + n_{i1}log(p_{i1}) + \cdots + n_{ik}log(p_{ik}) + logB \ + m_{i1}log(q_{i1}) + \cdots + m_{ik}\log(q_{ik})$$

Under $H0$, $p_{ij} = q_{ij} = \eta_{ij}$

$$l(\eta_i|n_i, m_i) = logA + n_{i1}log(\eta_{i1}) + \cdots + n_{ik}log(\eta_{ik}) + logB + m_{i1}log(\eta_{i1}) + \cdots + m_{ik}\log(\eta_{ik})$$

$$= (n_{i1} + m_{i1}) \log \eta_{i1} \ + \cdots + (n_{ik} + m_{ik}) \log \eta_{ik} + \lambda(1 - [(n_{i1} + m_{i1}) \log \eta_{i1} \ + \cdots + (n_{ik} + m_{ik}) \, log\eta_{ik}])$$

Using Lagrange multipliers under the constraints $\sum_{j=1}^{k} \eta_{ij} = 1$ and $\eta_{ij} \geq 0 \ \forall_j$

$$= \sum_{j=1}^{k} (n_{ij} + m_{ij}) log\eta_{ij} + \lambda(1 - \sum_{j=1}^{k} \eta_{ij})$$

$$\frac{\delta l}{\delta \eta_{ij}} = \frac{n_{ij} + m_{ij}}{\eta_{ij}} - \lambda = 0, \quad \frac{\delta l}{\delta \lambda} = (1 - \sum_{j=1}^{k} \eta_{ij}) = 0$$

$$n_{ij} + m_{ij} = \lambda \eta_{ij}$$

$$\sum_{j=1}^{k} (n_{ij} + m_{ij}) = \lambda \sum_{j=1}^{k} \eta_{ij}$$

$$n_i + m_i = \lambda$$

Since $n_i + m_i = \lambda$, we can write the maximum likelihood estimator (MLE) under $H0$ as:

$$n_{ij} + m_{ij} = (n_i + m_i)\eta_{ij}$$

$$\hat{\eta}_{ij}^{ML} = \frac{n_{ij} + m_{ij}}{n_i + m_i} \tag{2.31}$$

Under $H1$, $p_{ij} \neq q_{ij}$

$$l(p_i, q_i|n_i, m_i) = logA + n_{i1}log(p_{i1}) + \cdots + n_{ik}log(p_{ik}) + logB + m_{i1}log(q_{i1}) + \cdots + m_{ik}\log(q_{ik})$$

$$\frac{\delta l}{\delta p_i} = \frac{n_{ij}}{p_{ij}} - \lambda_1 = 0$$

$$n_{ij} = \lambda_1 p_{ij}$$

$$\sum_{j=1}^{k} n_{ij} = \lambda_1 \sum_{j=1}^{k} p_{ij}$$

$$n_i = \lambda_1$$

$$n_{ij} = n_i p_{ij}$$

$$\hat{p}_{ij}^{ML} = \frac{n_{ij}}{n_i} \tag{2.32}$$

$$\frac{\delta l}{\delta p_i} = \frac{m_{ij}}{q_{ij}} - \lambda_2 = 0$$

$$m_{ij} = \lambda_2 q_{ij}$$

$$\sum_{j=1}^{k} m_{ij} = \lambda_2 \sum_{j=1}^{k} q_{ij}$$

$$m_i = \lambda_2$$

$$m_{ij} = m_i q_{ij}$$

$$\hat{q}_{ij}^{ML} = \frac{m_{ij}}{m_i} \tag{2.33}$$

The GLRT for each row i is in the form:

$$D_i = -2 \, log \left(\frac{L_0}{L_A}\right) = 2 \, log \, (L_A - L_0)$$

Substituting yields,

$$-2\,log\Delta = 2([\log A + \log B + n_{i1}\log p_{i1} + \cdots + n_{ik}\log p_{ik} + m_{i1}\log q_{i1} + \cdots + m_{ik} log\, q_{ik}]$$

$$- [logA + logB + (n_{i1} + m_{i1})log\eta_{i1} + \cdots + (n_{ik} + m_{ik})log\eta_{ik}])$$

$$= 2\left(\sum_{j=1}^{k} n_{ij}\log p_{ij} + \sum_{j=1}^{k} m_{ij}\log q_{ij} - \sum_{j=1}^{k}(n_{ij} + m_{ij})\log\eta_{ij}\right)$$

$$D_i = 2\left(\sum_{j=1}^{k} n_{ij}\log\left(\frac{p_{ij}}{\eta_{ij}}\right) + \sum_{j=1}^{k} m_{ij}\log\left(\frac{q_{ij}}{\eta_{ij}}\right)\right)$$

However, this only allows us to test whether row i in matrices A and B are significantly different.

To test the whole matrix, the sum must be taken over the rows such that:

$$D_{Total} = 2\sum_{i=1}^{k}\left(\sum_{j=1}^{k} n_{ij}\log\left(\frac{\hat{p}_{ij}}{\hat{\eta}_{ij}}\right) + \sum_{j=1}^{k} m_{ij}\log\left(\frac{\hat{q}_{ij}}{\hat{\eta}_{ij}}\right)\right) \quad (2.34)$$

This expression requires $\hat{p}_{ij}$, $\hat{q}_{ij}$ and $\hat{\eta}_{ij}$ to be non-zero. One strategy is to add a constant to each and every frequency, which guarantees $\hat{p}_{ij}$, $\hat{q}_{ij}$ and $\hat{\eta}_{ij}$ are all > 0 but as some of the frequencies are small, this drastically affects results. Another option is to use the asymptotically equivalent Pearson χ^2 test statistic as this expression does not require $\hat{p}_{ij}$ or $\hat{q}_{ij}$ to be > 0.

$$2\sum_{i=1}^{k}\left(\sum_{j=1}^{k} n_{ij}\frac{(\hat{p}_{ij} - \hat{\eta}_{ij})^2}{\hat{\eta}_{ij}} + \sum_{j=1}^{k} m_{ij}\frac{(\hat{q}_{ij} - \hat{\eta}_{ij})^2}{\hat{\eta}_{ij}}\right) \sim \chi^2(F_i - 1) \quad (2.35)$$

$$F_i = \{j : \hat{\eta}_{ij} > 0\}$$

Note that no reference could be found for this test within the scientific literature.

2.11 Survival analysis

Hypothesis 4 states that recent size is a more valuable predictor of mortality than growth over a recent time period. Survival analysis using Cox proportional hazard models allow us to determine what the best predictors of “time until death” are. By determining whether models across datasets are in agreement, the best set of anthropometric measurements which predict mortality can be determined.

2.11.1 Survival function, hazard rates and censoring

Let T be a non-negative random variable which represents the time until the occurrence of the event of interest. The probability density function of T is:

$$f(t) = \lim_{\delta_t \to 0} \frac{P(t \leq T < t + \delta_t)}{\delta_t}$$

and the cumulative distribution function of T is:

$$F(t) = P(T \leq t)$$

and the survival function of T is:

$$S(t) = P(T > t) = 1 - F(t) = \int_t^{\infty} f(x).dx$$

This provides the probability of being alive just before duration t, or in other words the probability that the event of interest has not occurred by time t.

The instantaneous risk of death can be expressed as the hazard function $h(t)$:

$$h(t) = \lim_{\delta_t \to 0} \frac{P(t \leq T < t + \delta_t | T \geq t)}{\delta_t} = \lim_{\delta_t \to 0} \frac{P(t \leq T < t + \delta_t)}{\delta_t P(T \geq t)} = \frac{f(t)}{S(t)} \qquad (2.36)$$

therefore,

$$h(t) = \frac{d}{dt}\big(-log.S(t)\big)$$

and since $-f(t)$ is the derivative of $S(t)$ we can re-write $S(t)$ as:

$$S(t) = \exp\left(-\int_0^t h(x).dx\right)$$

The hazard function is the probability of death in the next moment, given that you are alive now. We can denote the cumulative hazard as,

$$H(t) = \int_0^t h(x).dx$$

This can be thought of as the sum of the risks going from 0 to t.

Given the simplest survival distribution, which is a constant risk over time,

$$h(t) = \lambda$$

then the corresponding survival function is,

$$S(t) = \exp(-\lambda.t)$$

which is an exponential distribution with parameter λ. By multiplying the survival function by the hazard function we obtain the probability distribution function $f(t)$:

$$f(t) = \lambda.exp(-\lambda.t) \qquad (2.37)$$

which has mean $1/\lambda$ and variance $1/\lambda^2$.

One problem with survival analysis is the issue of censoring. For some individuals the event of interest might not occur within a certain specified timeframe. In our models we only wish to model up until 2 years of age – most children will not die within that 2 year period. Subjects who are still alive at the end of the study are right censored, as are those who leave before the end of the study (therefore their outcome is not known) (Miller 2011).

2.11.2 Cox Proportional Hazards Model

Linear regression is not adequate while predicting mortality because it does not cope properly with censoring and, in any case, survival times are not normally distributed. Therefore a regression model tailored to survival data is needed. A suitable approach is provided by the Cox proportional hazards model (Cox 1972). The package $coxph$ in R can be used to apply the model.

In this model, the hazard function, denoted $h(t)$ is expressed as;

$$h(t) = h_0(t).\exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)\,, \quad t > 0$$

where $x_1, x_2, \ldots, x_n$ are a collection of covariates and $h_o(t)$ is some unknown baseline hazard at time t ($h_0(t)$ represents a person with 0 for all predictor variables, no assumptions are made about $h_0(t)$ apart from $h(t) > 0$).

By dividing both sides by $h_0(t)$ we obtain

$$\frac{h_i(t)}{h_0(t)} = \exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_n x_{in})$$

and by taking logs

$$\log\left(\frac{h_i(t)}{h_0(t)}\right) = (\beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_n x_{in}) \tag{2.38}$$

The ratio $\frac{h_i(t)}{h_0(t)}$ is called the hazard ratio for subject i.

Time dependent variables can be introduced by letting x depend on t such that

$$\log\left(\frac{h_i(t)}{h_0(t)}\right) = (\beta_1 x_{i1}(t) + \beta_2 x_{i2}(t) + \cdots + \beta_n x_{in}(t)) \tag{2.39}$$

The model is semi-parametric because the baseline hazard can take any form however the covariates enter the model linearly.

If we have two observations i and i' that differ in their x values with

$$h_i(t) = h_0(t).\exp(\beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_n x_{in}) = h_0(t).\exp(\eta_i)$$

and

$$h_{i'}(t) = h_0(t).\exp(\beta_1 x_{i'1} + \beta_2 x_{i'2} + \cdots + \beta_n x_{i'n}) = h_0(t).\exp(\eta_{i'})$$

then the hazard ratio for these two observations is,

$$\frac{h_i(t)}{h_{i'}(t)} = \frac{h_0(t).e^{\eta_i}}{h_0(t).e^{\eta_{i'}}}$$
$$= \frac{e^{\eta_i}}{e^{\eta_{i'}}}$$
$$= e_i^{\eta - \eta_{i'}}$$

Model comparisons will be carried out via BIC, with the best model minimising the criterion (Miller 2011). BIC was chosen as it penalises against the number of parameters more heavily.

2.11.3 Stratification

A stratified Cox model allows the form of the hazard function to vary across levels of a categorical stratification variable Z, without estimating the effects of the outcome. This allows us to adjust for variable Z when making inferences with the model. This can also be used when the variable does not satisfy the proportional hazards assumption.

2.12 Chapter conclusions

In this chapter we described the statistical methodology that will be used throughout this thesis. Sections 2.4 to 2.9 described methodology used to address aims which focus on summarising and modelling data (aims 1-3). Sections 2.10 to 2.11 presented methodology which can be used to quantify how anthropometry can be used to predict future adverse outcomes (aims 4-6).

Aim 3 investigated the impact that external references have on some common measures of weight gain and investigated how to modify the measures to allow for them. Section 2.7.4 thus investigated how to adapt conditional weight gain for use with an external reference. Both unconditional (Section 2.7.1) and conditional (Section 2.7.2) measures can be used to identify the magnitude of growth of a child relative to their population. Caution should however be taken when using the expression for unconditional weight gain (Cole 1997) since unconditional measures are biased (Cole 1993). This is because unconditional weight gain does not take into account RTTM, and as a consequence, should only be used in age groups where correlation is high (Cole 1997) since any change in centile may be attributed to a change in weight rather than expected movement.

Conditional weight gain (Cole 1995) corrects for RTTM by taking into account the correlation between Z scores at t_1 and t_2. Without this correction, the amount of bias

increases as the correlation decreases, thus, the longer the time frame in which unconditional weight gain is calculated, the larger the bias. However, the expression for conditional weight gain only allows comparative growth with the distribution of values in the specific population to which the infant belongs, known as an 'internal' reference. If an 'external' reference is used, then the condition of zero mean and unit variance does not hold and the assumptions for the expression break down. Roelants (2013) suggests using models from Cameron (1980) or Berkey at al. (1983) with appropriate estimates for the mean and standard deviation of the SD scores to estimate conditional growth under these conditions, although, the advantage of working on the SD scale is largely lost. The generalised model, Equation (2.19), handles external references while on the Z scale, which in turn can be converted to a centile. However, the expression requires estimates of $\mu_{Z_1}, \mu_{Z_2}, \sigma_{Z_1}, \sigma_{Z_2}$ and r. These parameters can easily be calculated if data are available at t_2 and t_2. However, if we wish to estimate these between measurement dates, we need to model these parameters. The models which can be used to do this are presented in Chapters 3 and 4. The GLRT developed in Section 2.10.1 is also novel. No reference could be found within the literature which describes this test, although, a similar "one sample" test was presented by Bickenbach & Bode (2001), applied to financial data.

In Chapter 3, we introduce three developing world datasets which have been made available for analysis. We apply the LMS method to the datasets via GAMLSS, creating 'internal' growth charts of the available datasets. This provides insight into how children from these datasets grow in terms of average growth trajectory and variability. The datasets can then be compared and contrasted. Nevertheless, there are disadvantages in using internal reference charts to assess child growth since these charts describe growth at a specific time and place. Ideally, references such as the WHO standard should be used as this is an idealised 'external' reference which describes growth under optimal conditions, is well known and is widely used. In Chapter 3 we use GAMLSS to assess how children from the datasets made available for this thesis fit the standard, modelling the mean, μ_{Z_t}, and standard deviation, σ_{Z_t}, over time to assess how developing world datasets fit the standard. These models can be used to interpolate estimates of μ_{Z_1}, $\mu_{Z_2}, \sigma_{Z_1}, \sigma_{Z_2}$ for use in the expression for generalised conditional weight gain.

Chapter 3

Datasets and growth characteristics

3.1 Introduction

Three datasets have been made available for analysis within this thesis, datasets from: Malawi, Pakistan and South Africa. This chapter focuses on introducing these datasets and reviewing how the datasets were collected. Papers which have been produced as a result of the studies are reviewed here. Refer to Tables 3.1 – 3.3 for summary evidence tables. Observational summaries and comparisons between the three datasets can be found throughout the chapter and towards the end the LMS method is used to develop growth reference charts. Serial measurements can be plotted on these charts to assess growth. If a child's measurements deviate from a previously established growth pattern then there may be cause for concern. There are, however, disadvantages in using internal growth chart references as they are developed from children at a particular time and place. The populations of children who make up the references may not grow at an adequate rate. The WHO state that all children grow the same under optimal conditions and advise universal use of their WHO 2006 growth standard, developed from populations of healthy children from around the globe. Work has been carried out by Roelants (2013), Hui et al. (2008) and Wright et al. (2008), assessing how populations of pre-school children from the developed world fit the standard.However, little is known about how children from the developing world fit the distribution of children which make up this idealised reference distribution. One of our aims is to determine how real populations of pre-school children from the developing world fit the standard.

To do this, we convert measurements from our datasets to Z scores relative to the WHO standard and use GAMLSS to model the transformed series assuming a normal distribution, allowing us to model the mean and SD over time. By plotting the mean and SD curves we can determine how children from our datasets grow relative to the standard. Conditional weight gain was introduced in the last chapter which requires $Z_X \sim N(0,1)$. If using the expression with an external reference, we no longer assume $\mu_{Z_t} = 0$ and $\sigma_{Z_t} = 1$. Both of these parameters can be estimated from the transformed series at t_1 and t_2 to obtain $\mu_{Z_1}, \mu_{Z_2}, \sigma_{Z_1}$ and σ_{Z_2}, or interpolated from the GAMLSS model of the transformed series. This analysis can be found after the introduction of the datasets.

Note that all analyses have been restricted from 0 to 24 months, the maximum age of the Pakistani set, to make them comparable.

3.2 Lungwena Child Survival Study

The Lungwena Child Survival Study (LCSS) was an observational cohort study which set out to analyse the varying factors which influence child mortality in an impoverished community within Lungwena, Malawi. Mothers were enrolled between June 1995 and August 1996. Data were collected on the health of both pregnant women and their children under five years old. Both pre and post-natal factors which influence infant mortality were of interest as well as the magnitude of their influence. Referring to Figure 3.1, from Vaahtera et al. (2000), 795 pregnant women were enrolled in the study and 813 foetuses (18 twin pregnancies) started follow up. However, only 760 were born alive and of the 760, only 733 were made available (27 dropped out). Note that the number evaluable for neonatal outcome and infancy outcome only take abortions, stillbirths and dropouts into account, not neonatal deaths or postnatal infant deaths.

A number of maternal and child health factors have already been analysed as predictors of infant mortality as well as socioeconomic factors such as the age of parents, number and age of people in the household, religion and occupation of parents and size of building and material of house. For those with a first recorded weight of >2500g, the infant mortality rate was 103 per 1000 whereas those with a first recorded weight of <2500g had an infant mortality rate of 205 per 1000, an RR of 1.99. Other high risk factors included maternal primiparity with an RR of 2.21 and birth date between May to July with an RR of 2.1 (Vaahtera et al. 2000). Ashorn et al. (2002) used the data to study male

biased mortality among 1 and 2 year old children within the Malawian dataset. A total of 147 1-2 year old infants died during follow up resulting in infant and under 3 year mortality rates of 136 and 202 deaths/1000 live births, respectively. Mortality was higher among boys than girls immediately after birth, but not during the subsequent 6-8 months. Between 9 and 35 months of age, there was a notable excess of male mortality (RR 1.9, 95% CI 1.0 to 3.0, p = 0.04, Cox regression). Previous studies show that three biological factors have been associated with male biased mortality in infants and older children: malnutrition, immunisations and accidents (Naeye 1971). However, in general, it is females that have a higher mortality rate within developing world countries, whereas the reverse is true within the developed world (Waldron 1987). The prevalence of underweight, stunting or wasting was higher among boys than girls in the study and close to half the deaths were malnutrition associated. However, the sex difference in mortality was not significantly affected by adjustment for the incidence of moderate wasting, other forms of malnutrition, or children's anthropometric measurements at 9 months of age. Ashorn states that malnutrition might have explained in part, but not the majority of the sex specific mortality trends among 1–2 year old children. Accident related deaths were infrequent and did not account for the difference (Ashorn et al. 2002).

Growth curves for this dataset have already been developed using the LMS method by Maleta et al. (2003), up to 36 months. Both the NCHS/WHO and CDC reference populations were assessed against the Lungwena population to determine average deviations from the population median. New-borns in Lungwena were on average 0.5kg lighter and 2.5cm shorter than the CDC reference babies and 0.3kg lighter and 3.1cm shorter than the NCHS/WHO standard babies. Compared to both references, children fell away from the median after three months and then grew parallel to the 3rd centile after 12 months. The median height was well below the 3rd centile after 12 months compared to both references.

Maleta (2003) conducted additional research into growth, studying the effect of seasonality in the dataset. It is known that childhood growth follows a seasonal pattern in both low income and high income countries. Maleta argued that according to most available evidence, weight gain appears to peak in autumn and winter and height gain in spring and summer. In low income countries this is attributed to seasonal variation in food availability and the prevalence of infectious disease. Those under 6 months old had the greatest height and weight gains between May and September. After 12 months,

weight gains peaked in May to July and height gains 3 months later between August and October.

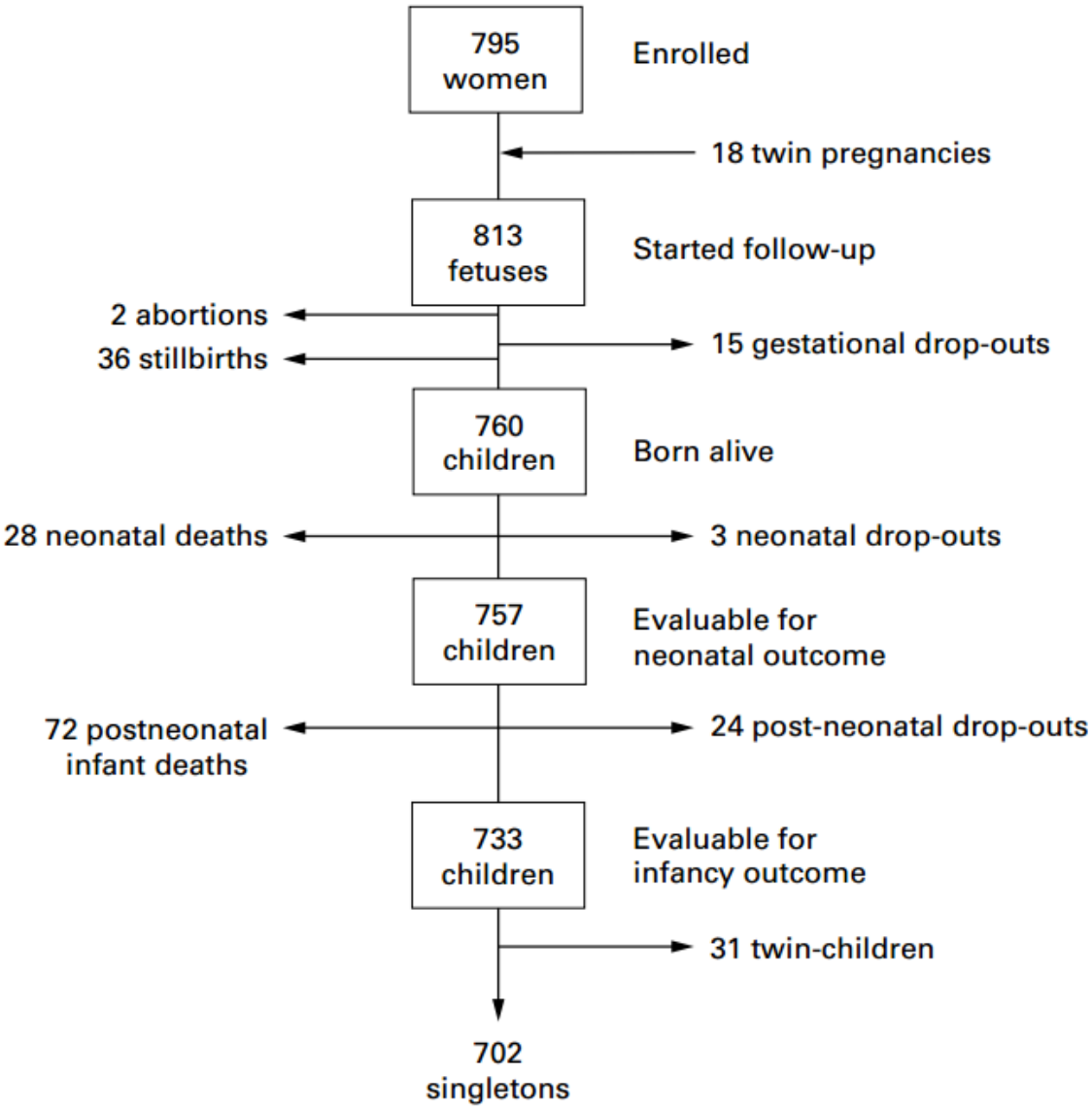

Figure 3.1: Flow diagram of women eligible for enrolment and numbers lost to follow-up during the study (Vaahtera et al. 2000)

Maleta's further research set out to determine the predictors of malnutrition in the Malawian data. The reference population used was adopted from the 1977 United States NCHS reference, based largely on non-breast fed American infants and children. Factors associated with severe underweight (WAZ < -3) were: delivery attended by trained assistant and morbidity in infancy. Factors associated with severe stunting were (HAZ < -3): delivery attended by trained assistant, sex and morbidity in infancy.

Espo et al. (2002) also studied predictors of severe stunting within the dataset. Variables associated with stunting using a univariate analysis were: preterm birth, low maternal height, low weight gain of mother during pregnancy, gender, low quality weaning diet,

total number of illness episodes, total number of diarrhoea episodes, place of birth and literacy of father. All of these variables apart from literacy of father (excluding duration of pregnancy and diarrhoea morbidity because they are associated with birth weight and total morbidity, respectively) were significant predictors when included in a logistic regression model together (Espo et al. 2002).

Author, year	Outcome	Risk factors	Results
(Ashorn, Maleta, Espo, & Kulmala, 2002) (Ashorn et al. 2002)	Mortality	Gender	Mortality higher for boys than girls after birth Mortality similar during months 6-8 Mortality higher for boys than girls during months 9-35
(Espo et al. 2002)	Stunting	WAZ during first 3 mo Weaning diet Total number of illness episodes Maternal height Gender of child Place of birth Weight gain of mother >200 g/wk during pregnancy Weight gain of mother >200 g/wk during pregnancy Weight gain of mother >200 g/wk during pregnancy Weight gain of mother	All associated with stunting
(Maleta et al. 2003)	Growth	Growth compared to other references	Gained weight quicker than CDC at first, then fell to 3rd centile Gained weight at the same rate as NCHS/WHO then drops to 3rd centile Median height below 3rd centile at 12 months compared to both references

(Maleta et al. 2003)	Growth	Season	Weight gain peaks in autumn and winter, height gain in spring and summer
(Maleta et al. 2003)	Malnutrition (underweight, wasting or stunting)	**Socio-economic factors** Improper water and sanitation No possessions Distance to health centre ≥ 5 km (vs. <5 km) 1.4 [1.0, 1.9]* 1.2 [0.8, 1.8] 1.3 [0.9, 1.8] **Maternal characteristics** Mother illiterate HIV-ELISA test Height Body mass index Pregnancy weight gain Delivery attended by trained assistant **Child factors** Male sex Average size 1–3 months Complementary feeds Birth at < 37 completed gestation weeks ≥ 2.4 episodes/month infancy morbidity (vs. < 2.4 months) ≥ 2.4 episodes/month infancy morbidity	Significant predictors: **Underweight** HIV-ELISA test Delivery attended by trained assistant Average size 1–3 months ≥ 2.4 episodes/month infancy morbidity **Stunting:** HIVE-ELISA test delivery attended by trained assistant male sex birth <37 completed gestation weeks (vs. ≥ 37 weeks), ≥2.4 episodes/month infancy morbidity **Wasting**: HIVE-ELISA test positive delivery attended by trained assistant average size 1-3 months ≥2.4 episodes/month infancy morbidity (vs. <2.4 months)

Table 3.1: Evidence table: LCSS (Malawi)

3.2.1 LCSS observational summary

Measurements were taken each month (including birth) up until 18 months and every 3 months thereafter (a total of 34 scheduled measurement dates over the 5 years). Within the original dataset, there were 34 rows for each child, with NA values recorded if information was not available (27642 rows).

Our analyses, however, have been restricted to the 0-24 month range. This smaller dataset includes only 21 scheduled measurements.

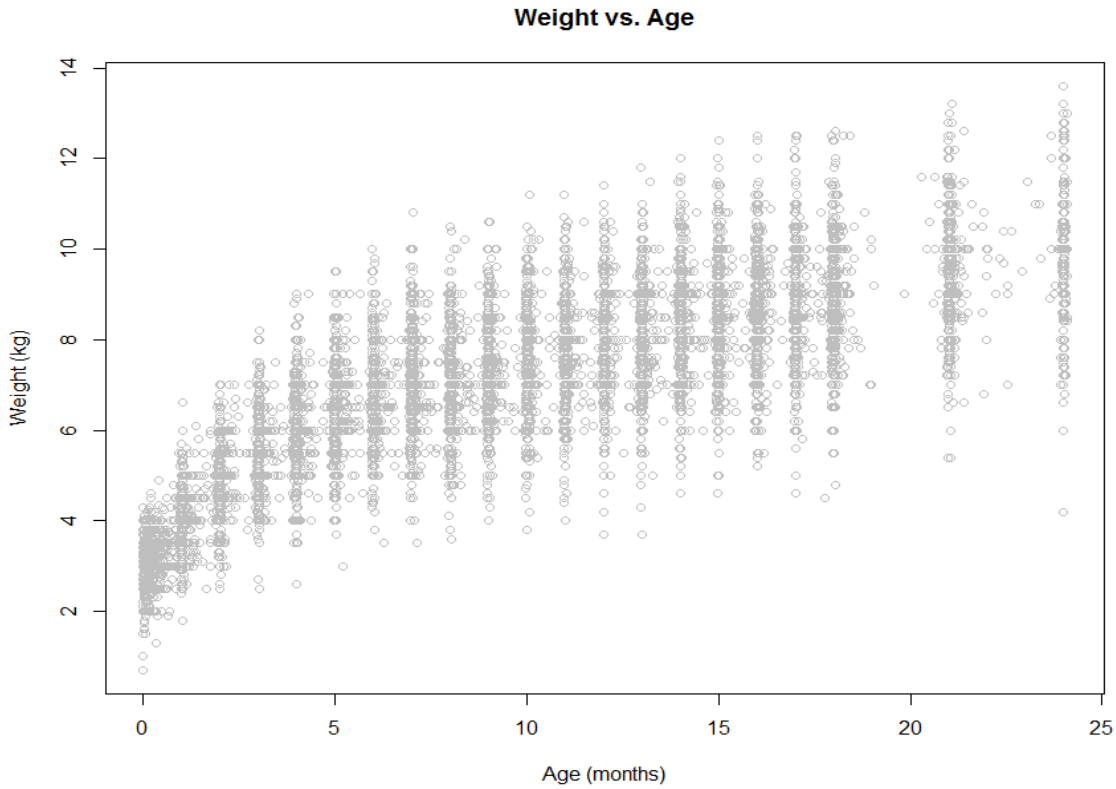

Figure 3.2: Weight vs. Age (Malawi)

It can be seen within Figure 3.2 that age measurements around month 0 (the first scheduled measurement) are more spread out along the x axis than the other 20 measurements. This is because measurements were taken as soon as a trained anthropometrist was available and not necessarily as soon as children were born. Figure 3.3 shows the frequency of children per number of measurements.

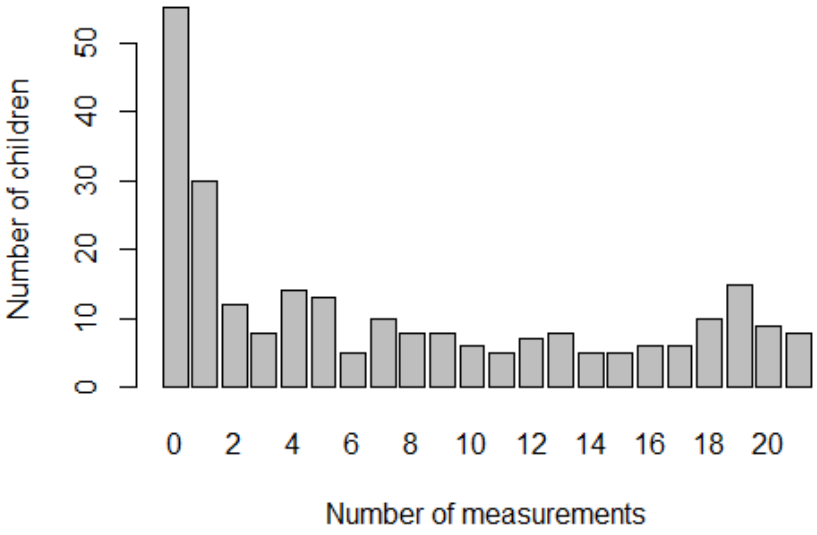

Figure 3.3: Number of children per number of measurements (Malawi)

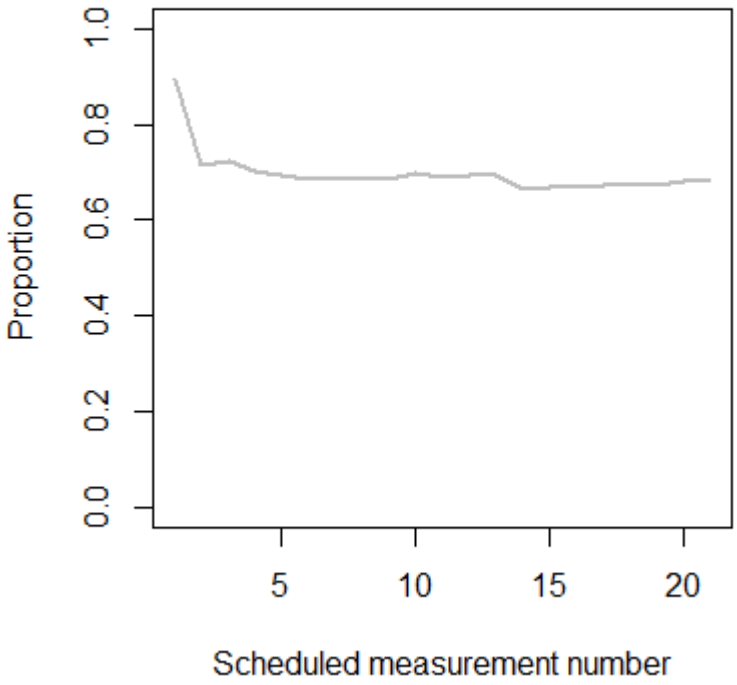

Figure 3.4: Proportion of observations at each scheduled measurement number (Malawi)

Figure 3.4 shows the proportion of total observations at each scheduled measurement date, with 813 as the maximum. The calculations were made by counting the number of observations between the mid points of scheduled measurement dates. Therefore, the

proportions include deaths. The proportion should therefore treated simply as the proportion of available observations. See Figure A1, Appendix A, for more information on missing data. The structure of the dataframe restricted the number of measurements at each scheduled measurement date by each child to be 1.

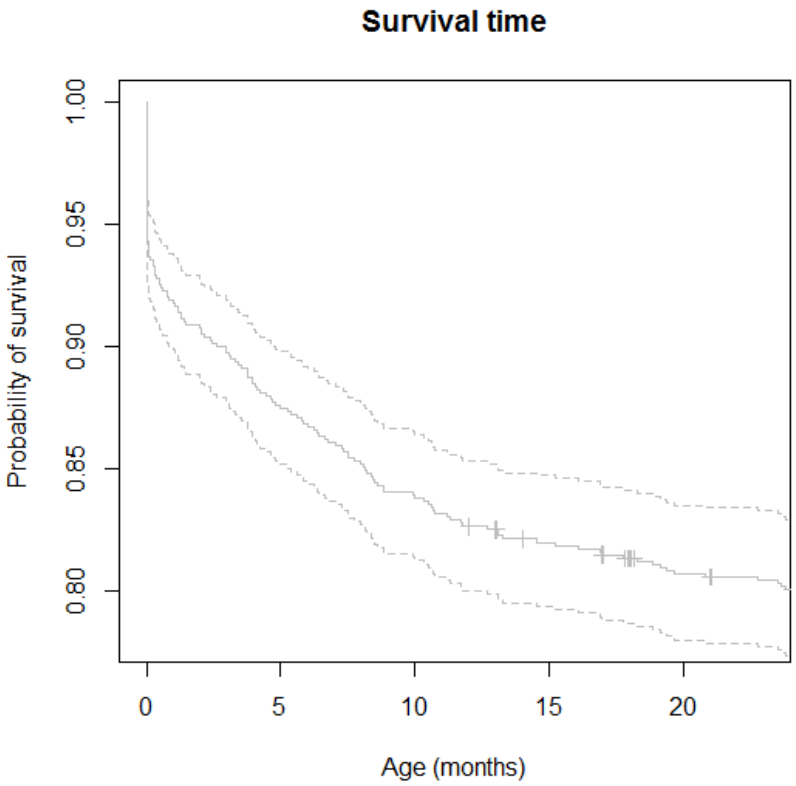

Figure 3.5: Kaplan Meier plot (Malawi)

A Kaplan Meier plot for the Malawi dataset can be seen in Figure 3.5. The plot shows the survival curve for the Malawi dataset. Still births can be seen at month 0. Right censored data are marked with crosses. The dotted lines represent 95% confidence intervals.

3.3 Lahore Longitudinal Study

The second dataset, from the Lahore Longitudinal Study (LLS) was collected in Lahore, Pakistan. Four areas were involved in this study: a village, a peri-urban slum, an urban slum and an upper middle class control group.

The village was represented by the villages of Halloki and Sadoki, about 40km from Lahore. The peri-urban slum was selected from a 20km stretch of dwellings along the railway track at the periphery of the old city. The urban slum group was taken from inside the old walled city. The middle class group was scattered over the city suburbs (Jalil, Lindblad, Hanson, Khan, Ashraf, et al. 1993).

The aim of this study was to "characterise the determinants of child health in a rapidly urbanising community". Firstly, a cross sectional survey collecting socio-economic and demographic background information was carried out between March 1984 and August 1984, then this was followed up by a longitudinal study of 1,476 infants representing the outcome of the all pregnancies registered from September 1984 to March 1987.

The original dataset included 1,476 infants which were followed monthly from birth up to 36 months and less frequently thereafter. The data provider, Professor Shakila Zaman, specified there were no known cases of HIV. At 24 months, 70% of the infants were still in the study, 11% had died, 13% had moved from their respective areas and 6% refused to participate. A number of longitudinal, socio-economic and demographic background variables were recorded. Feeding practices, screening tests, vaccination coverage, psycho-motor development, measurements and diagnosis of various diseases were taken continuously throughout the study. The number of adults and children in the household, education, physical status of the parents and housing standard were also taken at the start of the study. Furthermore, childcare and hygiene scores were calculated. A number of papers have been published using this dataset, analysing different variables and their effects on the health of children from the four areas (Jalil, Lindblad, Hanson, Khan, Ashraf, et al. 1993).

The dataset used within this thesis is an extended dataset which was collected from 1984-1994. This dataset contains a total of 3146 children, 1,314 of which were from the village area, 572 from the peri-urban slum, 921 from the urban slum and 339 from the middle class area. Published papers reviewed in this section all analysed only the smaller dataset with 1,476 children.

Saleemi et al. published a paper in 2001 which studied stunting at 6, 12, 24 and 60 months as well as linear growth in the children. A number of risk factors were analysed which included maternal factors, size at birth and duration of breastfeeding. Maternal age, weight, height, education (illiterate, number of years of schooling) and gestational age were analysed as maternal factors. Weight and length at birth were also analysed as risk factors as well as duration of breastfeeding at 6, 12, 24 and 60 months. The findings showed that if the duration of breastfeeding was less than 4 months for the first 6 months of life, the risk of stunting was increased twofold at 6 months of age. With a gestational age of less than 36 weeks, the risk of stunting almost doubled. Those who

were stunted at 6 months were three times as likely to be stunted at 12 months (Saleemi et al. 2001).

Zaman et al. (1993) also studied morbidity within the dataset, studying patterns according to age, area of living, sex and season. Infections were responsible for 87% of morbidity in the first 24 months of life. In 20,911 examinations, 16,099 recorded at least one disease (77%). Diarrheal diseases were recorded in 30.3% of all monthly intervals and Acute Respiratory Infections (ARIs) were recorded in 22.4%. Only 0.5% of recorded morbidity was accounted for by vaccine preventable infection diseases. Morbidity was evenly distributed between sexes, but age and the area that children inhabited influenced patterns of morbidity.

Mahmud et al. (1993) studied the incidence and home management of infantile diarrhoeal disease. On average, each child had 3.6 episodes of diarrhoea during year one: 4.3 episodes in the village, 3.4 in the urban slum, 4.4 in the peri-urban slum and 1.4 in the middle class area. No difference in sex was found. However, there was seasonal variation with a peak between April and June. Diarrhoea was persistent in 14% of all acute cases, the being highest in the village at 18%, followed by the peri-urban slum at 14%, the middle class area at 10% and finally the urban slum at 8% (Mahmud et al. 1993). Khalil et al. (1993) studied the causative agents of acute diarrhoea within the population, with the aim to determine the frequencies of the most common enteropathogens within the faeces. The total isolation rate was 73.4%. One third was of viral origin and two thirds were of bacterial origin (Khalil et al. 1993).

Khan et al. (1993) studied mortality rates between each Pakistani area. At age 6 months, the cumulative mortality rate per 1000 was 116 within the peri-urban slum population, 105 within the village, 90 within the urban slum and 16 within the middle class area. By 12 months, these rates had increased to: 145, 118, 96 and 16 respectively. By 2 years, the rates were: 158, 132, 106 and 16 respectively. The biggest associated morbidity measures were chronic diarrhoea, acute diarrhoea, respiratory infection and asphyxia neonatorum.

Karlberg et al. (1993) studied infant growth within the dataset.

Karlberg created growth reference charts for all four areas (weight, length and head circumference). Those within the peri-urban slum were on average shorter than those in the village area, who were in turn shorter than those in the urban area. Karlberg found that those living in the poorer areas were more likely to be stunted at 24 months of age: 63% in the peri-urban slum, 54% in the village and 26% in the urban slum (middle class

reference) however, fewer differences could be seen between the areas in weight for length (10-20%).

Ashraf et al. (1993) studied feeding patterns and practices. 87-98% of the infants in all four areas were breastfed at one month of age but only 9% were exclusively breastfed (declining with age and highly influenced by season) with mothers adding water to human milk or feeding their children with fresh animal milk and/or commercial formula.

Yaqoob et al. (1993) studied psychomotor development between the areas. Psychomotor development within the middle class area was on par with European and North American standards, however there were significant delays in psychomotor development in infants belonging to the three poorer areas. Infants were on average about 3 months delayed in walking and fine motor activity (building a tower of 3 cubes) in comparison to the middle class area.

Author, year	Outcome	Risk factors	Results
(Ashraf et al. 1993)	Feeding patterns	Between cohorts	See text
(Jalil, Lindblad, Hanson, Khan, Yaqoob, et al. 1993)	Weight-for-length at birth < -2 % Mortality rates Mortality		Weight-for-height <-2 ranges from 12-31% for boys, 12-25% for girls See Karlberg et al. 1993 for mortality rates Mortality significantly related to birth weight
(Karlberg et al. 1993)	Infant growth	Between cohorts	Length by cohort: middle class, urban slum, village, peri-urban slum
(Khan et al. 1993)	Mortality	Between cohorts	Mortality rates (2y): middle class (0.016), urban slum (0.106), village(0.132), peri-urban slum (0.158)
(Mahmud et al. 1993)	Incidence and home management of diarrhoeal disease	Between cohorts, sex, season	Cases during one year: 4.4 peri-urban, 4.3 village, 3.4 urban slum, 1.4 middle class No difference between sexes. Peak episodes April-June

(Saleemi et al. 2001)	Stunting at 6, 12, 24 and 60 months	Maternal factors: age, maternal weight, height, education, gestational age Size at birth: weight, length Duration of breastfeeding: at 6m, 12m, 24m, 60m	Predictors of stunting at: 6m: duration of breastfeeding, gestational age 12m: maternal height, birthweight, stunting at 6m 24m: stunting at 18m 60m: stunting at 18m, stunting at 24m
(Yaqoob et al. 1993)	Psychomotor development	Between cohorts, sex	Development of middle class area on par with European and North American countries. Delays in development in the three poorer areas. No differences between sexes (apart from in the middle class area, girls developed faster than boys by one month)
(Zaman et al. 1993)	Morbidity	Studied patterns according to age, cohort, sex, season	Morbidity evenly distributed between sexes Age and area influence morbidity

Table 3.2: Evidence table: LLS (Pakistan)

3.3.1 Comparing growth over time

Since measurements within the Pakistani dataset were taken over a period of 10 years, it was assumed that growth patterns may have changed over time. We investigated whether or not this was true. Dates were arbitrarily split up into two different groups: measurements taken between 1984 and 1987, and those taken between 1988 and 1994 (inclusive). Figure 3.6 shows the distribution of measurements between the two groups.

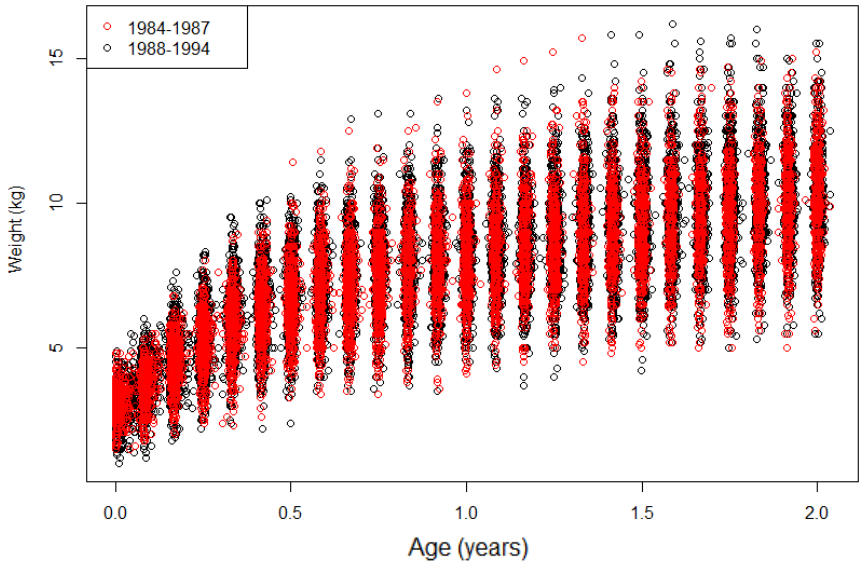

Figure 3.6: Weight vs. Age between time groups (Pakistan)

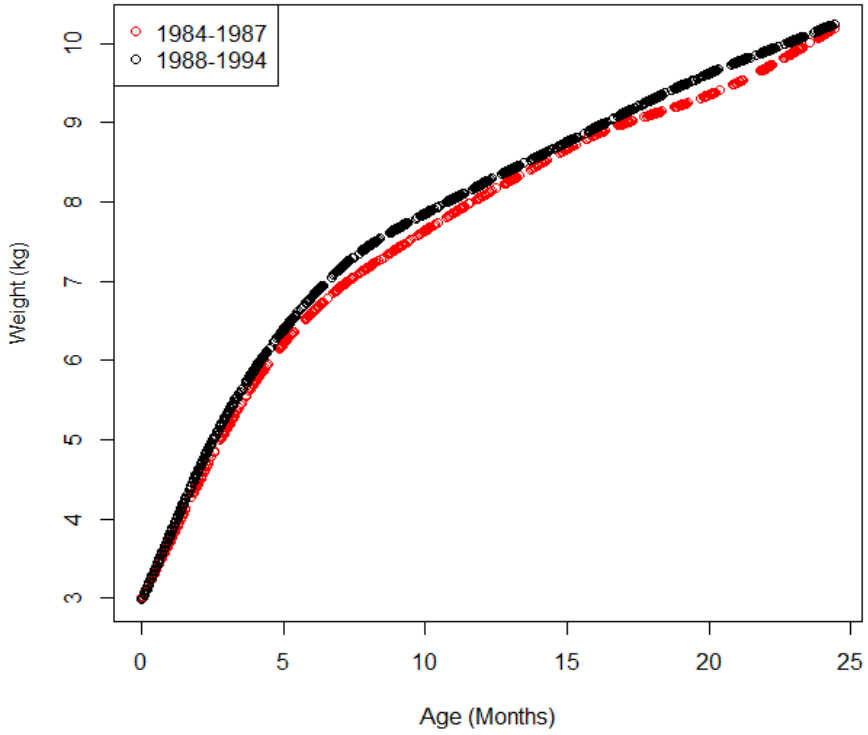

Figure 3.7: Median weight curves between time groups vs. Age (Pakistan)

Figure 3.7 above shows the two median curves from the 1984 to 1987 group (red) and the 1988 to 1994 group (black). There is clear overlap between the two groups with the 1984-1987 median, slightly lower than the 1988-1994 median after about 3 months. From this

preliminary analysis, there does not seem to be any difference in structure between age. No adjustment will be made for time within our analyses.

3.3.2 Comparing growth between socio-economic groups

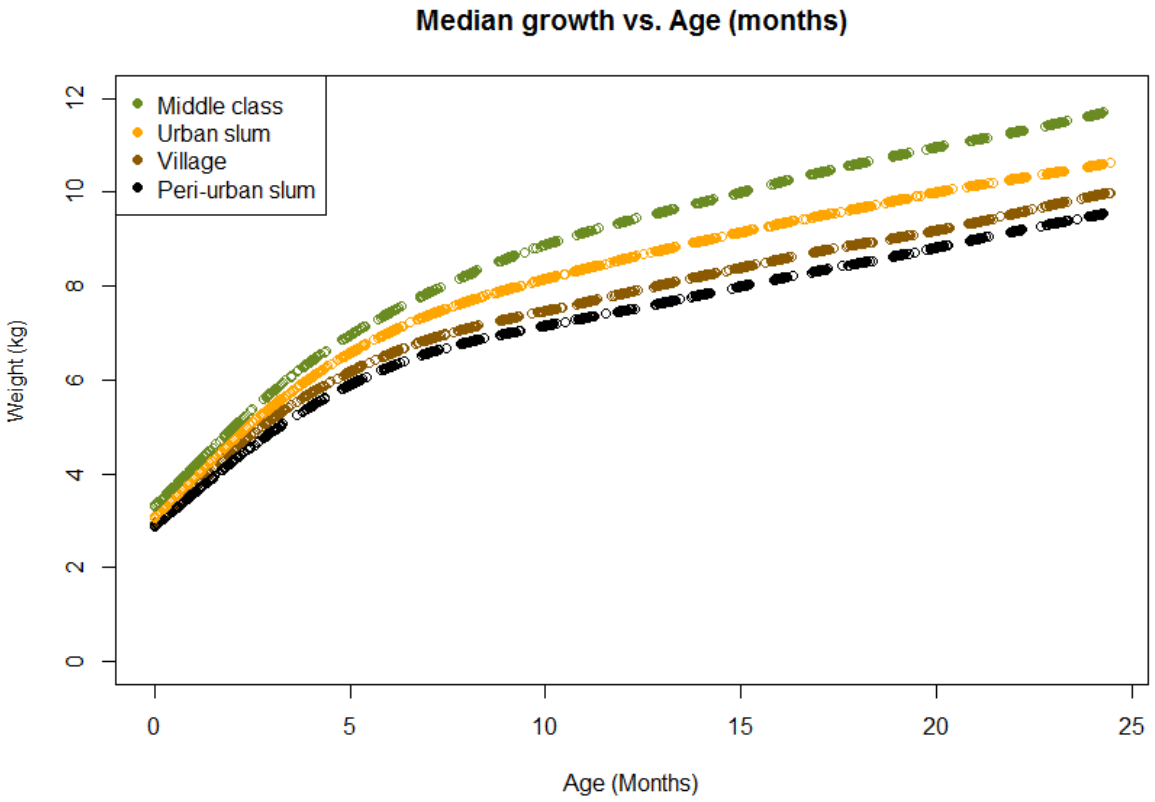

Figure 3.8: Median weight curves vs. Age (Pakistan)

Within Figure 3.8 the four median curves from the four areas within the Pakistan dataset can be seen for comparative purposes. All four curves sit relatively close at month 0 then slowly spread out. It is clear that children within the middle class area grow best with the highest median curve, followed by the urban slum, the village area and finally the peri-urban slum. The difference in medians between groups in Figure 3.8 indicates the four different areas should be modelled separately.

3.3.3 LLS observational summary

The total number of children included within the analysis was 3,146 with 1,314 within the village area, 572 within the peri-urban slum, 921 within the urban slum and 339 within the middle class area.

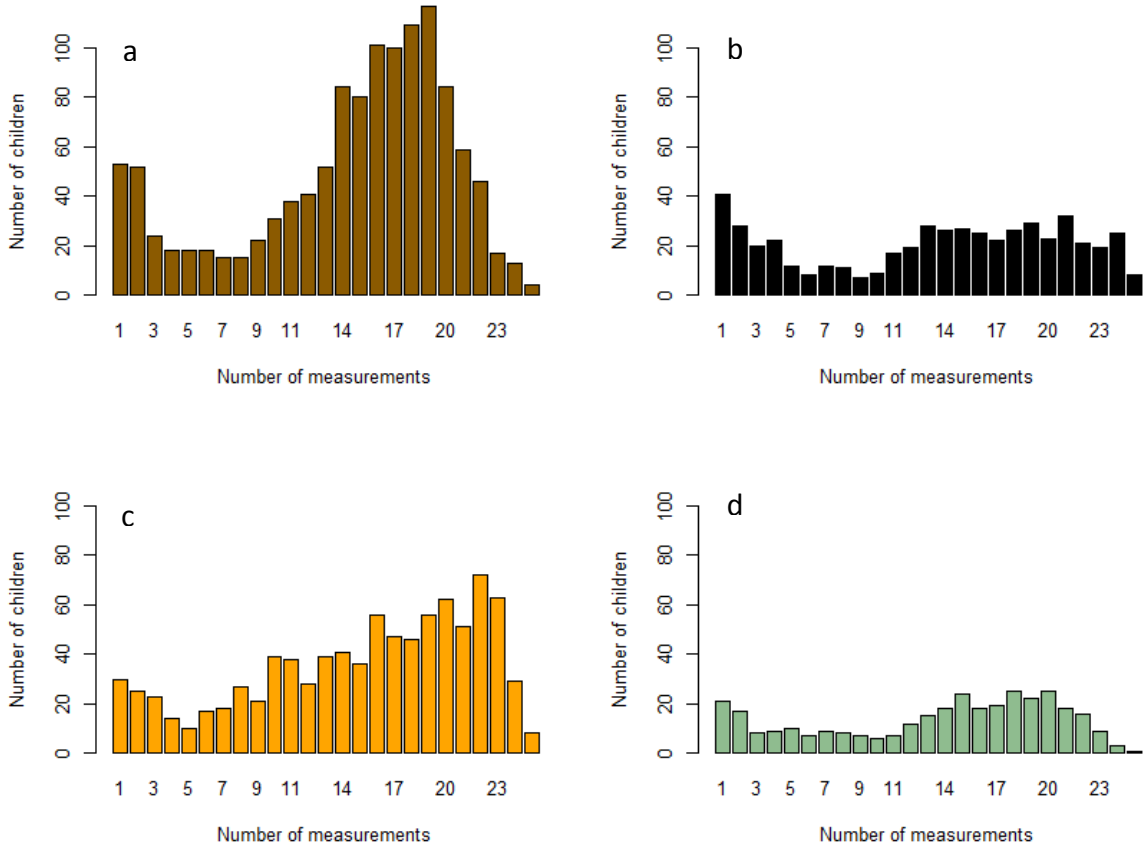

Figure 3.9: Number of children per number of measurements (Pakistan): (a) village, (b) peri-urban slum, (c) urban slum, (d) middle class

The number of measurements per number of children within each area can be seen in Figure 3.9. There are minimal numbers of repeat measurements per child within each inclusion region when counting the number of measurements at each scheduled measurement date. Calculations were made by taking the number of observations between midpoints of the scheduled measurement dates. Note that if no measurements were taken for a given child then they were excluded from the data frame, therefore there are no 'no measurements' for any of those children. The proportion of measurements at each scheduled measurement date can be seen in Figure 3.10, with the total number of measurements equal to the number of children within each cohort.

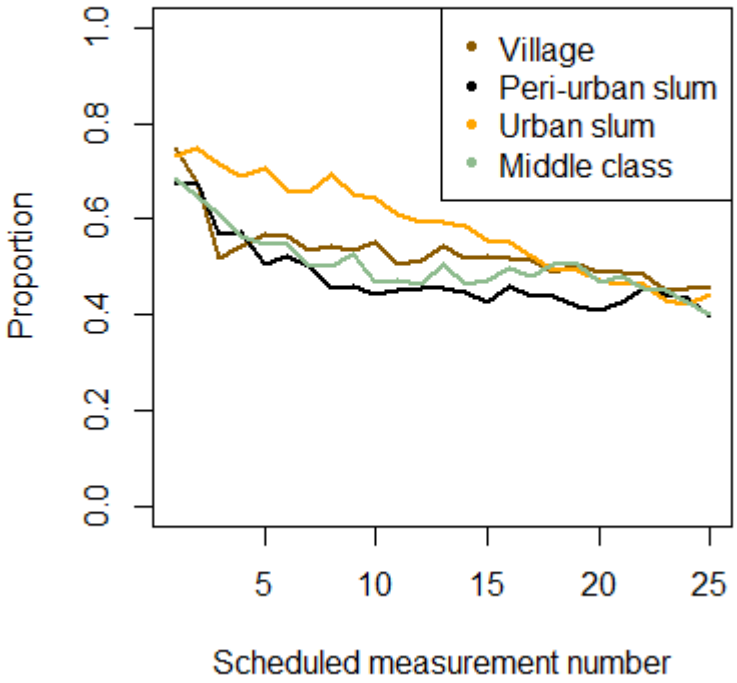

Figure 3.10: Proportion of observations at each scheduled measurement number (Pakistan)

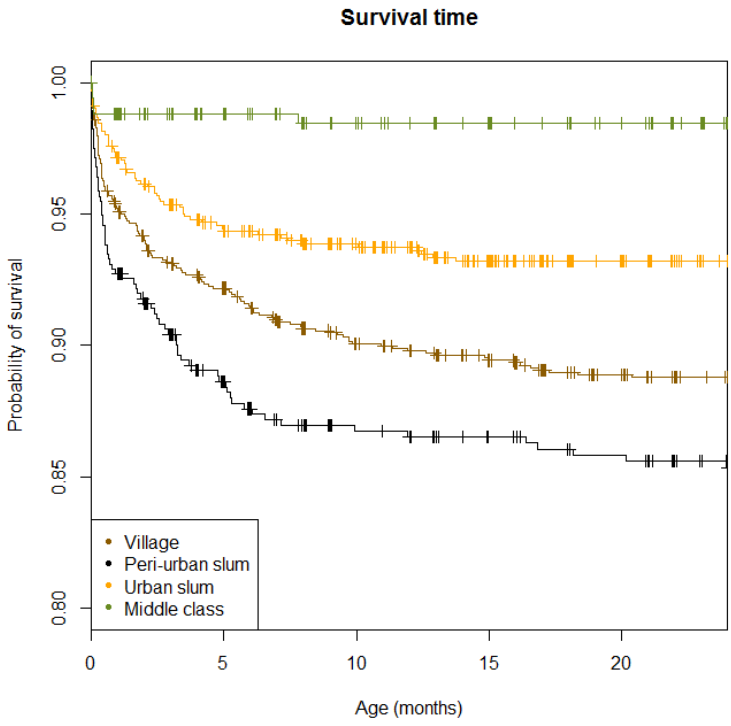

Figure 3.11: Kaplan Meier plot (Pakistan)

The Kaplan Meier plot in Figure 3.11 shows survival curves for the four different areas in Pakistan. Steps indicate deaths whereas crosses indicate right censored data. It is clear that those within the peri-urban slum have the lowest probability of survival at all time points, followed by the village area, urban slum and finally the middle class area.

3.4 Africa Centre Vertical Transmission Study

The third dataset is the Africa Centre Vertical Transmission Study (VTS). The study was designed to examine patterns of infant feeding, specifically exclusive breast feeding associated with HIV transmission. Furthermore, it was designed to determine whether exclusive breastfeeding carries a lower risk of mother to child transmission of HIV compared to mixed breastfeeding (Africa Centre 2007).

The study was conducted from 1st August 2001 until the 31st July 2007, which followed mothers and their children over 2 years after delivery. The study had two field sites, the largest of which was in the Mpukunyoni and Hlabisa Tribal Areas of the Hlabisa Health District (in the rural Umkhanyakude district of northern KwaZulu Natal) where enrolment commenced in August 2001 at 8 clinics. An additional site from Durban was included to increase numbers. The Durban site was urban and the population were very mobile with people migrating at very short notice into and out of the area during pregnancy and post-delivery (Bland et al. 2010).

In total, 3,445 woman were enrolled during pregnancy: 2,705 from the Hlabisa health district (1,586 rural, 1,119 semi-urban) and 740 from the urban site in Durban. 2,938 children were enrolled and 1,833 of them were followed up to 24 months. Socio-economic status and HIV status were collected antenatally. The study covered feeding practices, HIV status, morbidity, mortality, growth and development of children born to HIV infected and uninfected women (Bland et al. 2010). After birth, daily feeding practices and morbidity data were collected for the first 9 months of life and anthropometric measurements were taken at 6, 10, 14, 18, 22, and 26 weeks then at months 7, 8 ,9 ,12, 15, 18, 21 and 24 (Africa Centre 2007). In total, 1,769 (51.4%) of women were HIV positive, 1,662 (48.3%) were negative and 12 (0.3%) were indeterminate. In terms of the highest level of education attained 224 (6.5%) had no formal education, 1,196 (34.7%) had a primary education and 2,025 (58.8%) had a secondary education (Bland et al. 2010).

In terms of feeding choices, most HIV infected women (80%) chose to exclusively breastfeed their children as most did not have the resources for safe replacement feeding (Africa Centre 2007).

Rollins et al. (2008) studied infant feeding, HIV transmission and mortality at 18 months within the dataset. The aim of the paper was to determine the late HIV transmission and survival risks associated with early infant feeding practices. In total, 1193 live born infants were included and studied over 18 months.

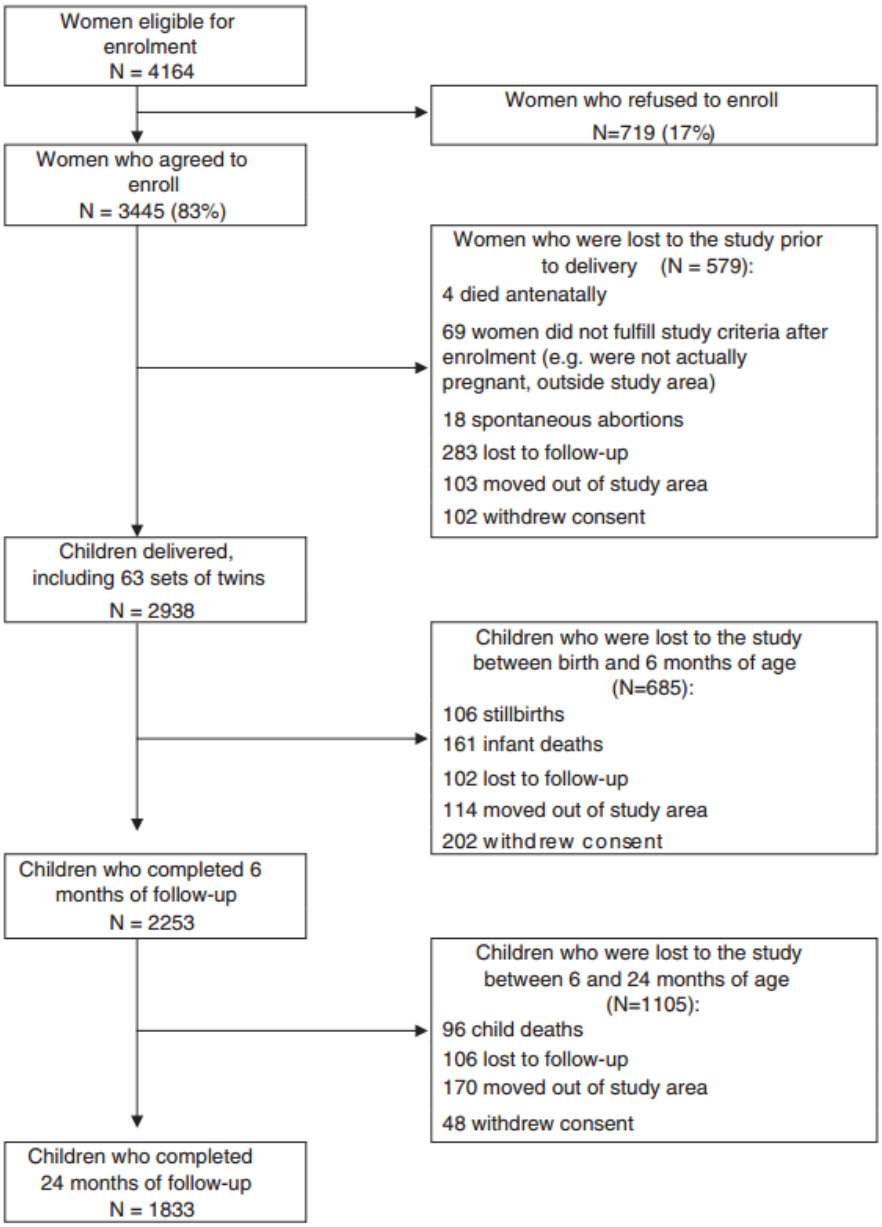

Figure 3.12: Flow diagram of women and children eligible for enrolment and numbers lost to follow-up during the study (Bland et al. 2010).

The probability of death was 0.04 (95% CI 0.03-0.06) and 0.53 (95% CI 0.46-0.60) for uninfected and infected children respectively. It was found that there was no statistically significant difference in 18 month probability of survival between HIV uninfected infants breastfed or replacement fed from birth. The probability of HIV free survival beyond 18 months in children alive at 6 months was 0.98 amongst infants replacement fed from birth. Rollins also conducted a multivariable analysis which found that maternal unemployment and low antenatal CD4 counts (a type of white blood cell that fights infection and their count indicates the stage of HIV or AIDS in a patient) were independently associated with more than threefold increased risk of infant HIV infection or death (Rollins et al. 2008).

Patel et al. (2010) set out to examine the growth of children by maternal and infant HIV status, allowing for infant feeding mode (breastfed/non breastfed), up to 2 years. It was shown that the Z scores for HIV infected children were consistently lower than those of HIV exposed but uninfected children. Breastfed HIV infected infants had consistently higher Z scores for weight (Patel et al. 2010).

In terms of postnatal HIV transmission, the Kaplan-Meier estimated risk of HIV infection between 4-6 weeks and 6 months of age was 4.04% (with a 95% CI of 2.29-5.76%). Those who received solids were significantly more likely to get infected with HIV than those who were exclusively fed breast milk, as were those who received both breast milk and formula milk at 12 weeks. Any HIV infected women who had serious breast health problems (such as bleeding nipple) were 3.55 times more likely to transmit HIV (Africa Centre 2007).

Author, year	Outcome	Risk factors	Results
(Africa Centre 2007)	Postnatal HIV transmission		Kaplan-Meier estimated risk of infection between 4-6 weeks and 6 months of age was 4.04%. Those who received solids and those who were breast fed + formula were significantly more likely to get infected with HIV than exclusively breast fed.
(Patel et al. 2010)	Growth	HIV status, Breastfed/ replacement fed	HIV infected children consistently lower Z scores than non-infected, irrespective of feeding mode. Mean difference of 420g for males, 405g for females at 52 weeks.
(Rollins et al. 2008)	HIV transmission, mortality at 18 months	Breastfed/ replacement fed	$P(dead\ at\ 18\ months|HIV) = 0.53$ $P(dead\ at\ 18\ months|noHIV) = 0.04$ No significant difference in 18 months probability of survival between HIV uninfected breastfed or replacement fed from birth. Probability of survival at 18 months given alive at 6 months with replacement food (HIV free) =0.98 Breastfed for more than 6 months = 0.91 Breasted for less than 6 months = 0.91

Table 3.3: Evidence table: VTS (South Africa)

3.4.1 VTS observational summary

In total, there were 2938 children included in the study. There are 15 scheduled measurement dates within the 0-24 month range.

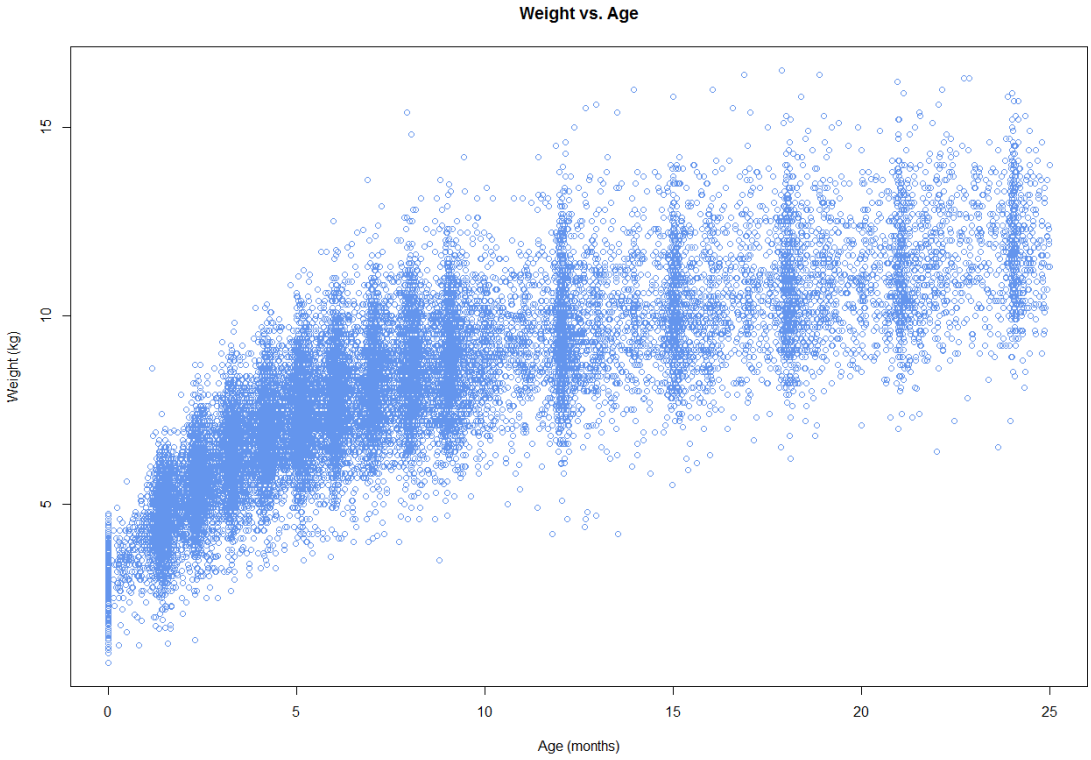

Figure 3.13: Weight vs. Age (South Africa)

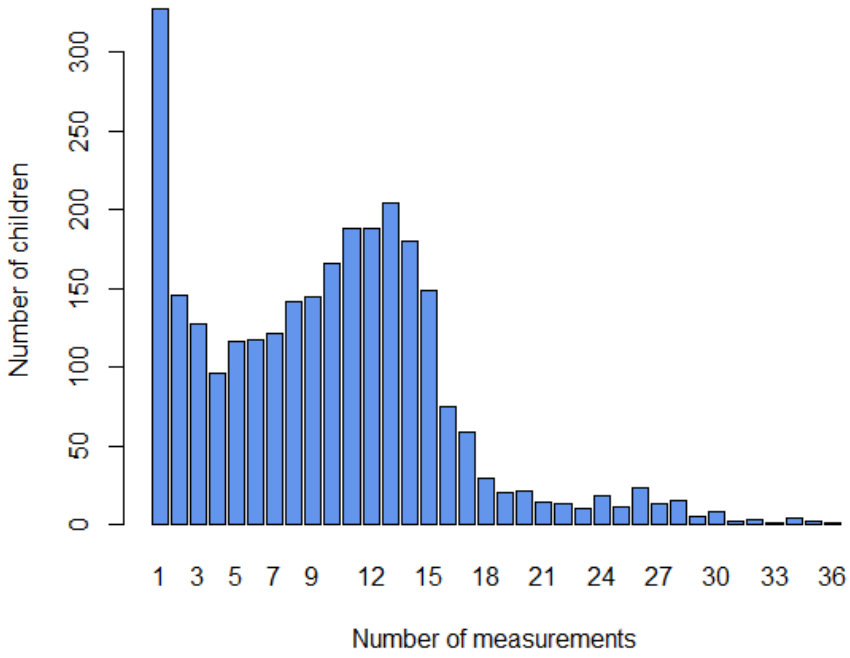

Figure 3.14: Number of children per number of measurements (South Africa)

A plot of weight versus age can be seen in Figure 3.13. The number of measurements per number of children can be seen in Figure 3.14. Many children have been measured more than scheduled (in between measurement dates or after 24 months).

The number of observations at each scheduled measurement date varies and is calculated by taking the sum of the number of observations between the mid points. Within the South African dataset there is a lot of overlap such that one observation included at a certain month, with an associated age close to that month, may actually be a measurement for the month before. This means that the number of observations at each scheduled measurement date may have replicates from the same children.

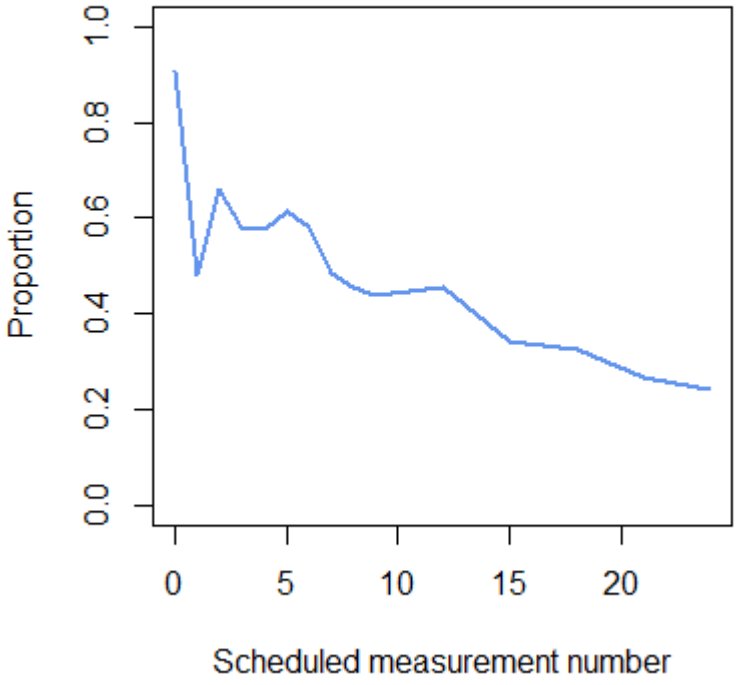

Figure 3.15: Number of observations at each scheduled measurement date (South Africa)

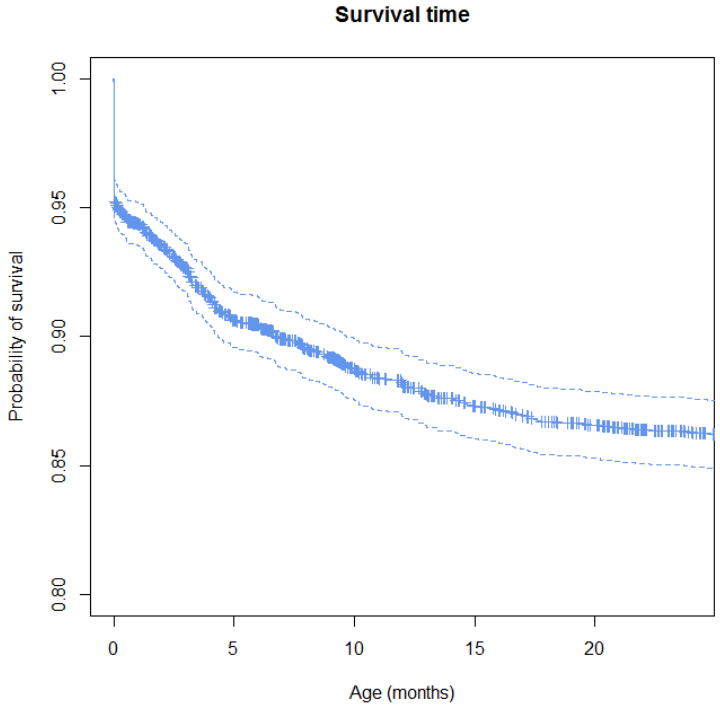

Figure 3.16: Kaplan Meier plot (South Africa)

A Kaplan Meier survival curve can be seen for the South African dataset above in Figure 3.16. Steps indicate where a death has occurred whereas crosses indicate right censored data. Due to the number of still births, there is a large drop in the curve at month 0. The dotted lines represent the upper and lower bounds of a 95% confidence interval.

3.5 Comparison of datasets

In this section, we compare the datasets. The background of those within the datasets are discussed in order of affluence in this section.

The middle class Pakistani cohort was the most affluent. Children from this area fitted criteria including the family earned more than US $275 per month (not adjusted for inflation), owned a house with more than three bedrooms and had access to health facilities (Jalil, Lindblad, Hanson, Khan, Ashraf, et al. 1993).

Mothers of children within the South African Cohort lived in rural, semi-rural and urban areas, 58.8% had a secondary education, 34.7% had a primary education and 6.5% had no formal education. Children in this area had access to health facilities (Bland et al. 2010).

The Pakistani urban slum group consisted of an organised community consisting of small scale businessmen and low-grade government servants. Children in this area had access

to public health facilities however these were under-utilised and were only used in emergencies (Jalil, Lindblad, Hanson, Khan, Ashraf, et al. 1993).

Those in the Malawi cohort were around 5km from local health facilities. Around 70% of children lived in unburned brick houses, 19% in burned brick houses and 11% in mud houses. 77% of children used traditional pit latrines and 23% had no toilet facilities at all. Around 58% of families had kitchens outdoors, only 1% had kitchens indoors (Kulmala et al. 2000). Children in the Pakistani village area had no access to health facilities. The economy of the village was based on agriculture and women frequently helped in the fields (Jalil, Lindblad, Hanson, Khan, Ashraf, et al. 1993).

People in the Pakistani peri-urban slum were nomads. There was no committee to guide or control social life. Health facilities were available, but these were not used (Jalil, Lindblad, Hanson, Khan, Ashraf, et al. 1993).

Table 3.4 compares the three studies' objectives, with numbers included in the study and mortality rates. The Pakistani dataset is the largest with 3146 children in total however this is split into four cohorts. The South African dataset is the second largest with 2938, followed by the Malawian dataset with 813.

Coupled with Figure 3.17b, it is the middle class area of Pakistan that has the best survival rate at 24 months, followed by the urban slum, the village, South Africa, peri-urban slum then Malawi.

			Survival rates			
Dataset	n	Purpose of study	6m	12m	18m	24m
Malawi	813	Designed to provide data on the health of both pregnant women and their children under 5 years old.	0.86	0.82	0.81	0.80
Pakistan: Village	1314	To "characterise the	0.91	0.89	0.89	0.88
Peri-urban	572	determinants of child	0.87	0.86	0.86	0.85
Urban slum	921	health in a rapidly	0.94	0.94	0.93	0.93
Middle class	339	urbanising community".	0.98	0.98	0.98	0.98
South Africa	2938	To examine patterns of infant feeding, specifically exclusive breast feeding associated with HIV transmission.	0.90	0.88	0.86	0.86

Table 3.4: Comparisons of the three datasets

A variable which is used later on within this thesis is HIV status. HIV status at each month was available for children within the South Africa dataset. Within the Pakistani dataset, the data provider specified that there were no cases of HIV within any of the subsets. Within the Malawi dataset, no data were available to indicate whether children had HIV at any point in their lives. However, maternal HIV status at birth was available for the Malawi dataset. Both HIV status and maternal HIV status is incorporated into survival analyses within Chapter 6.

		% missing in month				
Dataset	Scheduled measurement dates	3	6	9	12	24
Malawi	Every month to 18 months, then months 21 and 24. (21 in total)	0.27	0.29	0.28	0.28	0.32
Pakistan: Village	Every month. (25 in total)	0.46	0.47	0.45	0.46	0.54
Peri-urban		0.43	0.49	0.55	0.54	0.60
Urban slum		0.31	0.34	0.36	0.40	0.56
Middle class		0.39	0.45	0.51	0.48	0.58
South Africa	Weeks: 6, 10, 14, 18, 22, 26. Months: 7, 8, 9, 12, 15, 18, 21, 24. (15 in total)	0.39 (week 14)	0.39 (week 26)	0.56	0.66	0.85

Table 3.5: Scheduled measurement dates and missing values

Scheduled measurement dates and percentage of missing at a selection of months can be found in Table 3.5 The Pakistani set has the most scheduled measurement dates followed by Malawi then South Africa.

The percentage missing was calculating by using a search algorithm, allowing us to determine whether for each scheduled measurement date, at least one measurement lay within 9.12 days (0.025 years). While datasets like the Malawi dataset which was set up in such a way that there were as many rows per child as scheduled measurement dates, datasets such as the South African dataset was set up such that if data was missing at a scheduled date then no NA value would be recorded, so there were variable number of rows per child. We therefore had to search for measurements within the specified range. Furthermore, measurements were sometimes taken in between measurement dates making it hard to determine which scheduled date each measurement belonged to. As

there were no scheduled measurement dates for the South African dataset at months 3 and 6, weeks 14 and 26 were used instead.

The median weight curves in Figure 3.17a show that the middle class and South African datasets were heaviest, followed by the urban slum, Malawi/village and finally the peri-urban slum.

One problem within all datasets is that missing values may not be missing at random. This means that there may be missing values dependent on the unobserved data, and may be subject to bias.

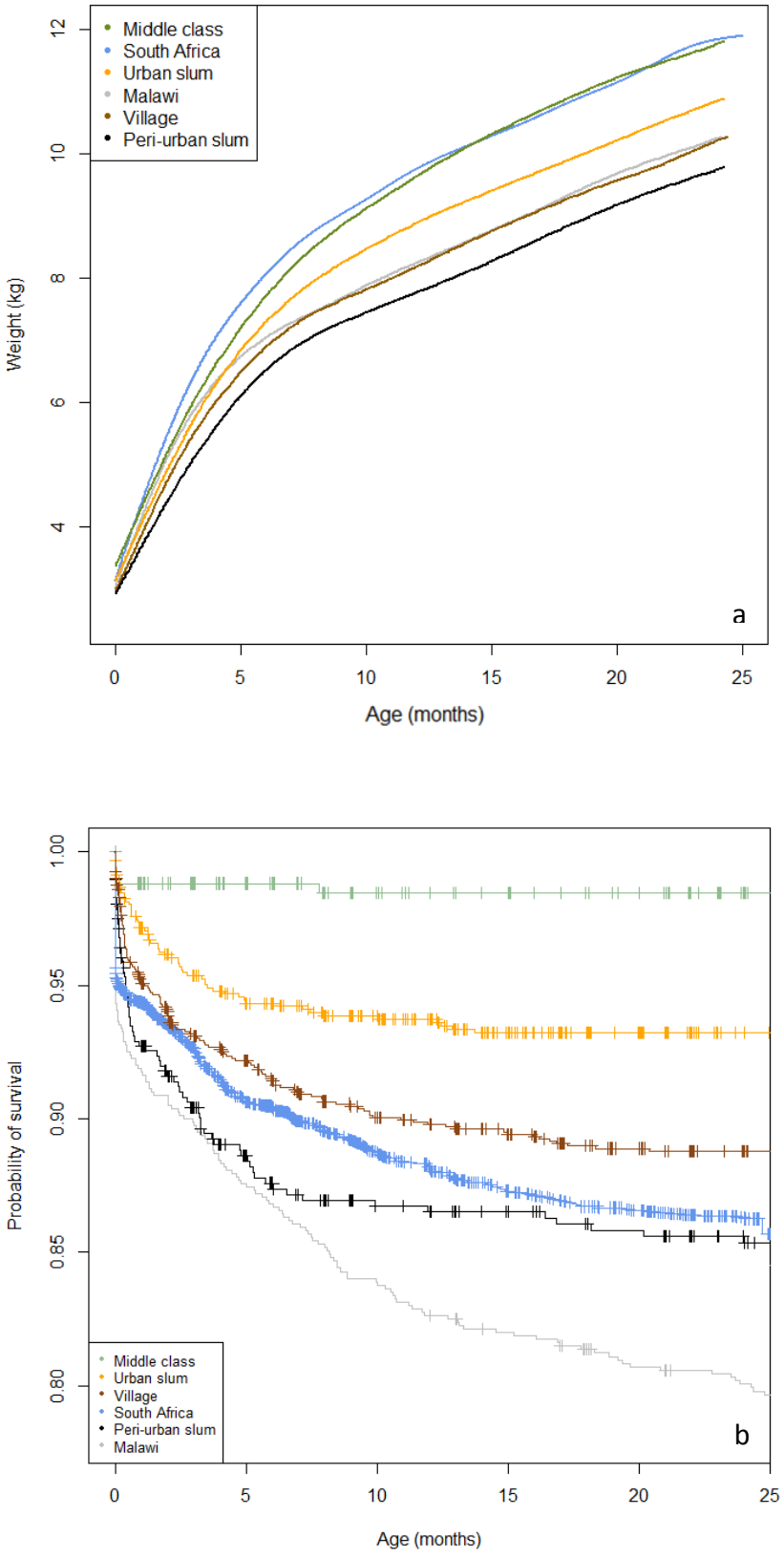

Figure 3.17: (a) Median (μ) curves, (b) Kaplan Meier curves

3.6 Internal and external references

In this section we create growth reference charts for the datasets. Growth chart modelling allows us to compare a given child's measurements to a population of children. This allows us to determine where that child sits on the chart relative to children of the same age group. When modelling growth charts, we assume that the reference population is normally distributed (perhaps after transformation). Therefore by comparing a given child's weight measurement with the reference, we can express as a Z score or centile where a child sits in relation to that specific reference.

GAMLSS, described in Section 2.4, allows us to create growth charts. These charts allow us to compare children with populations of children. By using Equations (2.6) and (2.7), we can also superimpose centile curves on charts. Centile curves allow us to map the growth chart such that certain proportions of the data lie below certain centiles. For example, roughly 45% of the data should lie below the 45^{th} centile.

3.6.1 Internal references - application of GAMLSS

By modelling the growth chart of the given child's parent population via GAMLSS, we can obtain an estimate of where that child sits relative to that population. The result is three smooth curve estimates of $\mu(t) = M(t), \sigma(t) = S(t)$ and $\lambda(t) = L(t)$ where $0 \leq t \leq 24$ (months). This allows us to convert any measurement to a Z score based on $M(t), S(t)$ and $L(t)$ at time t, and also allows us to calculate centile curves throughout the series. By then applying the estimates of those parameters to Equations (2.4) and (2.5), we can obtain a Z score to determine which centile they lie on.

The smoothness of each curve is based on the values of the smoothing parameters of the curves, $\alpha_\lambda, \alpha_\mu$ and α_σ (see Equation (2.13)), which need to be chosen. Cole & Green (1992) state "Fitting smooth centile curves has always been something of a subjective exercise, or even a black art" as certain features of a perceived growth pattern may be down to sampling error, rather than a true pattern. The tuning parameters $\alpha_\lambda, \alpha_\mu$ and α_σ can be chosen in a number of ways. We can specify the number of degrees of freedom for L, M and S, however as mentioned this is subjective. If the number of degrees of freedom is set too low, growth features may not be captured. If they are set too high, we over fit, resulting in a model which captures the random features of the sample.

While using the GAMLSS package, it is recommended that the function $pb()$ is used which is a function based on penalised beta splines (P-splines) of Eilers & Marx (1996). The

function $pb()$ uses local maximum likelihood estimation to automatically choose $\alpha_\lambda, \alpha_\mu, \alpha_\sigma$ (Rigby & Stasinopoulos 2013). We use penalised beta splines, utilising local maximum likelihood to automatically choose our tuning parameters.

Cole (1990) states that the centile curves must be 'properly calibrated', meaning that the proper proportions of the population must lie within certain bands of the centile curves. One way to check if this is done correctly is to express each measurement as a Z score and check if they are distributed as $N(0,1)$ throughout the age range. This was done for all the growth charts that we developed.

A brief explanation of the output is presented for the Malawi growth chart. The other charts are presented without text as the output is similar.

3.6.1.1 LCSS (Malawi)

The first application of GAMLSS was made to the Malawian dataset. Shown below in Figure 3.18 is weight versus age in months. Superimposed on the chart data are 7 lines which represent the 3rd, 10th, 25th, 50th, 75th 90th and 97th centiles.

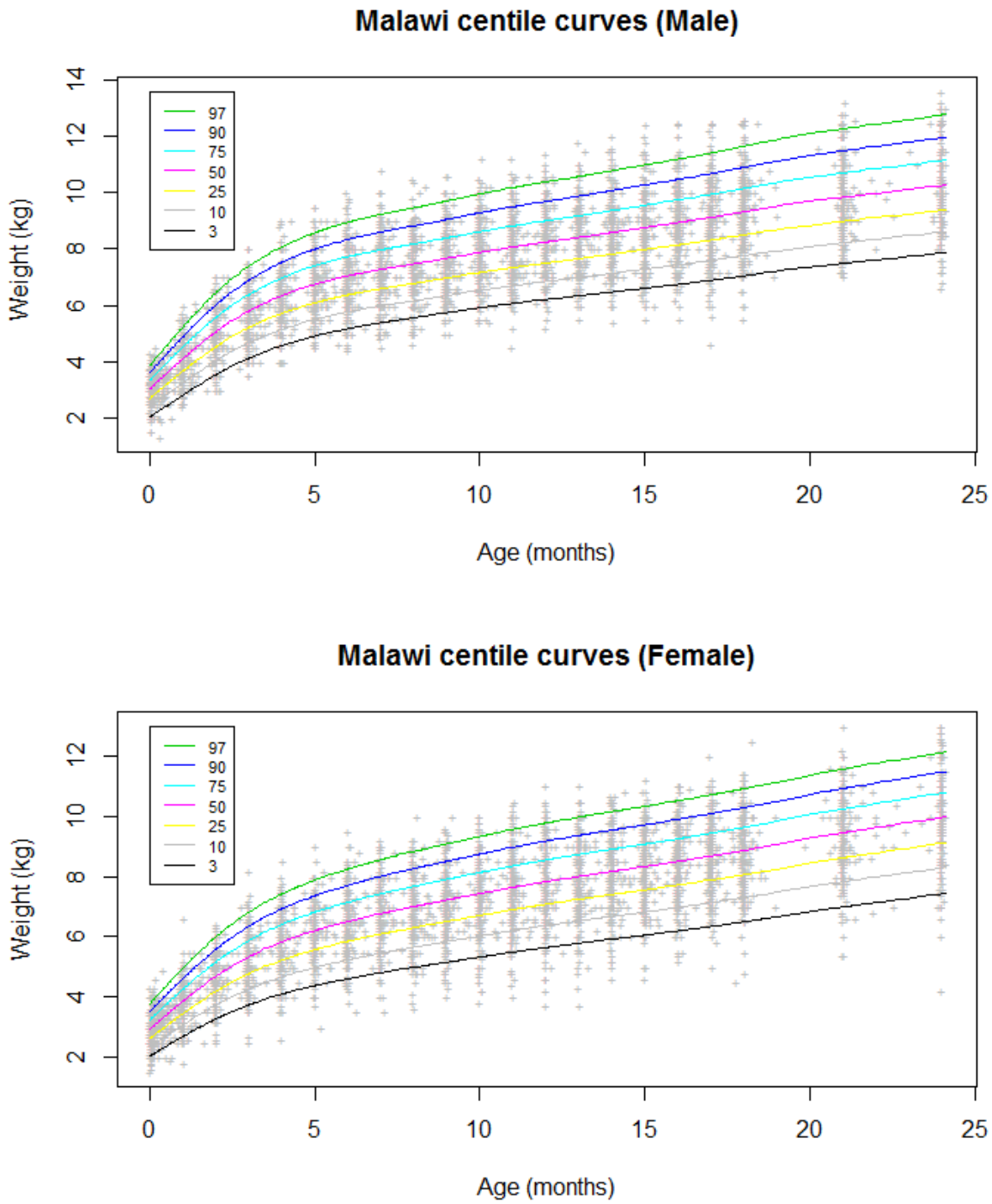

Figure 3.18: Weight vs. Age with superimposed centile curves (Malawi)

Since the LMS method takes both variability and skewness into account, all of the centile lines should accurately depict the proper proportions of data below the curves. For example, around 3% of the data should lie below the green 3rd centile curve, 25% should lie between the 25th centile curve, 75% should lie below the 75th centile curve and 97% should lie below the 97th centile curve. The proportion of measurements which lay under each of the centiles shown in Figure 3.18 were calculated. These proportions were very close to the implied centiles.

Figure 3.19 shows the L, M and S curves for the Malawi dataset. The M curve corresponds to the 50th centile which can be seen in Figure 3.18 (pink line).

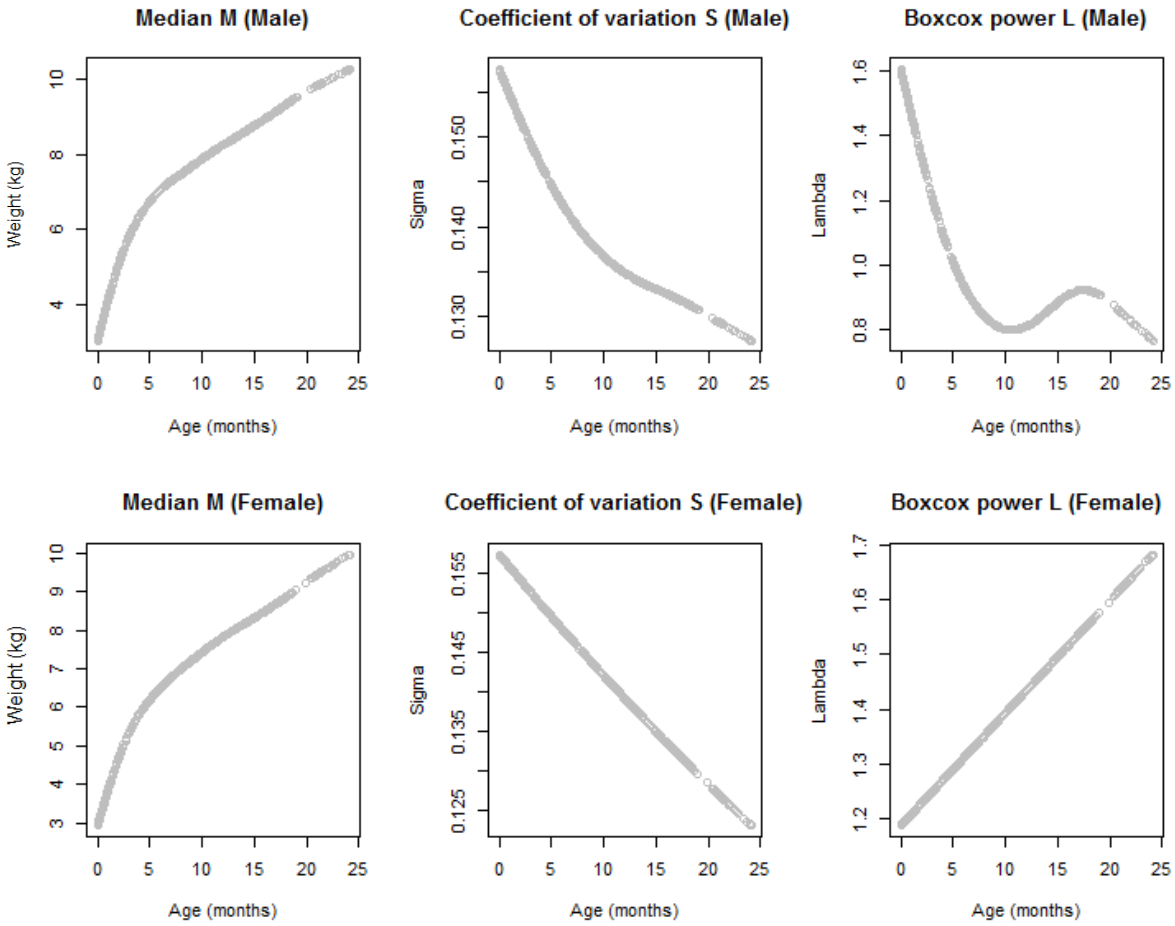

Figure 3.19: L, M and S curves (Malawi)

The number of effective degrees of freedom can be seen in Table 3.6 below, as chosen by the $pb()$ function. In total, there were 18.41 degrees of freedom for males where the M curve had 11.06, the S curve had 3.41 and the L curve had 3.93. There were 10.52, 2.02 and 2.00 degrees of freedom for M, S and L curves for females.

Source	Degrees of freedom (Male)	Degrees of freedom (Female)
Mu	11.06	10.52
Sigma	3.41	2.02
Lambda	3.93	2.00
Total	18.41	14.54

Table 3.6: Degrees of freedom (Malawi)

Notably, the S and L curves of the female model are straight, with only around 2 degrees of freedom. The Box-Cox power for the male model falls with age, implying that skewness decreases with time, whereas for the female model the skewness is rising.

3.6.1.2 LLS (Pakistan)

Area	Source	Degrees of freedom (Male)	Degrees of freedom (Female)
Village	Mu	11.20	10.22
	Sigma	3.24	2.03
	Lambda	2.00	2.99
	Total	16.45	15.25
Peri-urban slum	Mu	8.60	8.65
	Sigma	4.05	2.01
	Lambda	2.00	3.01
	Total	14.66	13.68
Urban slum	Mu	9.93	9.57
	Sigma	2.02	2.02
	Lambda	3.00	2.29
	Total	14.96	13.89
Middle class	Mu	8.50	8.09
	Sigma	3.70	3.95
	Lambda	3.75	3.33
	Total	15.96	15.38

Table 3.7: Degrees of freedom (Pakistan)

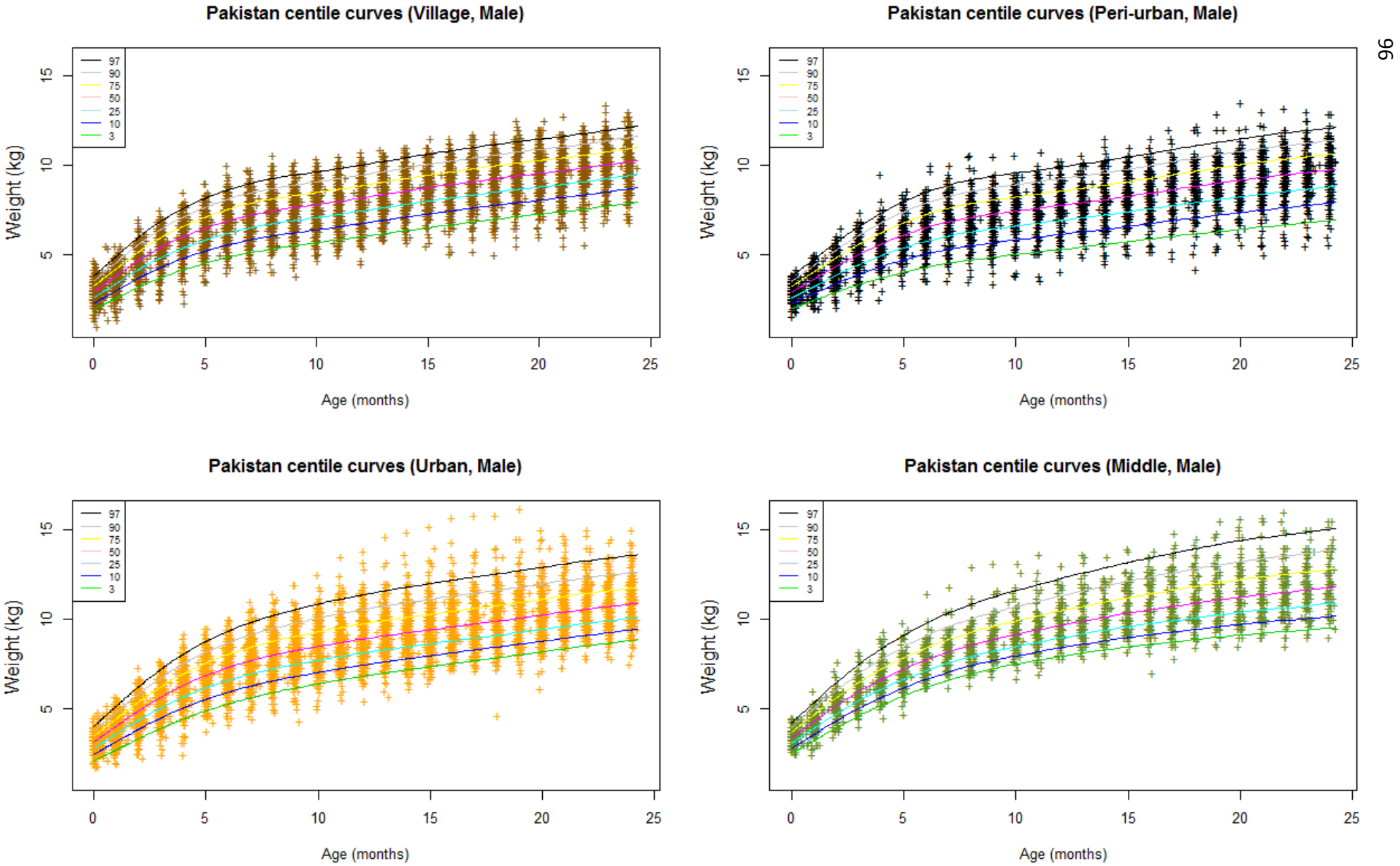

Figure 3.20: Weight vs. Age with superimposed centile curves (male, Pakistan)

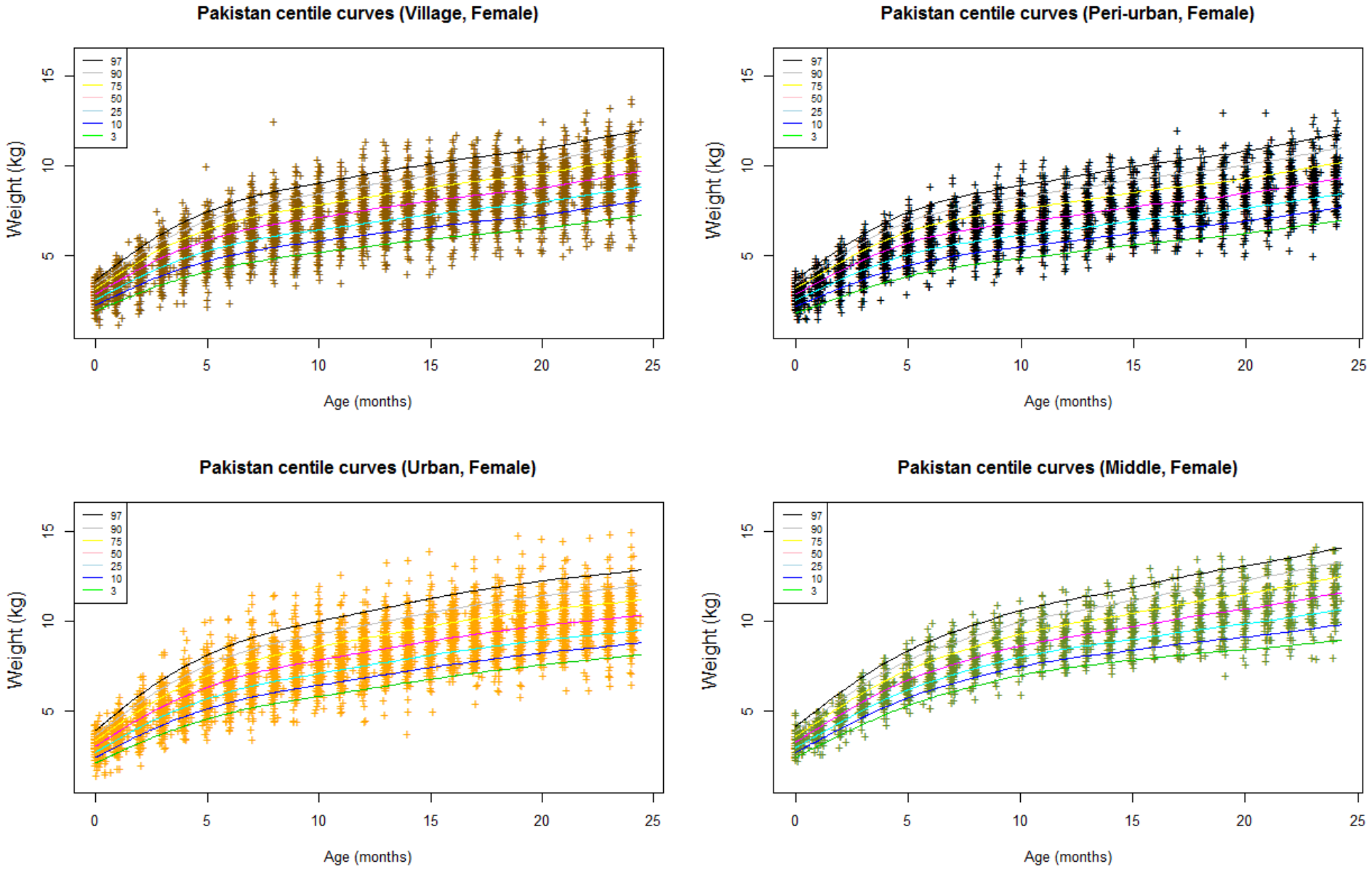

Figure 3.21: Weight vs. Age with superimposed centile curves (female, Pakistan)

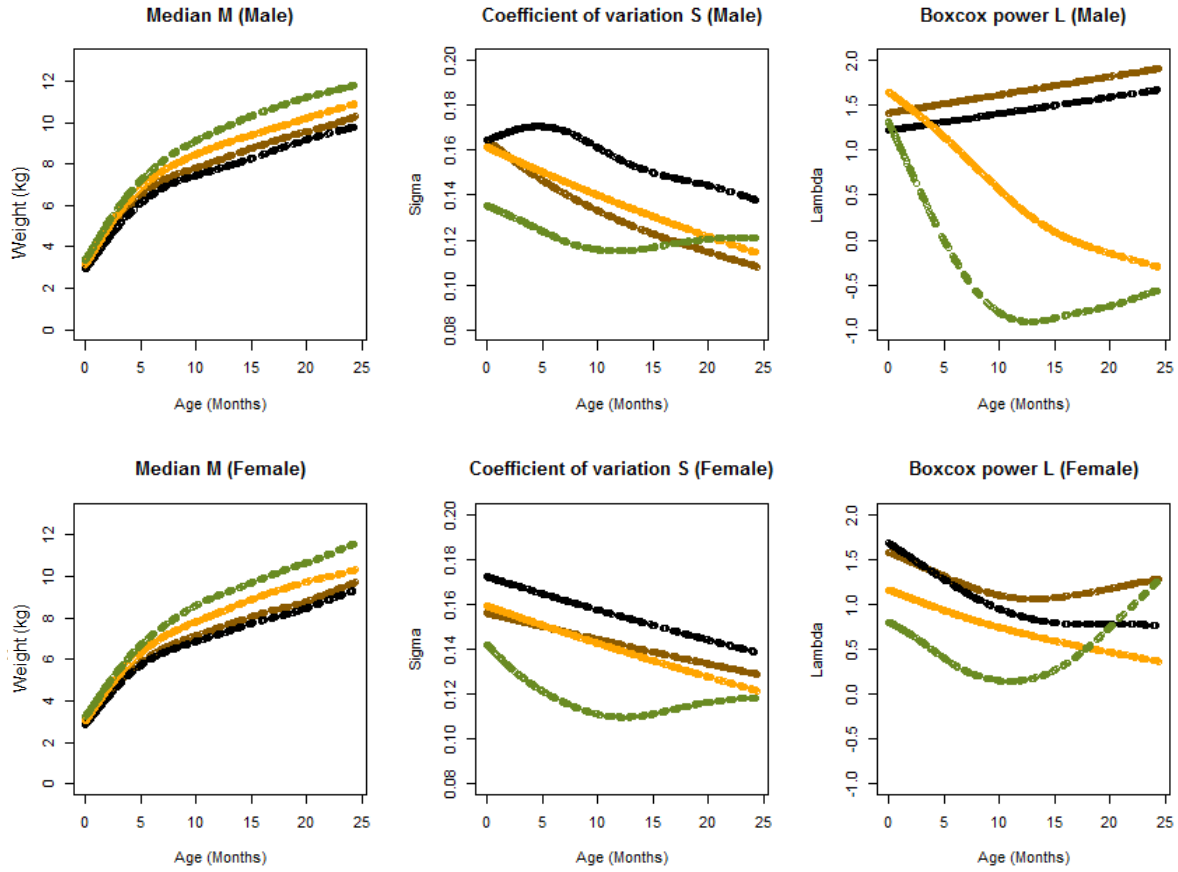

Figure 3.22: Comparison of L, M and S curves (Pakistan)

3.6.1.3 VTS (South Africa)

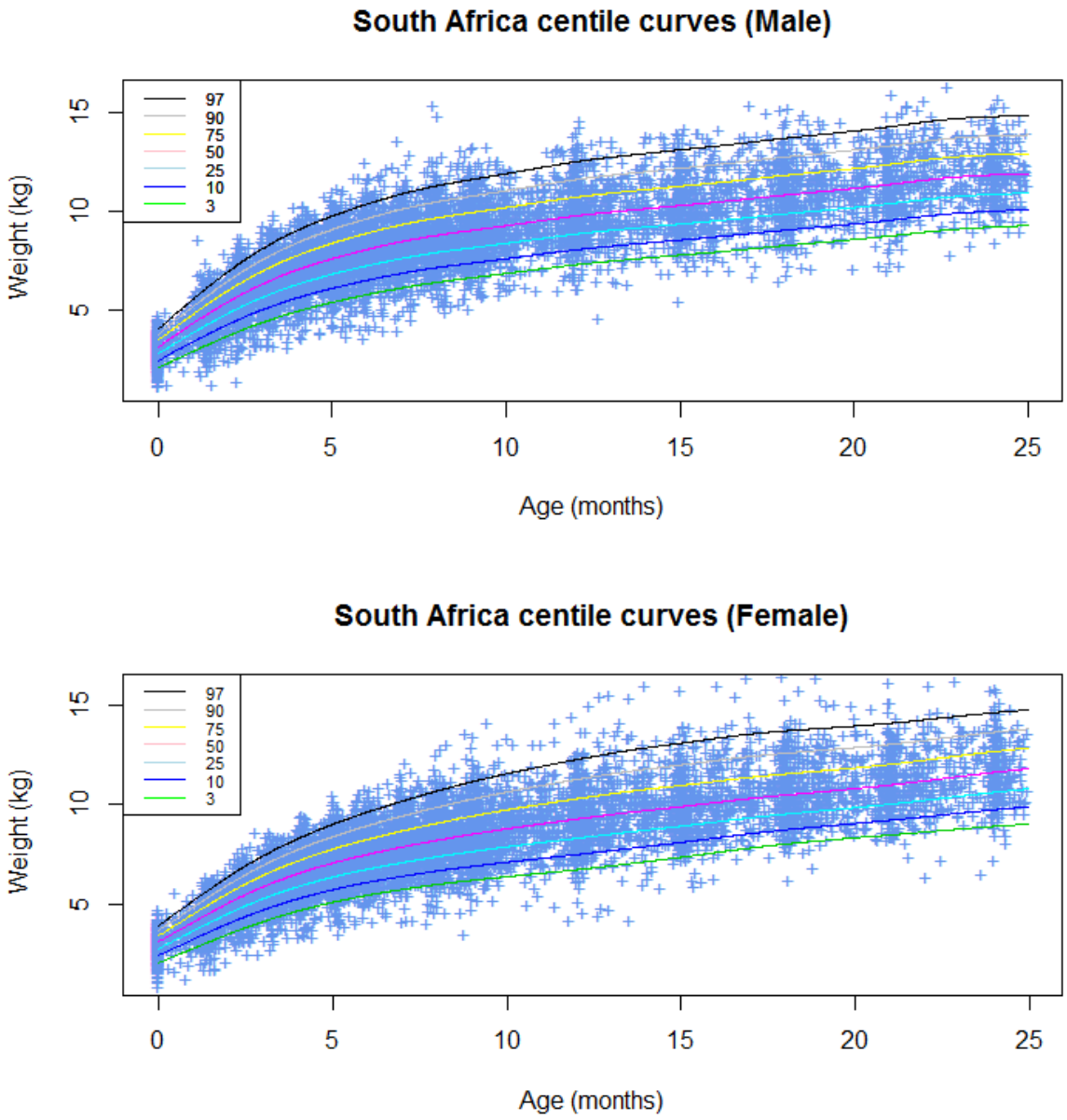

Figure 3.23: Weight vs. Age with superimposed centile curves (South Africa)

Source	Degrees of freedom (Male)	Degrees of freedom (Female)
Mu	11.50	10.86
Sigma	2.03	5.76
Lambda	3.27	5.27
Total	16.81	21.90

Table 3.8: Degrees of freedom (South Africa)

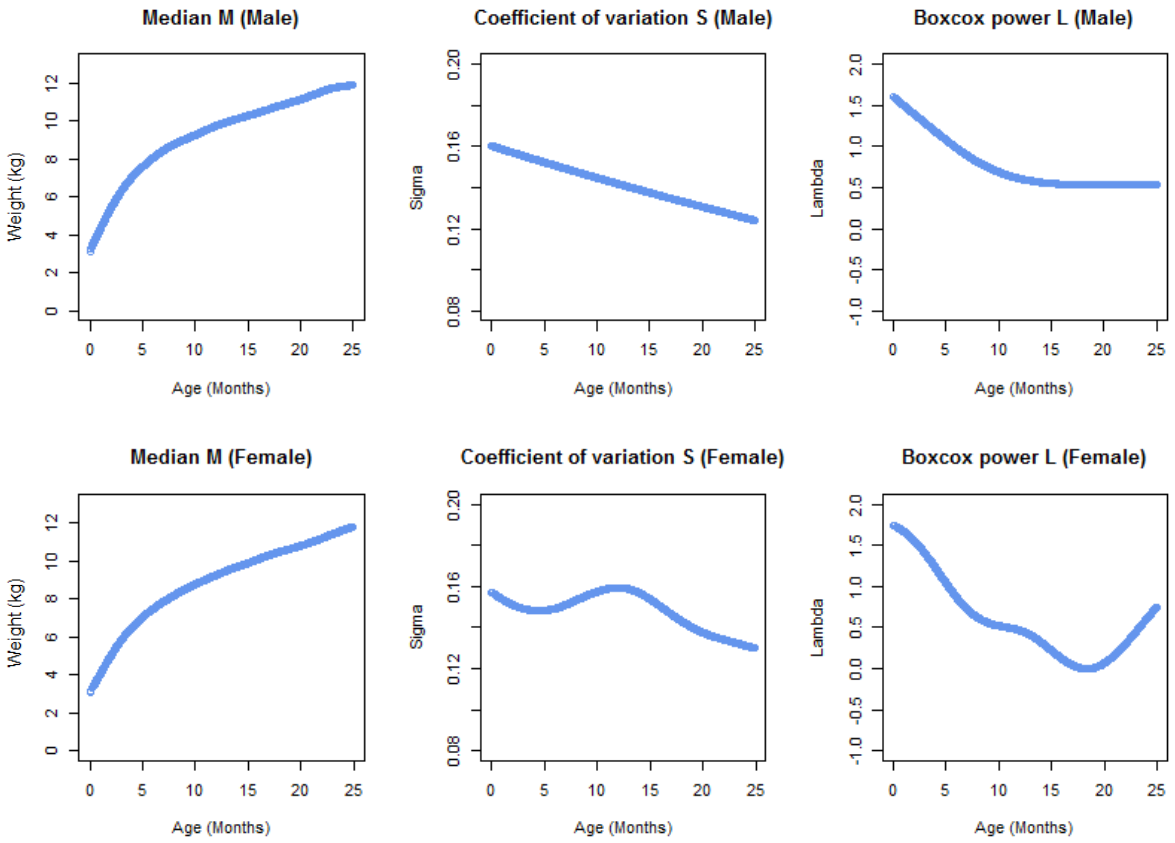

Figure 3.24: L, M and S curves (South Africa)

3.6.2 External references – the WHO standard

The internal references shown in Section 3.6.1 allow us to describe growth within each of the datasets. These can be used to calculate weight-for-age Z scores for individual children. It is however of interest to determine where these children from the developing world sit relative to external populations, such as the WHO standard, as it is normal healthy children that should be used as a reference population.

Modelling boys and girls together would increase the sample size, however, this would increase variance since girls and boys grow at different rates on average.

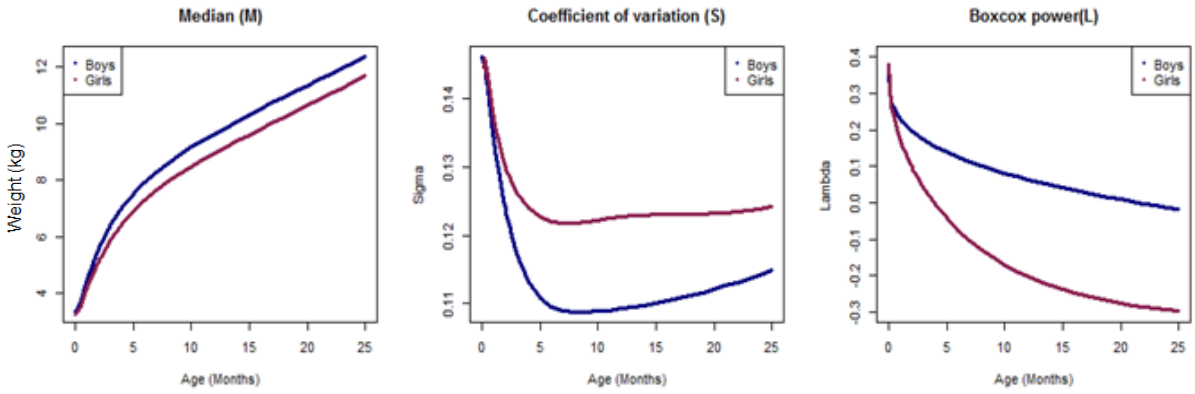

Figure 3.25: L, M and S curves (WHO)

The WHO L, M and S curves can be seen in Figure 3.25, separately for boys and girls, with navy representing boys and purple representing girls. It can be seen from the plot of the medians versus age that the boys have a slightly higher median than girls.

3.6.3 How well do pre-school children from the developing world fit the WHO standard?

Weight measurements were converted to Z scores relative to the WHO LMS model using Equations (2.4) and (2.5), creating weight-for-age Z scores. We then applied GAMLSS to each of the transformed series, assuming a normal distribution for the response, modelling both the mean and standard deviation. This allowed us to gain insight into how children grow on average relative to the WHO standard.

The mean curves are plotted within Figure 3.26 and standard deviations in Figure 3.27. The GAMLSS model used to model the transformed Z scores generates output for $N(\mu_t, \sigma_t)$ such that $\mu_t = M(t)$ is equal to the mean and $\sigma_t = S(t)$ is equal to the standard deviation at time t. The degrees of freedom for each curve can be found in Table 3.9.

Within Figure 3.26 values close to 0 indicate that the measurements for the corresponding dataset lie close to the population median of the WHO dataset, while low values indicate that the measurements of the corresponding dataset lie far away from the WHO median curve.

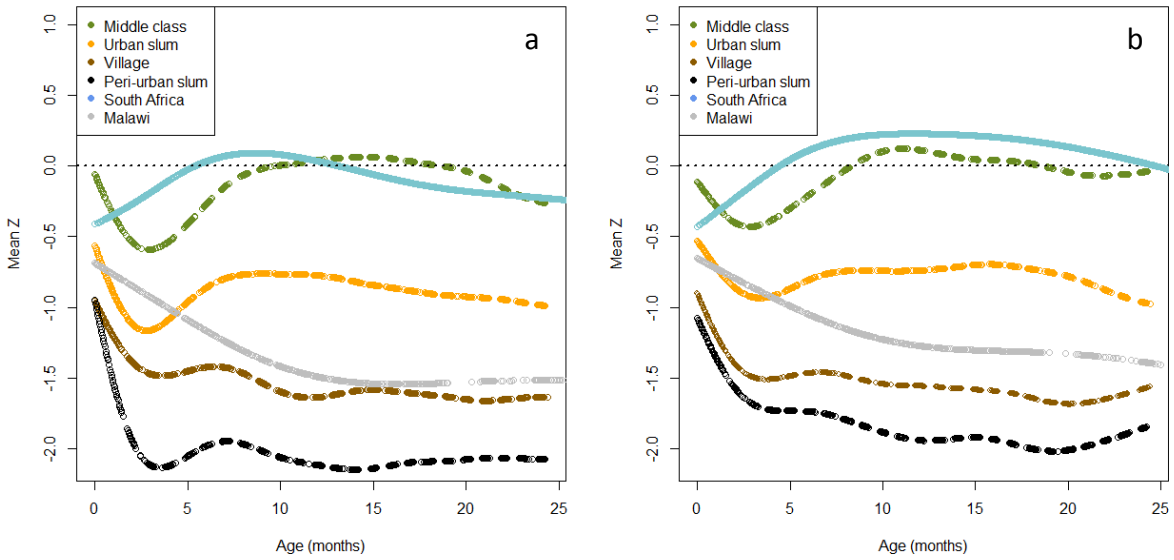

Figure 3.26: Mean Z scores (WHO) vs. Age (a) male, (b) female

As expected, of the 4 Pakistani areas, the peri-urban slum deviates most from 0, followed by the village area, then the urban slum and finally the middle class area. Growth relative to the WHO standard is slow for these datasets in the first 3 months, however all 4 cohorts show catch-up until around 6 months, especially boys. From 6 months onwards growth is roughly on par with the standard but each cohort is a number of SD below 0 (peri-urban slum: -2SD, village: -1.5SD, urban slum: -1SD, middle class: 0SD).

Both South African males and females are around -0.5SD at birth but catch up with the standard around month 6. Notably, South African females outgrow the standard until around 12 months where the rate of growth gradually slows until 2 years.

The mean curve for those in the Malawi dataset exhibits a similar pattern for both males and females, showing a steep decline within the first year from around -0.5SD to around -1.5SD.

	Boys		Girls	
Area	μ_{Z_t}	σ_{Z_t}	μ_{Z_t}	σ_{Z_t}
Malawi	7.92	3.14	7.52	3.84
Village	9.14	5.10	9.11	2.96
Peri-urban slum	9.44	5.70	7.47	2.96
Urban slum	10	5.46	8.20	4.23
Middle class	8.48	2	7.56	3.34
South Africa	7.91	2	7.40	8.12

Table 3.9: Degrees of freedom (of transformed series)

Therefore, children in the developing world tend to grow at different rates relative to the WHO standard in the first 6 months, depending on the background of the child. After 6 months, children parallel the WHO standard tracking along a low line within the normal range, apart from boys in the peri-urban slum. Generally, more children from more affluent areas fit the standard well, tracking along the 50^{th} centile of the WHO standard or above.

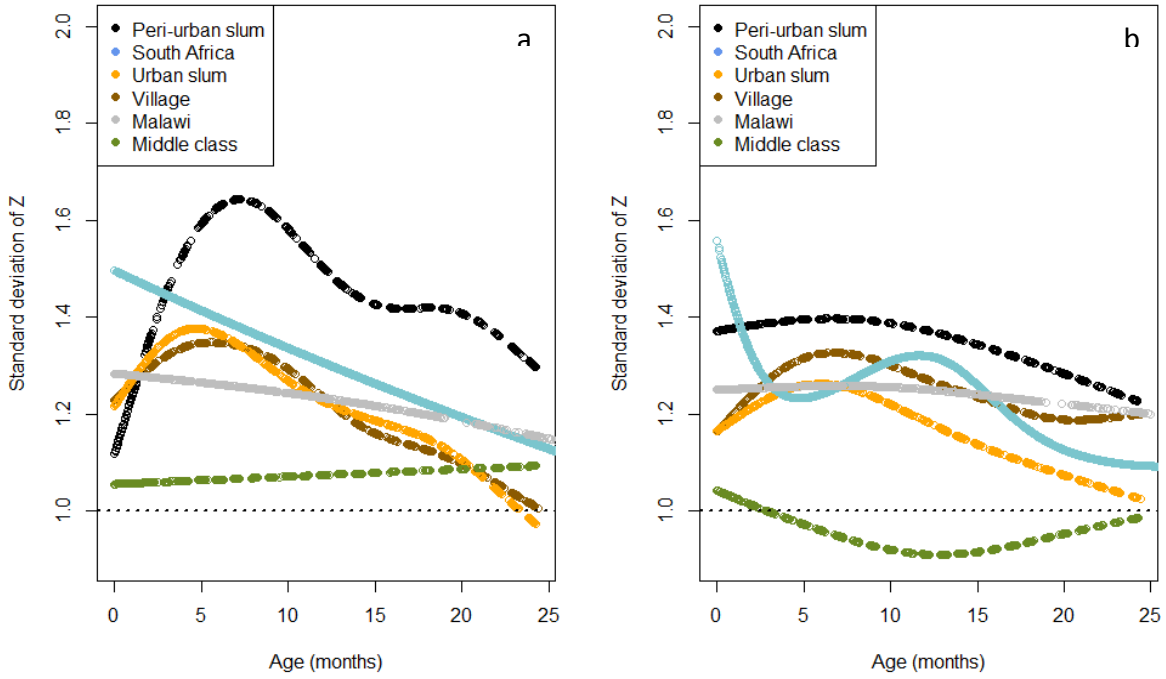

Figure 3.27: Standard deviation of Z scores (WHO) vs. Age: (a) male, (b) female

3.7 Chapter conclusions

This chapter focused on introducing the datasets and creating internal references based on the available data. The datasets were then compared with the WHO standard, allowing us to assess their fit relative to the reference.

Three developing world datasets were introduced - one from Malawi, one from Pakistan (with 4 cohorts) and one from South Africa. Since the Pakistani dataset was collected over a long period of time, we split the data to determine if there were any differences in growth over the two time periods. There were no differences between the two periods.

The internal charts were developed to demonstrate the use of GAMLSS in creating growth references. The three parameter Box-Cox normal distribution was applied within the GAMLSS framework, another term for the LMS method (Cole & Green, 1992), where the L curve represents the Box-Cox power (skewness), M represents the median (location) and S represents the coefficient of variation (scale). The LMS method worked well and the application of the method via GAMLSS provided the appropriate modelling framework for our cohorts, allowing us to accurately represent growth while adjusting for non-normality within the developing world datasets. Local maximum likelihood estimation was used to automatically choose our smoothing parameters for λ, μ and σ via the function $pb()$ within the $gamlss$ package in R, which uses penalised beta splines (piecewise

polynomials) to generate the curves (Eilers & Marx 1996). Penalised beta splines were chosen as smoothing functions as these are recommended by Rigby & Stasinopoulos (2004). However, other methods are available to generate curves such as cubic splines, locally weighted least squares regression, polynomials and fractional polynomials. Cubic splines were applied and we considered presenting these, however the degree of smoothing required is somewhat subjective to the user. We found finding the balance between different degrees of freedom was hard, as our models were either under fitting or over fitting. The automatic procedure found a balance, which resulted in the development of growth curves with the expected proportions below each centile.

The LMS method has been applied to large numbers of datasets since publication. Notable early applications contributed to the development of the original UK 1990 growth reference for BMI (Cole, 1995) height and weight (Freeman et al. 1995). However, other distributions were taken into consideration to model growth within our datasets, such as: the Box-Cox power exponential (BCPE) (Rigby & Stasinopoulos 2004), Box-Cox t (Rigby & Stasinopoulos 2004), Johnson's SU (Johnson 1949) and the Modulus-exponential-normal (Royston & Wright 1998). In fact, the WHO originally used the BCPE distribution to create the WHO standard curves (de Onis & WHO Multicentre Growth Reference Study Group 2006) but reverted to the LMS method. The CDC 2000 growth reference charts were also based on the LMS method (Flegal & Cole 2013). Cole & Green (1992) state that a limitation of the LMS method is that it does not deal with the presence of kurtosis. Kurtosis however tends to be less important than skewness as a contributor to non-normality (Cole & Green 1992).

In this chapter we also assessed how real populations of pre-school children from the developing world fit the WHO standard. The WHO standard was developed using measurements from approximately 8500 children from a wide range of backgrounds and cultural settings. It was designed to represent the best description of physiological growth – an international standard for all children. Previous research which has investigated how populations of pre-school children fit the standard has mainly focused on assessing the fit of children from the developed world.

Roelants (2013), Hui et al. (2008) and Wright et al. (2008) compared Belgian, Norwegian, Hong Kongese and British children with the WHO standard. They found that children from the developed world generally match or outperform children from the MGRS study on average, although Hui found that healthy Hong Kongese toddlers were shorter, which

challenges the assumption that children in all countries can achieve their full growth potential when nurtured under optimal conditions (Hui et al. 2008) since Hong Kong is considered a developed dependency of China.

Our work adds to the literature by focusing on assessing the fit of children from the developing world. To determine where children within the developing world datasets sit relative to the WHO standard, the L, M and S curves were used to create Z scores for every measurement within each population. GAMLSS was then applied to each of these transformed variables, assuming a normal distribution so that the mean and standard deviation for each transformed dataset could be modelled over time. Karlberg (1993) adopted this approach when analysing child growth within the Pakistani cohort using the middle class group as a reference and Maleta et al. (2003) compared the Malawian infant's growth with both the CDC and NCHS references. However, Maleta superimposed centile curves on top of each other, rather than model the transformed data.

In Figure 3.26, departures from 0 can be interpreted as a mean difference from the WHO median at any point over the x axis. Boys and girls share quite similar mean growth trajectories relative to the WHO standard, suggesting that these sets could possibly have been pooled. Including sex as a categorical variable within a pooled model then comparing the t-value with the t-distribution would have allowed us to determine if sex was a significant variable for both the median and SD curves.

Of the 6 mean curves, it is the middle class Pakistani and South African curves which most closely fit the WHO standard. In fact, both of these curves are a reasonable match, outperforming the standard over some portions of the 0-24 month age range. The children in these cohorts are from much more affluent backgrounds compared to the other cohorts, which is most likely the reason for considerably better growth rates. The South African population start around -0.5SD, but quickly catch up with the reference. Notably, South African females outperform the standard from around 4 months onwards. The middle class infants follow a similar trend with the South African cohort in the sense they catch up to the WHO median around month 6. However, they firstly grow slower for the first 3 months. The urban slum shares a very similar trend with the middle class infants (these two are the most affluent of the Pakistani cohorts), showing slower growth than the WHO infants during months 0-3 but catching during months 3 to 6. However the mean is around 1SD below 0 at month 6. These children come from a less affluent

background than the middle class area, but from an organised community with access to health facilities (Jalil, Lindblad, Hanson, Khan, Ashraf, et al. 1993).

The children in the urban slum, therefore, do not fit the standard well, along with the Malawi, village and peri-urban slum cohorts – these children come from even less affluent backgrounds, without the same facilities as those from the middle class and South African cohorts, paralleling the WHO median on average at around -1SD, -1.5SD, -1.5SD and -2SD from 6 months onwards, respectively.

In summary, more affluent populations fit the WHO standard well as children's mean weight-for-age Z scores sit close to the WHO median. In populations which are considered less affluent, children tend to track a low centile within the normal range, indicating these population do not fit the reference.

It should be noted that the design and inclusion criteria of the MGRS shows that in developing countries the standard is an "elite standard", where as in the developed world it is simply a reference for breastfed children (Roelants 2013).

Chapter 4
Modelling correlation matrices
4.1 Introduction

One of the parameters needed within the expression for generalised conditional weight gain, Equation (2.19), is the correlation between groups of measurements as Z scores at the times at which growth of the child was measured. The correlation is needed as this indicates the amount that a child's weight (or height) Z score will be expected to regress towards the mean between t_1 and t_2.

Obtaining the correlation coefficient, r, between time points t_1 and t_2 is simple when t_1 and t_2 are scheduled measurement occasions. This is because each group of measurements corresponds to a single age. A problem arises when t_1 and t_2 lie between scheduled measurement occasions as there are no intermediate reference data. It therefore needs to be estimated. To do this, correlations can be calculated between all ages in which data are available, resulting in a correlation matrix which can be modelled. This model can be used to interpolate correlations between any combination of ages.

This chapter focuses on obtaining correlation matrices for the three datasets, creating useable models to interpolate correlations for use within Equation (2.19). We aimed to develop a general model but found that on the whole, a general model did not fit the individual matrices sufficiently. Heimendinger & Laird (1983) state that correlations should be "pertinent to the target population", suggesting that models used to interpolate correlations should fit well. Therefore, we apply models to the individual datasets and adopt a number of statistical modelling approaches and assess their fit, using criteria presented within Section 2.9. Estimates of the mean and standard deviation

of the Z scores at t_1 and t_2 are also required for Equation (2.19). These can be estimated from the same procedure used to plot Figure 3.26 - measurements can be converted to Z scores relative to the WHO standard and GAMLSS can be applied to the transformed variables assuming a continuous normal distribution, obtaining estimates for $\mu_{Z_1}, \mu_{Z_2}, \sigma_{Z_1}$ and σ_{Z_2}.

4.2 Problems with missingness

If a dataset contains the same number of rows per child and each row represents a scheduled measurement occasion, we simply take our data and calculate correlations for each pair of occasions. This is because correlations can be calculated between vectors of the same length.

For studies that have varying number of measurements per child, like the South African dataset, grouping data is a little more complicated. Children may have missing measurements or multiple measurements around the same scheduled measurement date.

A simulated example of this can be seen in Figure 4.1 where 3 simulated children's measurements are plotted against their population. The child represented by the blue line has only 4 rows to represent 4 weight measurements, the red line represents a child with 18 measurements. However, measurements are not necessarily at consecutive months. The green line represents a child with 23 measurements with no value at month 18. This results in three vectors of weights of length 4, 18 and 23. Within the example, the correlation between measurements at months 3 and 4 can be calculated. However correlations cannot be calculated between months 4 and 5, as the vectors are of varying length.

To resolve this problem, we follow a rule where if a child's measurement lies within 0.025 decimal years (9.125 days) of the scheduled age, that measurement is taken as the measurement for that child at that date. If more than one measurement is within the interval for a specific scheduled date, one of the measurements are taken at random and selected as the measurement. If no measurement lies within 0.025 decimal years of the scheduled measurement date, an NA value is stored for that child at that scheduled measurement date. Obviously, one problem with this is that much of the data are lost in this process. However, for a fair comparison of correlation matrices this procedure must

take place. The choice of 0.025 years was found to balance the loss of data and minimisation of standard error.

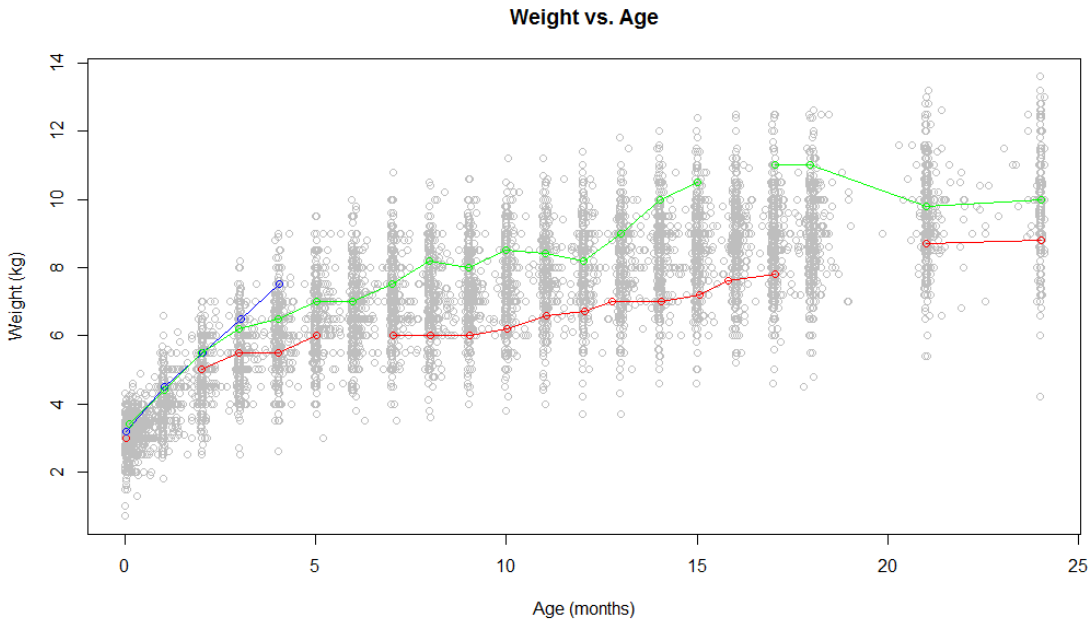

Figure 4.1: Weight vs. Age with simulated superimposed individual growth tracking (Malawi)

4.2 Producing correlation matrices

4.2.1 Malawi

For the Malawi dataset calculating correlations is relatively easy. Data tends to be closely grouped around the scheduled measurement dates and there is the same number of measurements per child as there are scheduled measurement dates.

Following our rule for obtaining the MSE in Section 2.9.1, we remove ¼ of the correlations in our matrix and model using the remaining correlations ($\frac{1}{4}.210 = 52.5$). Therefore, we will remove 53 correlations and predict these using our models (based on the remaining 158), then calculate the MSE using the residuals.

The correlation matrix of Z scores for this dataset can be seen in Table 4.1, ranging from age 0 to 24 months. This shows the correlation between all combinations of ages. A contour plot of the correlation matrix can be seen in Figure 4.2.

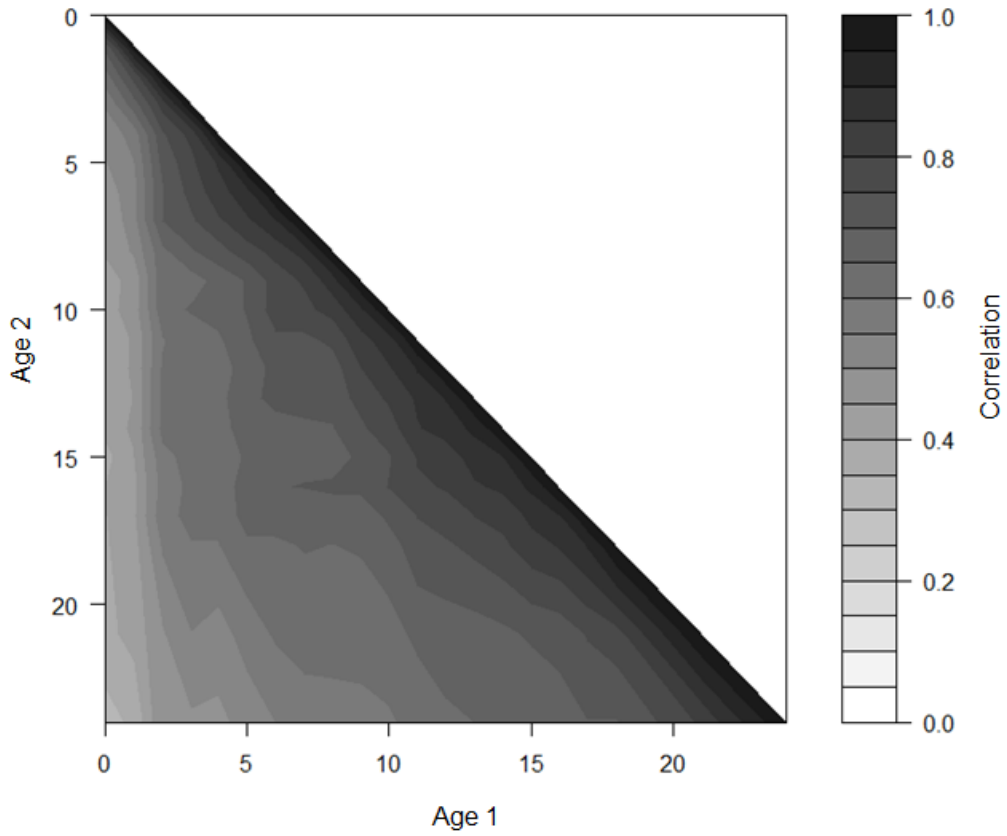

Figure 4.2: Correlations between t_1 and t_2 (Malawi)

Age	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	21	24
0	1																				
1	0.72	1																			
2	0.63	0.77	1																		
3	0.56	0.67	0.82	1																	
4	0.50	0.57	0.74	0.81	1																
5	0.49	0.54	0.71	0.78	0.85	1															
6	0.46	0.53	0.69	0.76	0.82	0.88	1														
7	0.44	0.53	0.69	0.74	0.79	0.84	0.89	1													
8	0.45	0.50	0.63	0.69	0.73	0.77	0.81	0.87	1												
9	0.42	0.47	0.62	0.63	0.66	0.70	0.77	0.80	0.87	1											
10	0.43	0.47	0.62	0.65	0.66	0.70	0.77	0.78	0.82	0.89	1										
11	0.41	0.45	0.59	0.62	0.64	0.69	0.74	0.73	0.76	0.83	0.87	1									
12	0.40	0.45	0.60	0.61	0.63	0.66	0.72	0.70	0.72	0.80	0.84	0.90	1								
13	0.40	0.45	0.60	0.62	0.63	0.68	0.71	0.70	0.72	0.76	0.80	0.85	0.90	1							
14	0.41	0.46	0.61	0.61	0.62	0.68	0.68	0.69	0.69	0.74	0.77	0.84	0.87	0.91	1						
15	0.38	0.44	0.58	0.61	0.61	0.65	0.67	0.68	0.67	0.71	0.74	0.80	0.82	0.87	0.89	1					
16	0.40	0.43	0.57	0.61	0.62	0.67	0.69	0.70	0.70	0.71	0.75	0.79	0.80	0.84	0.86	0.91	1				
17	0.39	0.44	0.58	0.61	0.61	0.67	0.67	0.66	0.67	0.67	0.71	0.75	0.77	0.80	0.81	0.86	0.89	1			
18	0.39	0.42	0.55	0.59	0.59	0.64	0.63	0.65	0.64	0.65	0.67	0.73	0.74	0.76	0.78	0.83	0.85	0.90	1		
21	0.38	0.41	0.49	0.54	0.52	0.57	0.60	0.61	0.63	0.62	0.63	0.66	0.67	0.67	0.68	0.70	0.71	0.75	0.77	1	
24	0.32	0.36	0.49	0.49	0.48	0.52	0.55	0.58	0.56	0.58	0.59	0.62	0.64	0.65	0.66	0.65	0.67	0.69	0.69	0.80	1

Table 4.1: Matrix of correlations between Weight Z Scores at different ages, in months (Malawi)

4.2.2 Pakistan

Since growth patterns between the four areas differ, the four area's correlation matrices were modelled separately. The Pakistan dataset contains a varying number of rows for each of the 3146 children because some children had missing data at scheduled measurement dates. There were very few repeat measurements or measurements between scheduled measurement dates. If no data was recorded one month for a child, no NA value was recorded. Therefore, each row for each child therefore did not represent each scheduled measurement date so that if, for example, a child had three rows, they did not necessarily represent the first three measurements. The dataset was reformatted so that each child had the same number of rows where each row represented each scheduled measurement date - 25 rows for each of the 3146 children (78650 rows in total), with NA values for missing data.

In total, there are 625 cells in the matrix however the lower triangular section contains $\frac{25(25-1)}{2} = 300$ unique correlations. Following our rule for obtaining the MSE, we remove ¼ of the correlations in our matrices and model using the remaining correlations ($\frac{1}{4}.300 = 75$). Therefore, we will remove 75 correlations and predict these using our models (based on the remaining 225), then calculate the MSE using the residuals. Correlation matrices can be found in Tables 4.2 - 4.5, plots of these matrices can be found in Figures 4.3 – 4.6.

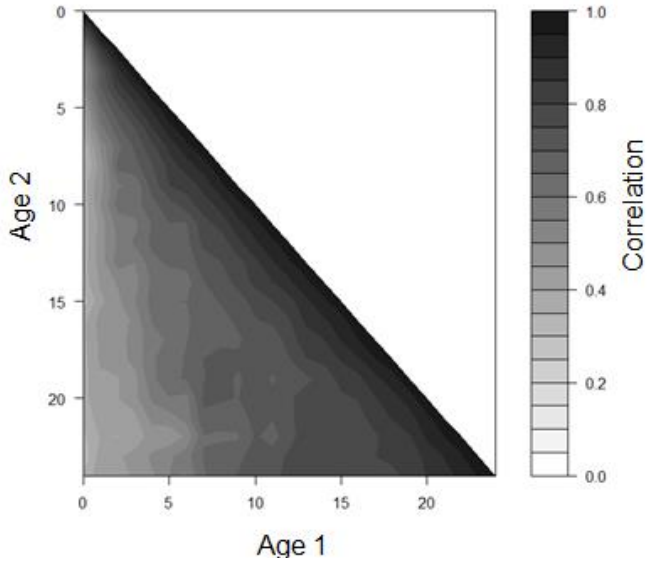

Figure 4.3: Correlations between t_1 and t_2 (village)

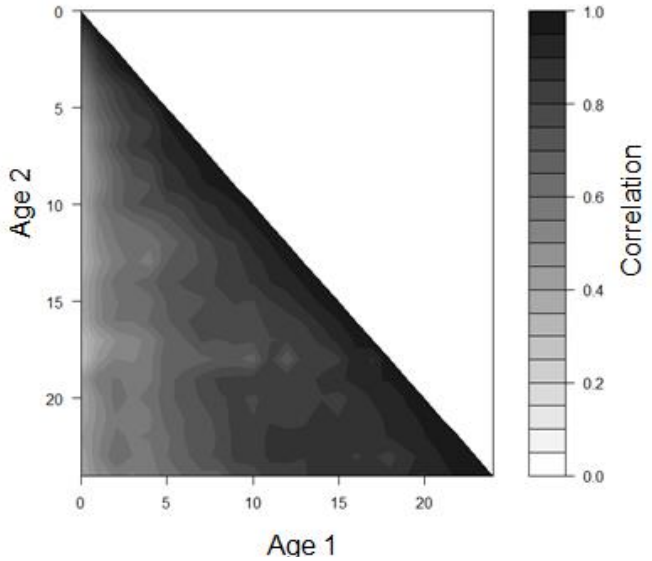

Figure 4.4: Correlations between t_1 and t_2 (peri-urban slum)

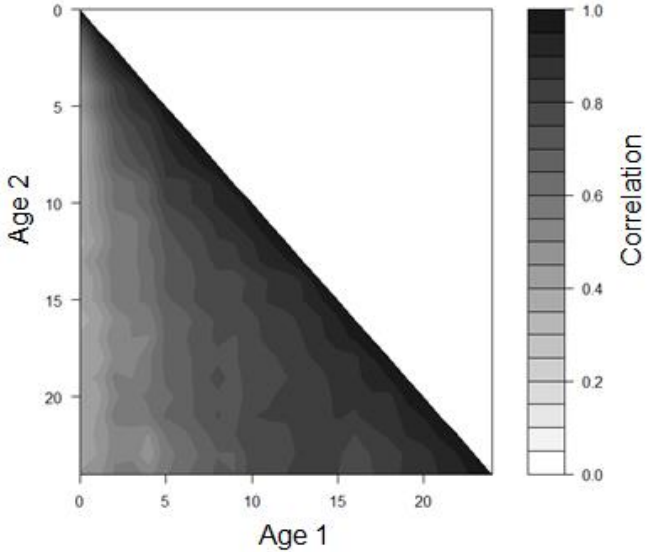

Figure 4.5: Correlations between t_1 and t_2 (urban slum)

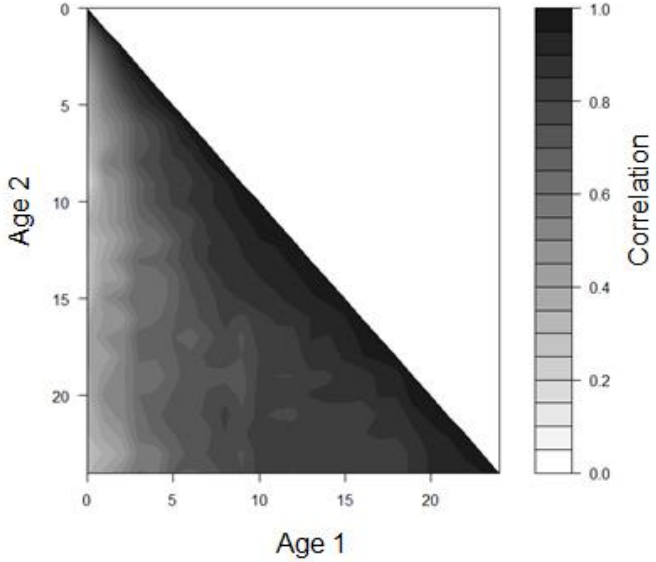

Figure 4.6: Correlations between t_1 and t_2 (middle class)

Age	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
0	1																								
1	0.77	1																							
2	0.58	0.84	1																						
3	0.49	0.72	0.89	1																					
4	0.5	0.69	0.82	0.9	1																				
5	0.44	0.64	0.74	0.83	0.91	1																			
6	0.41	0.59	0.69	0.75	0.85	0.93	1																		
7	0.35	0.52	0.63	0.72	0.79	0.87	0.93	1																	
8	0.34	0.47	0.67	0.68	0.75	0.82	0.88	0.92	1																
9	0.38	0.49	0.59	0.62	0.72	0.8	0.82	0.87	0.94	1															
10	0.36	0.5	0.63	0.63	0.68	0.73	0.78	0.82	0.87	0.93	1														
11	0.36	0.47	0.57	0.59	0.66	0.7	0.7	0.78	0.81	0.88	0.93	1													
12	0.36	0.48	0.56	0.61	0.66	0.66	0.69	0.76	0.77	0.83	0.89	0.93	1												
13	0.38	0.48	0.57	0.55	0.63	0.66	0.67	0.73	0.75	0.78	0.84	0.9	0.94	1											
14	0.38	0.46	0.51	0.53	0.61	0.61	0.63	0.7	0.72	0.77	0.81	0.85	0.89	0.94	1										
15	0.33	0.44	0.49	0.55	0.62	0.65	0.65	0.69	0.7	0.72	0.78	0.8	0.83	0.89	0.94	1									
16	0.38	0.48	0.45	0.52	0.6	0.63	0.64	0.67	0.67	0.72	0.74	0.76	0.8	0.86	0.89	0.92	1								
17	0.39	0.47	0.49	0.52	0.58	0.61	0.65	0.67	0.68	0.69	0.73	0.71	0.78	0.83	0.85	0.9	0.95	1							
18	0.38	0.46	0.47	0.52	0.59	0.63	0.66	0.7	0.73	0.71	0.74	0.72	0.75	0.8	0.82	0.85	0.89	0.95	1						
19	0.38	0.44	0.45	0.48	0.59	0.63	0.66	0.7	0.73	0.69	0.73	0.69	0.73	0.74	0.76	0.82	0.84	0.89	0.92	1					
20	0.38	0.44	0.45	0.49	0.56	0.6	0.58	0.71	0.71	0.7	0.72	0.72	0.74	0.77	0.76	0.79	0.8	0.85	0.89	0.92	1				
21	0.37	0.43	0.42	0.44	0.52	0.53	0.57	0.69	0.68	0.67	0.72	0.7	0.74	0.78	0.76	0.77	0.78	0.82	0.85	0.88	0.93	1			
22	0.34	0.4	0.4	0.43	0.47	0.47	0.51	0.65	0.64	0.64	0.71	0.68	0.73	0.78	0.77	0.78	0.76	0.79	0.83	0.85	0.91	0.95	1		
23	0.31	0.43	0.44	0.46	0.51	0.54	0.58	0.65	0.68	0.68	0.74	0.72	0.76	0.79	0.77	0.78	0.77	0.8	0.81	0.83	0.86	0.9	0.95	1	
24	0.36	0.43	0.44	0.46	0.5	0.55	0.55	0.65	0.69	0.7	0.73	0.72	0.78	0.8	0.8	0.8	0.77	0.79	0.79	0.81	0.85	0.89	0.91	0.95	1

Table 4.2: Matrix of correlations between Weight Z Scores at different ages, in months (village)

Age	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
0	1																								
1	0.81	1																							
2	0.68	0.89	1																						
3	0.59	0.81	0.93	1																					
4	0.51	0.71	0.82	0.89	1																				
5	0.49	0.68	0.79	0.86	0.9	1																			
6	0.41	0.62	0.72	0.81	0.83	0.94	1																		
7	0.39	0.61	0.75	0.83	0.87	0.92	0.95	1																	
8	0.38	0.56	0.67	0.77	0.79	0.88	0.92	0.96	1																
9	0.38	0.53	0.67	0.73	0.75	0.87	0.89	0.94	0.96	1															
10	0.37	0.58	0.67	0.74	0.77	0.79	0.82	0.88	0.91	0.95	1														
11	0.34	0.54	0.62	0.64	0.67	0.72	0.77	0.84	0.88	0.91	0.95	1													
12	0.36	0.52	0.6	0.63	0.61	0.67	0.71	0.77	0.82	0.85	0.92	0.95	1												
13	0.34	0.48	0.57	0.61	0.57	0.7	0.71	0.76	0.8	0.84	0.87	0.92	0.96	1											
14	0.39	0.53	0.63	0.65	0.69	0.73	0.76	0.78	0.81	0.83	0.86	0.9	0.92	0.96	1										
15	0.39	0.54	0.6	0.63	0.63	0.71	0.72	0.74	0.79	0.8	0.8	0.85	0.88	0.92	0.96	1									
16	0.38	0.53	0.61	0.6	0.66	0.69	0.74	0.77	0.78	0.78	0.78	0.81	0.83	0.86	0.91	0.95	1								
17	0.31	0.42	0.5	0.53	0.56	0.66	0.7	0.76	0.75	0.76	0.76	0.8	0.8	0.82	0.87	0.92	0.96	1							
18	0.29	0.47	0.55	0.54	0.59	0.7	0.67	0.68	0.71	0.71	0.67	0.82	0.7	0.84	0.81	0.84	0.94	0.89	1						
19	0.44	0.58	0.6	0.59	0.59	0.68	0.72	0.77	0.81	0.81	0.83	0.83	0.84	0.83	0.87	0.89	0.91	0.94	0.96	1					
20	0.4	0.53	0.63	0.59	0.62	0.68	0.73	0.75	0.79	0.81	0.79	0.82	0.81	0.8	0.85	0.83	0.87	0.9	0.92	0.96	1				
21	0.41	0.51	0.59	0.57	0.59	0.68	0.7	0.74	0.76	0.83	0.8	0.84	0.83	0.83	0.89	0.85	0.86	0.88	0.92	0.94	0.96	1			
22	0.45	0.54	0.61	0.6	0.6	0.64	0.7	0.73	0.75	0.82	0.82	0.86	0.86	0.87	0.89	0.87	0.87	0.88	0.9	0.92	0.95	0.97	1		
23	0.4	0.55	0.65	0.6	0.58	0.65	0.69	0.73	0.76	0.81	0.82	0.86	0.86	0.85	0.89	0.86	0.85	0.86	0.83	0.89	0.91	0.95	0.97	1	
24	0.4	0.51	0.56	0.54	0.57	0.62	0.65	0.69	0.68	0.77	0.78	0.82	0.83	0.84	0.88	0.88	0.87	0.87	0.87	0.9	0.9	0.94	0.96	0.96	1

Table 4.3: Matrix of correlations between Weight Z Scores at different ages, in months (peri-urban slum)

Age	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
0	1																								
1	0.74	1																							
2	0.59	0.86	1																						
3	0.52	0.76	0.91	1																					
4	0.42	0.64	0.81	0.88	1																				
5	0.45	0.64	0.79	0.86	0.9	1																			
6	0.37	0.57	0.73	0.79	0.84	0.93	1																		
7	0.37	0.55	0.71	0.77	0.81	0.9	0.95	1																	
8	0.4	0.54	0.68	0.73	0.78	0.86	0.91	0.95	1																
9	0.4	0.51	0.61	0.64	0.7	0.84	0.83	0.87	0.92	1															
10	0.39	0.5	0.61	0.63	0.69	0.8	0.83	0.87	0.91	0.9	1														
11	0.42	0.48	0.58	0.58	0.64	0.75	0.78	0.81	0.87	0.9	0.94	1													
12	0.38	0.46	0.56	0.56	0.66	0.72	0.76	0.79	0.85	0.85	0.9	0.95	1												
13	0.44	0.51	0.59	0.59	0.63	0.7	0.73	0.77	0.82	0.83	0.88	0.92	0.94	1											
14	0.41	0.46	0.57	0.58	0.62	0.71	0.73	0.76	0.78	0.76	0.82	0.85	0.87	0.91	1										
15	0.41	0.46	0.58	0.59	0.62	0.69	0.72	0.77	0.8	0.78	0.82	0.86	0.87	0.9	0.94	1									
16	0.38	0.42	0.51	0.56	0.57	0.67	0.68	0.72	0.76	0.78	0.78	0.82	0.84	0.86	0.9	0.94	1								
17	0.41	0.45	0.54	0.55	0.54	0.66	0.7	0.72	0.76	0.74	0.78	0.82	0.82	0.85	0.89	0.93	0.96	1							
18	0.4	0.4	0.52	0.52	0.59	0.67	0.68	0.73	0.74	0.75	0.77	0.78	0.82	0.84	0.87	0.88	0.89	0.95	1						
19	0.43	0.46	0.56	0.55	0.59	0.66	0.67	0.73	0.77	0.73	0.79	0.8	0.8	0.82	0.83	0.88	0.89	0.93	0.95	1					
20	0.4	0.42	0.54	0.58	0.61	0.63	0.67	0.72	0.74	0.72	0.78	0.79	0.82	0.82	0.83	0.86	0.87	0.89	0.88	0.96	1				
21	0.44	0.45	0.58	0.59	0.59	0.66	0.67	0.72	0.75	0.74	0.8	0.81	0.8	0.83	0.83	0.86	0.84	0.88	0.86	0.93	0.95	1			
22	0.44	0.46	0.52	0.52	0.5	0.6	0.64	0.67	0.72	0.75	0.77	0.78	0.79	0.82	0.81	0.83	0.8	0.83	0.86	0.9	0.92	0.96	1		
23	0.39	0.45	0.54	0.54	0.46	0.59	0.61	0.65	0.7	0.7	0.78	0.79	0.78	0.82	0.81	0.81	0.77	0.81	0.81	0.86	0.88	0.92	0.95	1	
24	0.46	0.48	0.56	0.58	0.53	0.6	0.64	0.68	0.71	0.7	0.77	0.8	0.8	0.81	0.81	0.8	0.77	0.8	0.84	0.86	0.86	0.9	0.93	0.96	1

Table 4.4: Matrix of correlations between Weight Z Scores at different ages, in months (urban slum)

Age	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
0	1																								
1	0.7	1																							
2	0.54	0.78	1																						
3	0.44	0.68	0.89	1																					
4	0.37	0.61	0.8	0.91	1																				
5	0.36	0.56	0.69	0.85	0.93	1																			
6	0.35	0.48	0.59	0.74	0.8	0.9	1																		
7	0.31	0.45	0.59	0.73	0.76	0.88	0.93	1																	
8	0.26	0.54	0.59	0.75	0.79	0.83	0.86	0.95	1																
9	0.19	0.45	0.52	0.67	0.76	0.82	0.84	0.88	0.94	1															
10	0.3	0.44	0.52	0.69	0.76	0.81	0.86	0.89	0.93	0.96	1														
11	0.35	0.44	0.52	0.63	0.69	0.76	0.78	0.87	0.9	0.94	0.97	1													
12	0.3	0.38	0.47	0.59	0.61	0.7	0.8	0.85	0.86	0.91	0.95	0.96	1												
13	0.29	0.43	0.54	0.65	0.67	0.74	0.78	0.86	0.88	0.89	0.93	0.94	0.96	1											
14	0.3	0.44	0.45	0.64	0.63	0.66	0.77	0.8	0.86	0.87	0.9	0.92	0.93	0.97	1										
15	0.3	0.5	0.57	0.65	0.61	0.67	0.74	0.77	0.84	0.85	0.88	0.89	0.92	0.94	0.96	1									
16	0.39	0.45	0.47	0.6	0.64	0.7	0.73	0.77	0.81	0.78	0.84	0.86	0.87	0.9	0.92	0.95	1								
17	0.4	0.5	0.52	0.67	0.69	0.7	0.67	0.72	0.79	0.74	0.83	0.84	0.83	0.91	0.9	0.93	0.96	1							
18	0.41	0.42	0.51	0.66	0.66	0.7	0.71	0.77	0.79	0.75	0.8	0.82	0.83	0.87	0.88	0.91	0.95	0.96	1						
19	0.38	0.48	0.54	0.62	0.61	0.69	0.72	0.74	0.73	0.73	0.81	0.8	0.8	0.84	0.84	0.87	0.86	0.91	0.93	1					
20	0.38	0.52	0.52	0.65	0.66	0.7	0.7	0.76	0.79	0.73	0.82	0.83	0.81	0.85	0.86	0.86	0.86	0.9	0.91	0.97	1				
21	0.34	0.47	0.52	0.63	0.69	0.73	0.73	0.76	0.81	0.76	0.81	0.79	0.79	0.82	0.83	0.83	0.81	0.86	0.85	0.93	0.96	1			
22	0.35	0.42	0.47	0.62	0.66	0.7	0.71	0.73	0.8	0.79	0.84	0.83	0.81	0.84	0.83	0.84	0.85	0.87	0.86	0.89	0.93	0.96	1		
23	0.3	0.36	0.41	0.57	0.56	0.67	0.72	0.75	0.79	0.74	0.83	0.84	0.82	0.82	0.83	0.81	0.81	0.8	0.82	0.86	0.91	0.9	0.97	1	
24	0.34	0.41	0.51	0.61	0.59	0.7	0.69	0.7	0.77	0.75	0.82	0.84	0.8	0.82	0.82	0.81	0.82	0.81	0.82	0.87	0.91	0.9	0.92	0.95	1

Table 4.5: Matrix of correlations between Weight Z Scores at different ages, in months (middle class)

4.2.3 South Africa

As with the Pakistan dataset, the South African dataset contains a varying number of rows per child, therefore a correlation matrix could not be obtained using the original data frame. A new data frame was set up to represent 15 measurements per child, with NA values present where a data point for that child was not available within 0.025 years (9.125 days) to the left or right of the scheduled date.

There are 2938 children within the data frame, each with 15 rows to represent scheduled measurement dates (44070 rows total in the new data frame).

The full correlation matrix is a $15\ x\ 15$ matrix of correlations between groups of Z scores. There are 225 cells in the matrix however the lower triangular section contains $\frac{15(15-1)}{2} = 105$ unique correlations. Following our rule for obtaining the MSE, we remove ¼ of the correlations in our matrix and model using the remaining correlations ($\frac{1}{4}.105 = 26.25$). Therefore, we will remove 26 correlations and predict these using our models (based on the remaining 79), then calculate the MSE using the residuals.

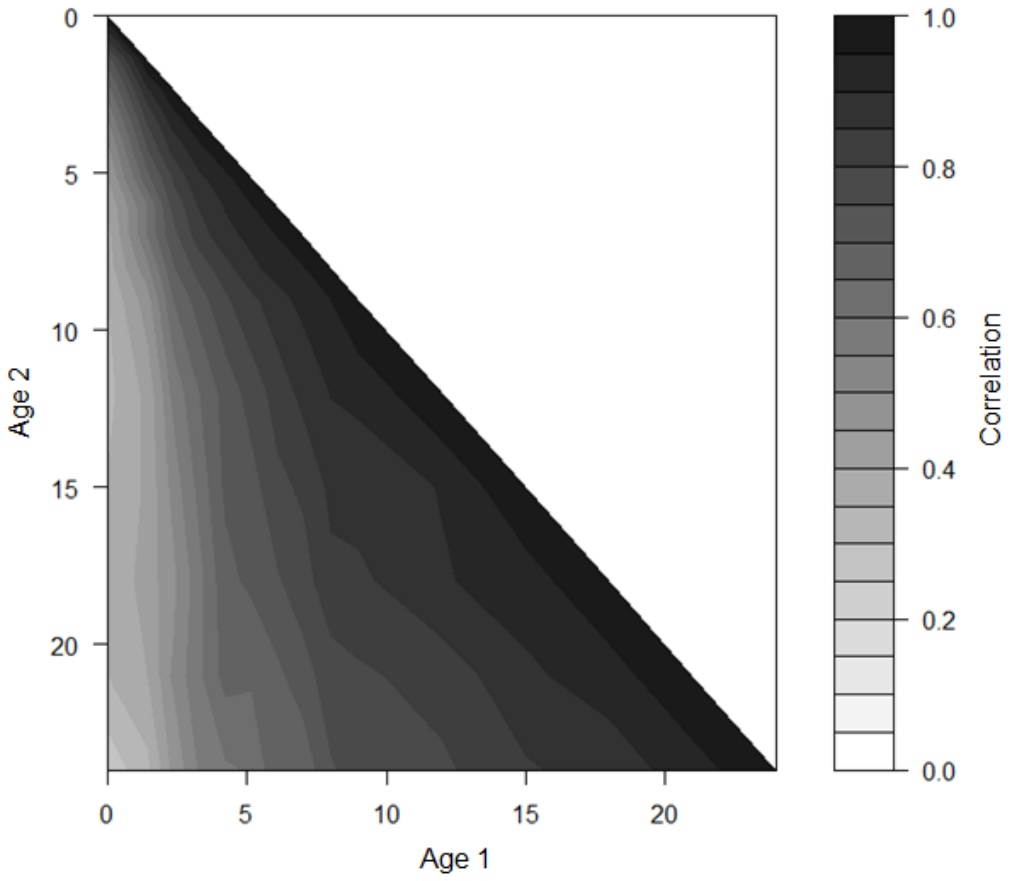

Figure 4.7: Correlations between t_1 and t_2 (South Africa)

Age	0	1.4	2.33	3.27	4.20	5.14	6.07	7	8	9	12	15	18	21	24
0	1														
1.4	0.71	1													
2.33	0.63	0.89	1												
3.27	0.55	0.81	0.92	1											
4.20	0.49	0.75	0.86	0.93	1										
5.14	0.45	0.68	0.80	0.89	0.93	1									
6.07	0.40	0.60	0.76	0.85	0.90	0.94	1								
7	0.38	0.60	0.73	0.82	0.88	0.91	0.95	1							
8	0.37	0.53	0.69	0.77	0.83	0.88	0.92	0.94	1						
9	0.34	0.47	0.64	0.72	0.79	0.83	0.88	0.91	0.95	1					
12	0.33	0.41	0.56	0.64	0.72	0.77	0.82	0.86	0.90	0.91	1				
15	0.35	0.40	0.51	0.61	0.70	0.72	0.78	0.80	0.86	0.87	0.90	1			
18	0.35	0.42	0.48	0.58	0.68	0.70	0.74	0.77	0.83	0.83	0.89	0.92	1		
21	0.35	0.39	0.50	0.59	0.66	0.65	0.68	0.72	0.77	0.79	0.81	0.88	0.92	1	
24	0.26	0.33	0.46	0.55	0.59	0.60	0.68	0.66	0.74	0.76	0.79	0.84	0.87	0.92	1

Table 4.6: Matrix of correlations between Weight Z Scores at different ages, in months (South Africa)

4.3 Model comparisons

We applied the fractional polynomial model developed by Cole (1995), and the Argyle model (Argyle et al. 2008) to our matrices as well as two others - fractional polynomials but let a selection algorithm choose the parameters (Equation (2.24)) and GAMs (Equation (2.22)). We chose fractional polynomials to model our datasets as they have been proven by Cole (1995) to provide good fit to data of this kind. Furthermore the selection algorithm has been shown to select terms which model non-traditional shapes (Nikolaeva et al. 2015).

However, Cole did not use the selection algorithm as it was not available at the time. The selection algorithm uses is a form of backward elimination. GAMs were chosen since they are a non-parametric alternative to parametric models which allow users to specify the smoothness of the models.

Model fit was assessed by four different criteria: R^2, RMS, AIC (adj.) and BIC (adj.). Since the Argyle model is on the log scale and all others are transformed by the Fisher transformation, AICs and BICs had to be adjusted by multiplying the likelihood by the Jacobian of the transformation. Refer to Section 2.9 for further description of the methods used to assess model fit.

	Model	BIC (adj)	AIC (adj)	$R^2(adj)$	Min RMS	Mean RMS	Max RMS
Malawi	GAM	-902	-986	0.974	0.020	0.024	0.028
	FP	-879	-902	0.959	0.026	0.029	0.033
	Cole	-800	-823	0.941	0.036	0.042	0.049
	Argyle	-512	-519	0.950	0.079	0.090	0.103
Village	GAM	-1163	-1257	0.968	0.026	0.030	0.032
	FP	-1372	-1402	0.979	0.019	0.025	0.030
	Cole	-1245	-1271	0.967	0.033	0.035	0.038
	Argyle	-861	-869	0.969	0.047	0.053	0.060
Peri-urban	GAM	-1093	-1185	0.950	0.027	0.035	0.044
	FP	-1140	-1170	0.944	0.029	0.038	0.047
	Cole	-1085	-1059	0.926	0.033	0.034	0.037
	Argyle	-840	-848	0.930	0.042	0.050	0.058
Urban	GAM	-1164	-1257	0.966	0.026	0.030	0.040
	FP	-1244	-1274	0.963	0.024	0.030	0.036
	Cole	-1223	-1249	0.959	0.027	0.033	0.040
	Argyle	-859	-867	0.964	0.044	0.049	0.056
Middle	GAM	-1130	-1221	0.961	0.031	0.037	0.044
	FP	-1207	-1233	0.961	0.026	0.029	0.034
	Cole	-1046	-1072	0.932	0.041	0.050	0.058
	Argyle	-726	-733	0.952	0.044	0.052	0.057
S.Africa	GAM	-442	-506	0.986	0.020	0.028	0.046
	FP	-510	-529	0.986	0.016	0.021	0.029
	Cole	-431	-449	0.971	0.032	0.051	0.076
	Argyle	-342	-349	0.981	0.027	0.035	0.045

Table 4.7: Model comparisons

Model comparisons between each of the 6 models, applied to each of the three datasets can be seen in Table 4.7 above. Values highlighted in red show the best fitting model for that criterion within that dataset.

Overall, across all criteria, GAMs and the FP models tend to perform best. However, the fractional polynomial models perform best in 4 of the 6 datasets. Furthermore, it is a parametric method and is not as computationally expensive as a GAM. The final 6 models

can be found in Table 4.8. In general, models not only have different parameters, but do not take the same form. Because these equations are different, it is hard to generalise the use of this methodology. A general model was built using the unique terms of each of the models. However, this did not perform well.

4.5 Chapter conclusions

Generalised conditional weight gain requires the correlation between groups of measurements as Z scores at the times at which growth of the child was measured. The correlation is required as it is the amount that a child's weight (or height) Z score will be expected to regress towards the mean between t_1 and t_2. Correlations can be calculated between scheduled measurements since data are available at these times. However, it must be noted that since measurements were not taken at the exact scheduled dates, the correlations may have been impacted, possibly reduced. Berkey et al. (1983) and Wright (1994) calculated correlations between exact ages and Cameron (1980) estimated correlations within age groups within their datasets for use within their conditional growth methodology. Correlations cannot be calculated in between scheduled measurement dates as no data are available. However, we can estimate correlations between measurement dates by interpolation.

Within the literature, Cole (1995) adopted fractional polynomials to model the correlation structure of the Cambridge Infant Growth study (Whitehead et al. 1989), England. This model had 6 terms (including intercept) and predicted the correlation coefficients using functions of the mean and difference between time points after applying the Fisher transformation (Fisher, 1921). Cole found that the correlation increased as the difference decreased and mean increased (Cole, 1995). Cole applied the Fisher transformation since it approximately stabilises the variance. This means the variance is approximately constant for the population correlation coefficient. Without the transformation, the variance gets smaller as the coefficient approaches 1.

Argyle et al. (2008) and Argyle (2002) used a form of linear regression to model the correlational surface of a dataset from Newcastle, England, which naturally estimated on the log scale. Argyle applied this model to the Cambridge dataset analysed by Cole (1995) but found that the function did not capture some of the observed trend within that dataset (there was the need for a quadratic term) which suggests that the approximation may not fit some datasets with any trend that deviates from the model's form. Roelants

(2013) also noted that this function is not always a satisfactory fit. However, Argyle (2002) states that the aim was simply to provide a simple model. This is of practical benefit but is not appropriate if structures are complex, thus requiring more flexible interpolation methods.

We applied both Cole and Argyle's models to our datasets, as well as a further two: FPs and GAMs. Although Cole's FP model was developed solely for the Cambridge Infant Growth study, we applied it to determine how well a model built for a developed world dataset would fit the developing world matrices. The model's coefficients were allowed to vary but we forced the model to have the same terms as specified by Cole (1995).

The $mfp()$ package in R estimates fractional polynomial terms based on a backward elimination selection algorithm. The FP models were chosen because they are parametric yet flexible, and the selection algorithm based on the "closed test procedure" (Marcus et al. 1976) which controls model selection (see Ambler & Royston (2001)) has been shown to be a useful modelling tool where non-traditional shapes cannot be modelled by logarithmic or quadratic functions (Nikolaeva et al. 2015).

The GAMs we applied had the number of basis functions set to 5 per dimension i.e. 25 in total. GAMs were chosen as a non-parametric alternative which allow the degree of smoothing to be chosen. Again, choosing the degree of smoothing is somewhat subjective although there are criteria to help chose smoothness, such as the GCV introduced by Craven & Wahba (1979). We assessed fit based on this criterion but eventually fixed the number of basis functions at 5 per dimension. This provided good fit over the correlations even though there were differences between the sizes of the matrices (Pakistani matrices: 300 correlations; Malawian: 210; South African: 105). All correlation coefficients were transformed using the Fisher transformation, for reasons given by Cole (1995), in that it is a variance stabilising transformation.

Fractional polynomials performed best overall, followed closely by GAMs. Fractional polynomials performed best in the village, urban slum, middle class area and South Africa. GAMs performed best in Malawi and the peri-urban slum.

In retrospect, using larger smoothing parameters for the GAM models would most likely provide a better fit than the FP models. However, GAM models are non-parametric and as a consequence they are harder to interpret than FPs. It is actually quite surprising that the GAM model did not perform well on the South African matrix as it is relatively small, with only a third as many data points as the Pakistani matrices and half as many as the

Malawian set. Roelants (2013) suggests that correlation between measurements may be comparable across populations despite differences in size. A general model was built by combining unique terms within the fractional polynomial models, however it did not perform very well across the datasets and was abandoned. In fact, not only were the parameters of the models different, the form of the models were also different. This highlights the need for complex modelling of correlation matrices when applied to data of this type, rather than a generalised model. As noted above, Heimendinger & Laird (1983) highlighted that correlations should be "pertinent to the target population", so it might be expected that developing a generalised model would not be possible

Coupled with interpolation from Figures 3.24 and 3.25 to obtain $\mu_{Z_1}, \mu_{Z_2}, \sigma_{Z_1}$ and σ_{Z_2}, the estimate of the correlation coefficient can be used in Equation (2.19), generalised conditional weight gain. Furthermore, these models can be used in any field in which correlations are calculated between successive time points, allowing users to estimate correlations for pairs of ages when no data is available.

Area	Model	Eq. number
Malawi	$E\left(\phi\left(r_{t_i,t_j}\right)\right) = \beta_0 + \beta_1 \log\left(\frac{\Delta t}{10}\right) + \beta_2\left(\frac{\Delta t}{10}\right)^2 + \beta_3\left(\frac{\bar{t}}{10}\right) + \beta_4\left(\frac{\bar{t}}{10}\right).\log\left(\frac{\bar{t}}{10}\right) + \beta_5\left(\Delta t.\frac{\bar{t}}{100}\right)$	(4.6)
Village	$E\left(\phi\left(r_{t_i,t_j}\right)\right) = \beta_0 + \beta_1 \log\left(\frac{\Delta t}{10}\right) + \beta_2\left(\frac{\Delta t}{10}\right)^2 + \beta_3\left(\frac{\bar{t}}{10}\right)^2 + \beta_4\left(\frac{\bar{t}}{10}\right)^3 + \beta_5 log\left(\Delta t.\frac{\bar{t}}{100}\right) + \beta_6\left(\Delta t.\frac{\bar{t}}{100}\right)^{0.5}$	(4.7)
Peri-urban slum	$E\left(\phi\left(r_{t_i,t_j}\right)\right) = \beta_0 + \beta_1 \log\left(\frac{\Delta t}{10}\right) + \beta_2\left(\frac{\Delta t}{10}\right)^2 + \beta_3 log\left(\Delta t.\frac{\bar{t}}{100}\right) + \beta_4\left(\Delta t.\frac{\bar{t}}{100}\right) + \beta_6\left(\frac{\bar{t}}{10}\right)^3 + \beta_7\left(\frac{\bar{t}}{10}\right)^3.\log\left(\frac{\bar{t}}{10}\right)$	(4.8)
Urban slum	$E\left(\phi\left(r_{t_i,t_j}\right)\right) = \beta_0 + \beta_1 \log\left(\frac{\Delta t}{10}\right) + \beta_2\left(\frac{\Delta t}{10}\right)^2 + \beta_3\left(\frac{\bar{t}}{10}\right)^{0.5} + \beta_4\left(\frac{\bar{t}}{10}\right)^{0.5}.\log\left(\frac{\bar{t}}{10}\right)^{0.5} + \beta_5\left(\Delta t.\frac{\bar{t}}{100}\right)$	(4.9)
Middle class	$E\left(\phi\left(r_{t_i,t_j}\right)\right) = \beta_0 + \beta_1 log\left(\frac{\Delta t}{10}\right) + \beta_2\left(\frac{\Delta t}{10}\right)^2 + \beta_3\left(\frac{\bar{t}}{10}\right) + \beta_4\left(\frac{\bar{t}}{10}\right).log\left(\frac{\bar{t}}{10}\right) + \beta_5\left(\Delta t.\frac{\bar{t}}{100}\right)$	(4.10)
South Africa	$E\left(\phi\left(r_{t_i,t_j}\right)\right) = \beta_0 + \beta_1\left(\frac{\Delta t}{10}\right)^{0.5} + \beta_2\left(\Delta t.\frac{\bar{t}}{100}\right) + \beta_3\left(\Delta t.\frac{\bar{t}}{100}\right).\log\left(\Delta t.\frac{\bar{t}}{100}\right) + \beta_4\left(\frac{\bar{t}}{10}\right)^{0.5} + \beta_5\left(\frac{\bar{t}}{10}\right)$	(4.11)

Table 4.8: Final FP models

Chapter 5
Exploring nutritional states
5.1 Introduction

The first 4 chapters of this thesis addressed aims which focus on data input, data summarising, data modelling, and developing generalised conditional weight gain.

The remainder of this thesis focuses on contributing to the understanding of how children grow by quantifying which measures of undernutrition are useful predictors of adverse future outcomes, testing the hypotheses developed in Section 1.3.4.

The original study question for this whole body of research was to determine what levels of generalised conditional weight gain predict later adverse outcomes, which we continue with in Chapter 6. The review of literature surrounding the original study aim highlighted that not only are size and growth predictors of adverse outcomes, nutritional status is also valuable. However, the pathways children take through nutritional states are still poorly described due to a general lack of detailed longitudinal data. Furthermore, little is known about how these presumed transitions change depending on the time frame that they are evaluated.

We therefore added an investigation of the relationships between nutritional states to our research agenda. We developed 4 hypotheses based on the review.

In this Chapter, we investigate the extent to which hypotheses 1-3 are true. The three hypotheses describe presumed transitions between 'nutritional states':

> *Hypothesis 1: those who are both wasted and stunted are at a higher risk of death than those who are wasted, who are in turn at a higher risk than those who are stunted*

Hypothesis 2: wasting tends to precede stunting

Hypothesis 3: children who are stunted are likely to stay stunted

Previous research suggested that children tend to move from certain states to others however the transitions are still poorly understood.

5.2 Method

To test our hypotheses, it was necessary to adopt an approach that: allows modelling of a discrete sample space (as we have a discrete number of nutritional states), takes the element of time into consideration and estimates the probability of moving from one state to another.

The approach we identified was a stochastic model, modelling the conditional probability of moving from one state to another, representing the probabilities of moving between states with a stochastic matrix. The statistical methodology used to generate the matrices can be found in Section 2.10.

In this case, we define states: dead (D), wasted-stunted (WS), wasted (W), stunted (S) and healthy (H). These are categorised using the definitions outlined in Section 1.2.7, including the alternative definition of wasting.

State	Description	Definition
1	Dead (D)	-
2	Wasted-Stunted (WS)	$Z\ BMI < -2\ \&\ HAZ < -2$
3	Wasted (W)	$Z\ BMI < -2\ \&\ HAZ \geq -2$
4	Stunted (S)	$Z\ BMI \geq -2\ \&\ HAZ < -2$
5	Healthy (H)	$Z\ BMI \geq -2\ \&\ HAZ \geq -2$

Table 5.1: State definitions

The probability of say, a child moving from a wasted state to a stunted state, can be represented by $P(S_t|W_{t-a}) = P_{W,S}$. There are therefore 5 probabilities of moving from one state to another, resulting in 25 probabilities in total. These can be represented in the matrix:

$$P_{Area\ t-a,t} = \begin{pmatrix} p_{D,D} & P_{D,WS} & P_{D,W} & P_{D,S} & P_{D,W} \\ p_{WS,D} & P_{WS,WS} & P_{WS,W} & P_{WS,S} & P_{WS,H} \\ p_{W,D} & P_{W,WS} & P_{W,W} & P_{W,S} & P_{W,H} \\ p_{S,D} & P_{S,WS} & P_{S,W} & P_{S,S} & P_{S,H} \\ p_{H,D} & P_{H,WS} & P_{H,W} & P_{H,S} & P_{H,H} \end{pmatrix}$$

where *Area* corresponds to the cohort. The time points t and $t - a$ can vary, as can a (so long as $t \geq a \geq 0$). State 1 (dead) is considered an 'absorbing state' - children who die cannot move state.

The probability that a child will move from one state to another is multinomial distributed. If children from different datasets follow the same multinomial distribution of moving between states, those datasets can be pooled. By utilising the GLRT developed in Section 2.10.1, we can assess whether one or more cohorts can be pooled. Pooling data is appropriate when the datasets are homogeneous with respect to the parameters of interest and by assessing whether children follow the same distribution and pooling data, we provide more power to analyses within Chapter 6, which focuses on testing hypothesis 4 - assessing whether growth over a recent time period is better than recent size in predicting mortality.

5.3 Proportion of children within each state

In this section we explore the proportion of children within each state before computing transition matrices. Figure 5.1 and Tables 5.2 – 5.8 display the proportions of each of the 6 populations (plus the pooled dataset) within each of the 5 states over months 0, 3, 6, 9, 12 and 24. It is important to note that these proportions exclude NA values, and so the tables describe the number of children within each state divided by the number of non-NA values at that time point. Therefore, the proportion of dead children (in state 1) does not necessarily need to monotonically increase, as the proportion missing within each month can vary. See Appendix A for the calculated proportions within each state including NA values.

Figure 5.1a and row 1 of Tables 5.2-5.8 show the proportions of children within state 1 (dead). With the exception of South Africa and the middle class area of Pakistan, curves are roughly the same monotonic shape but shifted, with some areas more likely to have a larger number of deaths than others.

Figure 5.1b and row 2 of Tables 5.2 – 5.8 show state 2, those who are both stunted and wasted over time. For most datasets, the prevalence of state 2 (wasted-stunted) varies slightly over time but there generally seems to be a negative linear trend. The peri-urban slum and village areas have the higher proportion of their respective populations in state 2, followed jointly by the urban slum and Malawi. The middle class area of Pakistan and

South Africa have very low proportions of their populations within state 2, starting off at around 1-3% at month 0 and dropping to 0-1% from 9 months onwards.

The proportion of those in state 3 (wasted) can be seen in Figure 5.1c and row 3 of Tables 5.2 – 5.8. There is a common trend among all 6 areas with an inverse relationship between proportion stunted and time. There is some overlap between the areas. Row 4 of 5.2 – 5.8 and Figure 5.1d illustrate the proportion of stunted children (state 4). Both the South African and middle class datasets follow a very similar trajectory with a linear downwards trend. The village and peri-urban slum follow an upward linear trend 24 months, as does the urban slum. The Malawi dataset follows its own distinct path on a log trajectory, with the proportion of children in a stunted state much higher than any of the other groups, rising from 0.13 at 0 months to 0.62 at 24 months.

The proportion considered healthy can be seen in row 5 of 5.2 – 5.8 and Figure 5.1e. Again, there is a clear trend within all areas with a decrease in the proportion and rate of change of those considered healthy as time progresses. The middle class area in Pakistan has the most children in state 5, followed by the urban slum, South Africa, and village, while the peri-urban slum and Malawi are last with very similar curves.

Figure 5.1 and Tables 5.2 – 5.8 only allow us to determine the proportion of children within each state at a given time. They are of limited value as they do not provide the probability of children moving from one state to another. What is of more value is the conditional probability that a child in state i at t_1 shifts to state j at t_2.

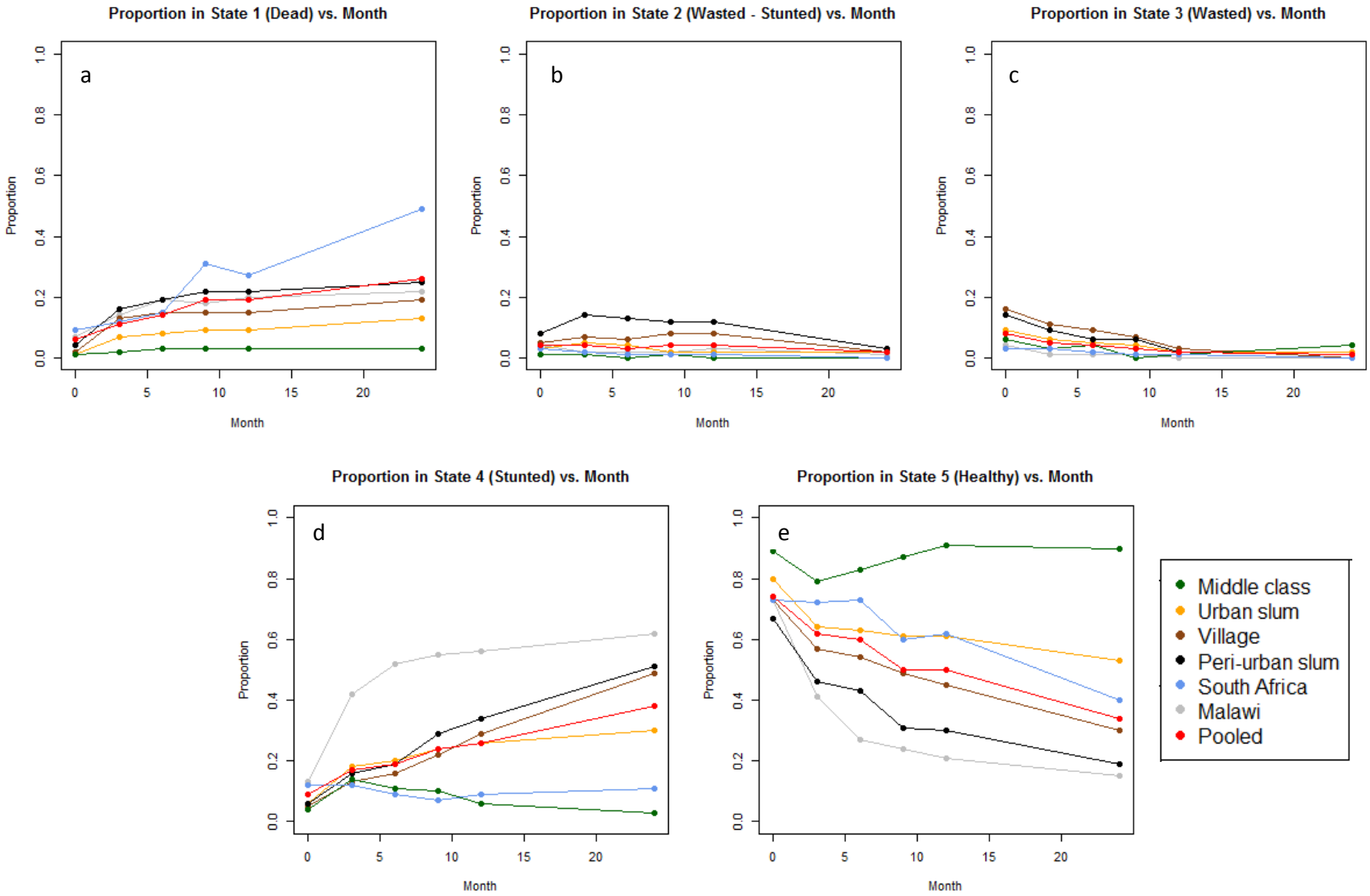

Figure 5.1: Proportion of population in states: (a) dead, (b) wasted-stunted, (c) wasted, (d) stunted, (e) healthy

Malawi	Month					
State	0	3	6	9	12	24
Dead (State 1)	0.07	0.14	0.19	0.18	0.20	0.22
Wasted and Stunted (State 2)	0.04	0.02	0.02	0.02	0.03	0.01
Wasted (State 3)	0.04	0.01	0.01	0.01	0	0
Stunted (State 4)	0.13	0.42	0.52	0.55	0.56	0.62
Healthy (State 5)	0.73	0.41	0.27	0.24	0.21	0.15

Table 5.2: Proportion of children in each state (Malawi)

Village	Month					
State	0	3	6	9	12	24
Dead (State 1)	0.02	0.13	0.15	0.15	0.15	0.19
Wasted and Stunted (State 2)	0.05	0.07	0.06	0.08	0.08	0.02
Wasted (State 3)	0.16	0.11	0.09	0.07	0.03	0
Stunted (State 4)	0.05	0.13	0.16	0.22	0.29	0.49
Healthy (State 5)	0.73	0.57	0.54	0.49	0.45	0.3

Table 5.3: Proportion of children in each state (village)

Peri-urban slum	Month					
State	0	3	6	9	12	24
Dead (State 1)	0.04	0.16	0.19	0.22	0.22	0.25
Wasted and Stunted (State 2)	0.08	0.14	0.13	0.12	0.12	0.03
Wasted (State 3)	0.15	0.09	0.06	0.06	0.02	0.01
Stunted (State 4)	0.06	0.16	0.19	0.29	0.34	0.51
Healthy (State 5)	0.67	0.46	0.43	0.31	0.30	0.19

Table 5.4: Proportion of children in each state (peri-urban slum)

Urban	Month					
State	0	3	6	9	12	24
Dead (State 1)	0.01	0.07	0.08	0.09	0.09	0.13
Wasted and Stunted (State 2)	0.03	0.05	0.04	0.02	0.02	0.02
Wasted (State 3)	0.09	0.06	0.05	0.04	0.02	0.02
Stunted (State 4)	0.06	0.18	0.20	0.24	0.26	0.30
Healthy (State 5)	0.80	0.64	0.63	0.61	0.61	0.53

Table 5.5: Proportion of children in each state (urban slum)

Middle class	Month					
State	0	3	6	9	12	24
Dead (State 1)	0.01	0.02	0.03	0.03	0.03	0.03
Wasted and Stunted (State 2)	0.01	0.01	0	0.01	0	0
Wasted (State 3)	0.06	0.03	0.04	0	0.01	0.04
Stunted (State 4)	0.04	0.14	0.11	0.1	0.06	0.03
Healthy (State 5)	0.89	0.79	0.83	0.87	0.91	0.90

Table 5.6: Proportion of children in each state (middle class)

South Africa	Month					
State	0	3	6	9	12	24
Dead (State 1)	0.09	0.12	0.15	0.31	0.27	0.49
Wasted and Stunted (State 2)	0.03	0.02	0.01	0.01	0.01	0
Wasted (State 3)	0.03	0.03	0.02	0.01	0.01	0
Stunted (State 4)	0.12	0.12	0.09	0.07	0.09	0.11
Healthy (State 5)	0.73	0.72	0.73	0.60	0.62	0.40

Table 5.7: Proportion of children in each state (South Africa)

Pooled	Month					
State	0	3	6	9	12	24
Dead (State1)	0.06	0.11	0.14	0.19	0.20	0.26
Wasted-stunted (State 2)	0.04	0.04	0.03	0.04	0.04	0.02
Wasted (State 3)	0.08	0.05	0.04	0.03	0.02	0.01
Stunted (State 4)	0.09	0.17	0.19	0.24	0.26	0.38
Healthy (State 5)	0.74	0.62	0.60	0.50	0.50	0.34

Table 5.8: Proportion of children in each state

5.4 Estimating transition matrices

Matrices on pages 133-135 show transition probabilities for all areas. As state 1 is an absorbing state, the probability of moving from this state to another is always 0. Probability calculations for some of the middle class transition matrices have frequencies of 0, returning NA probabilities, namely: from state 2 in the 6-9 month matrix and from states 2, 3 and 5 in the 9-12 month matrix. Generally frequencies are low when calculating transitions for the middle class area, which makes these probabilities imprecise.

As there are a large number of proportions, it is hard to determine whether two or more matrices differ significantly by visually interpreting the plots. By using the GLRT derived in Section 2.9.1, we can determine whether two matrices of transitions are significantly different from one another.

$$P_{Malawi\ 3,6} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.15 & 0.46 & 0 & 0.38 & 0 \\ 0.14 & 0 & 0 & 0.57 & 0.29 \\ 0.05 & 0.02 & 0.01 & 0.82 & 0.11 \\ 0.02 & 0 & 0.01 & 0.43 & 0.53 \end{pmatrix} \quad P_{Middle\ 3,6} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.67 & 0.33 \\ 0 & 0 & 0.25 & 0 & 0.75 \\ 0 & 0 & 0 & 0.50 & 0.50 \\ 0 & 0 & 0.03 & 0.03 & 0.94 \end{pmatrix} \quad P_{Village\ 3,6} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.14 & 0.35 & 0.05 & 0.40 & 0.07 \\ 0.07 & 0.11 & 0.37 & 0.19 & 0.26 \\ 0.05 & 0.03 & 0.06 & 0.55 & 0.31 \\ 0.01 & 0.03 & 0.10 & 0.08 & 0.79 \end{pmatrix}$$

$$P_{Malawi\ 6,9} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.10 & 0.20 & 0 & 0.70 & 0 \\ 0 & 0 & 0 & 0.75 & 0.25 \\ 0.02 & 0.03 & 0.01 & 0.77 & 0.16 \\ 0.03 & 0 & 0.01 & 0.33 & 0.64 \end{pmatrix} \quad P_{Middle\ 6,9} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ - & - & - & - & - \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0.08 & 0 & 0.54 & 0.38 \\ 0 & 0 & 0 & 0.07 & 0.93 \end{pmatrix} \quad P_{Village\ 6,9} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.08 & 0.51 & 0 & 0.38 & 0.03 \\ 0 & 0.13 & 0.38 & 0.15 & 0.33 \\ 0.01 & 0.09 & 0.01 & 0.67 & 0.22 \\ 0.01 & 0.03 & 0.06 & 0.13 & 0.77 \end{pmatrix}$$

$$P_{Malawi\ 9,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.14 & 0 & 0.86 & 0 \\ 0 & 0.33 & 0.17 & 0.17 & 0.33 \\ 0.01 & 0.02 & 0 & 0.84 & 0.12 \\ 0.01 & 0.02 & 0 & 0.39 & 0.59 \end{pmatrix} \quad P_{Middle\ 9,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ - & - & - & - & - \\ - & - & - & - & - \\ 0 & 0 & 0.08 & 0.58 & 0.33 \\ 0 & 0 & 0 & 0.01 & 0.99 \end{pmatrix} \quad P_{Village\ 9,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.02 & 0.52 & 0.02 & 0.43 & 0 \\ 0.05 & 0.10 & 0.26 & 0.18 & 0.41 \\ 0.01 & 0.06 & 0 & 0.80 & 0.12 \\ 0.01 & 0.01 & 0.03 & 0.14 & 0.81 \end{pmatrix}$$

$$P_{Malawi\ 0,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.41 & 0.07 & 0 & 0.48 & 0.04 \\ 0.17 & 0.09 & 0 & 0.65 & 0.09 \\ 0.13 & 0.05 & 0 & 0.71 & 0.11 \\ 0.12 & 0.02 & 0 & 0.59 & 0.27 \end{pmatrix} \quad P_{Middle\ 0,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0.18 & 0.82 \\ 0 & 0 & 0 & 0 & 1 \\ 0.01 & 0 & 0.01 & 0.06 & 0.93 \end{pmatrix} \quad P_{Village\ 0,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.46 & 0.18 & 0 & 0.26 & 0.10 \\ 0.15 & 0.15 & 0.05 & 0.27 & 0.39 \\ 0.14 & 0.14 & 0 & 0.51 & 0.22 \\ 0.09 & 0.06 & 0.03 & 0.30 & 0.52 \end{pmatrix}$$

$$P_{Malawi\ 12,24} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.07 & 0 & 0.93 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0.04 & 0.02 & 0 & 0.85 & 0.09 \\ 0.05 & 0.01 & 0.01 & 0.49 & 0.45 \end{pmatrix} \quad P_{Middle\ 12,24} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ - & - & - & - & - \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0.25 & 0.75 \\ 0 & 0 & 0.05 & 0.02 & 0.93 \end{pmatrix} \quad P_{Village\ 12,24} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.24 & 0 & 0.76 & 0 \\ 0.05 & 0 & 0 & 0.58 & 0.37 \\ 0.02 & 0.03 & 0 & 0.89 & 0.06 \\ 0.02 & 0 & 0 & 0.35 & 0.62 \end{pmatrix}$$

$$P_{Urban\ 3,6} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.17 & 0.42 & 0 & 0.42 & 0 \\ 0.03 & 0.12 & 0.25 & 0.16 & 0.44 \\ 0.01 & 0.07 & 0.03 & 0.61 & 0.28 \\ 0 & 0 & 0.04 & 0.10 & 0.86 \end{pmatrix} \quad P_{Peri\ 3,6} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.10 & 0.48 & 0 & 0.40 & 0.02 \\ 0.09 & 0.27 & 0.27 & 0.15 & 0.23 \\ 0.04 & 0.26 & 0.02 & 0.53 & 0.15 \\ 0.02 & 0.02 & 0.06 & 0.10 & 0.81 \end{pmatrix} \quad P_{S.Africa\ 3,6} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.16 & 0.23 & 0.03 & 0.48 & 0.10 \\ 0.07 & 0.02 & 0.11 & 0.17 & 0.63 \\ 0.04 & 0.06 & 0.02 & 0.38 & 0.51 \\ 0.02 & 0 & 0.02 & 0.05 & 0.92 \end{pmatrix}$$

$$P_{Urban\ 6,9} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.05 & 0.35 & 0.15 & 0.35 & 0.10 \\ 0 & 0.04 & 0.48 & 0.17 & 0.30 \\ 0.01 & 0 & 0 & 0.76 & 0.23 \\ 0 & 0.01 & 0.02 & 0.09 & 0.89 \end{pmatrix} \quad P_{Peri\ 6,9} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.05 & 0.62 & 0.03 & 0.31 & 0 \\ 0 & 0.07 & 0.36 & 0.15 & 0.43 \\ 0 & 0.12 & 0 & 0.83 & 0.06 \\ 0 & 0.01 & 0.06 & 0.26 & 0.66 \end{pmatrix} \quad P_{S.Africa\ 6,9} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.38 & 0.12 & 0 & 0.50 & 0 \\ 0 & 0 & 0.20 & 0.10 & 0.70 \\ 0.03 & 0.03 & 0.04 & 0.41 & 0.49 \\ 0.03 & 0 & 0.01 & 0.06 & 0.90 \end{pmatrix}$$

$$P_{Urban\ 9,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.40 & 0 & 0.50 & 0.10 \\ 0 & 0 & 0.32 & 0.05 & 0.64 \\ 0 & 0.02 & 0.01 & 0.82 & 0.16 \\ 0 & 0 & 0 & 0.07 & 0.93 \end{pmatrix} \quad P_{Peri\ 9,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.03 & 0.57 & 0.03 & 0.37 & 0 \\ 0 & 0.43 & 0.14 & 0.14 & 0.29 \\ 0 & 0.13 & 0 & 0.81 & 0.05 \\ 0 & 0.01 & 0.04 & 0.15 & 0.79 \end{pmatrix} \quad P_{S.Africa\ 9,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.86 & 0.14 & 0 & 0 \\ 0 & 0.09 & 0.91 & 0 & 0 \\ 0 & 0 & 0 & 0.63 & 0.37 \\ 0 & 0 & 0 & 0.07 & 0.93 \end{pmatrix}$$

$$P_{Urban\ 0,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.37 & 0.05 & 0 & 0.42 & 0.16 \\ 0.22 & 0.05 & 0.03 & 0.17 & 0.53 \\ 0.03 & 0.03 & 0 & 0.66 & 0.23 \\ 0.04 & 0.01 & 0.02 & 0.25 & 0.68 \end{pmatrix} P_{Peri\ 0,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.33 & 0.15 & 0 & 0.37 & 0.15 \\ 0.16 & 0.24 & 0 & 0.41 & 0.20 \\ 0.21 & 0.14 & 0 & 0.57 & 0.07 \\ 0.12 & 0.10 & 0.03 & 0.34 & 0.40 \end{pmatrix} P_{S.Africa\ 0,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.37 & 0.07 & 0 & 0.23 & 0.33 \\ 0.33 & 0 & 0 & 0.10 & 0.57 \\ 0.16 & 0.03 & 0 & 0.33 & 0.49 \\ 0.09 & 0.01 & 0.02 & 0.07 & 0.81 \end{pmatrix}$$

$$P_{Urban\ 12,24} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.12 & 0.12 & 0 & 0.62 & 0.12 \\ 0 & 0.50 & 0 & 0.17 & 0.33 \\ 0.01 & 0.03 & 0.02 & 0.72 & 0.23 \\ 0 & 0 & 0.03 & 0.14 & 0.83 \end{pmatrix} P_{Peri\ 12,24} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.04 & 0.25 & 0 & 0.71 & 0 \\ 0 & 0 & 0.33 & 0.67 & 0 \\ 0.01 & 0 & 0 & 0.93 & 0.06 \\ 0 & 0.02 & 0.02 & 0.36 & 0.60 \end{pmatrix} P_{S.Africa\ 12,24} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.57 & 0.14 & 0 & 0.14 & 0.14 \\ 0 & 0 & 0 & 0 & 1 \\ 0.08 & 0 & 0 & 0.62 & 0.30 \\ 0.02 & 0 & 0.01 & 0.11 & 0.86 \end{pmatrix}$$

$$P_{Pooled\ 3,6} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.14 & 0.37 & 0.02 & 0.42 & 0.05 \\ 0.07 & 0.10 & 0.24 & 0.18 & 0.40 \\ 0.04 & 0.06 & 0.02 & 0.59 & 0.29 \\ 0.01 & 0.01 & 0.03 & 0.10 & 0.85 \end{pmatrix}$$

$$P_{Pooled\ 6,9} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.09 & 0.46 & 0.04 & 0.39 & 0.03 \\ 0 & 0.08 & 0.36 & 0.17 & 0.39 \\ 0.02 & 0.04 & 0.01 & 0.72 & 0.21 \\ 0.02 & 0.01 & 0.02 & 0.12 & 0.84 \end{pmatrix}$$

$$P_{Pooled\ 9,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.02 & 0.50 & 0.03 & 0.45 & 0.01 \\ 0.02 & 0.14 & 0.33 & 0.12 & 0.39 \\ 0.01 & 0.04 & 0 & 0.80 & 0.15 \\ 0 & 0.01 & 0.01 & 0.11 & 0.87 \end{pmatrix}$$

$$P_{Pooled\ 0,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.39 & 0.11 & 0 & 0.34 & 0.16 \\ 0.18 & 0.12 & 0.03 & 0.28 & 0.39 \\ 0.14 & 0.05 & 0 & 0.51 & 0.31 \\ 0.08 & 0.03 & 0.02 & 0.27 & 0.60 \end{pmatrix}$$

$$P_{Pooled\ 12,24} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.07 & 0.20 & 0 & 0.71 & 0.02 \\ 0.03 & 0.09 & 0.03 & 0.46 & 0.40 \\ 0.03 & 0.02 & 0 & 0.83 & 0.11 \\ 0.02 & 0 & 0.02 & 0.23 & 0.73 \end{pmatrix}$$

5.5 Testing the equality of transition matrices

The transition matrices within the previous section describe how children transition from one state to another through time for each individual dataset. However, if children from different datasets share the same transition probabilities this suggests that children are moving between states in the same way. These datasets can therefore be pooled, providing more statistical power to our analyses while using the transition matrices. Furthermore, these pooled datasets can be used within further analysis, as long as the datasets are homogenous with respect to the parameters of interest for that specific analysis.

The statistical test used here was derived in Section 2.9.1. It must be noted that no adjustment has been made for multiple comparisons, although p-values which were less than 0.01 were small, indicating that our results would not have changed had a multiple comparisons procedure been used.

Results show that the peri-urban slum and the village data can be pooled as their matrices are not significantly different at all three transition times. The South African and middle class area can also be pooled. The four final transition matrices can be seen on pages 138 and 139.

Figure 5.2: Test of equality: (a) 3-6m, (b) 6-9m, (c) 9-12m

$$P_{SA/MC\ 3,6} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.15 & 0.20 & 0.03 & 0.50 & 0.12 \\ 0.06 & 0.02 & 0.12 & 0.16 & 0.64 \\ 0.03 & 0.05 & 0.02 & 0.39 & 0.51 \\ 0.02 & 0.01 & 0.02 & 0.05 & 0.90 \end{pmatrix}$$

$$P_{P/V\ 3,6} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.11 & 0.40 & 0.02 & 0.40 & 0.05 \\ 0.08 & 0.15 & 0.34 & 0.18 & 0.25 \\ 0.05 & 0.12 & 0.05 & 0.54 & 0.24 \\ 0.01 & 0.02 & 0.08 & 0.09 & 0.80 \end{pmatrix}$$

$$P_{SA/MC\ 6,9} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.37 & 0.13 & 0 & 0.5 & 0 \\ 0 & 0 & 0.15 & 0.07 & 0.78 \\ 0.02 & 0.04 & 0.04 & 0.43 & 0.47 \\ 0.03 & 0 & 0.01 & 0.06 & 0.90 \end{pmatrix}$$

$$P_{P/V\ 6,9} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.07 & 0.57 & 0.01 & 0.34 & 0.01 \\ 0 & 0.12 & 0.38 & 0.15 & 0.35 \\ 0.01 & 0.10 & 0.01 & 0.72 & 0.16 \\ 0.00 & 0.04 & 0.06 & 0.16 & 0.74 \end{pmatrix}$$

$$P_{SA/MC\ 9,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.86 & 0.14 & 0 & 0 \\ 0 & 0.10 & 0.90 & 0 & 0 \\ 0 & 0 & 0.01 & 0.63 & 0.36 \\ 0 & 0 & 0 & 0.07 & 0.93 \end{pmatrix}$$

$$P_{P/V\ 9,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.03 & 0.54 & 0.03 & 0.40 & 0.00 \\ 0.04 & 0.18 & 0.23 & 0.17 & 0.38 \\ 0.01 & 0.09 & 0 & 0.80 & 0.10 \\ 0 & 0.01 & 0.03 & 0.14 & 0.81 \end{pmatrix}$$

$$P_{SA/MC\ 0,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.35 & 0.08 & 0 & 0.22 & 0.35 \\ 0.24 & 0 & 0 & 0.12 & 0.64 \\ 0.15 & 0.03 & 0 & 0.31 & 0.51 \\ 0.08 & 0.00 & 0.02 & 0.07 & 0.83 \end{pmatrix}$$

$$P_{P/V\ 0,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.41 & 0.17 & 0 & 0.30 & 0.12 \\ 0.15 & 0.17 & 0.04 & 0.31 & 0.33 \\ 0.16 & 0.14 & 0 & 0.53 & 0.17 \\ 0.10 & 0.07 & 0.03 & 0.31 & 0.49 \end{pmatrix}$$

$$P_{SA/MC\ 12,24} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.58 & 0.14 & 0 & 0.14 & 0.14 \\ 0 & 0 & 0 & 0 & 1 \\ 0.07 & 0 & 0 & 0.59 & 0.34 \\ 0.02 & 0 & 0.02 & 0.08 & 0.88 \end{pmatrix}$$

$$P_{P/V\ 12,24} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.01 & 0.25 & 0 & 0.74 & 0 \\ 0.05 & 0 & 0.05 & 0.59 & 0.31 \\ 0.02 & 0.02 & 0 & 0.90 & 0.06 \\ 0.02 & 0.01 & 0.01 & 0.35 & 0.61 \end{pmatrix}$$

$$P_{Malawi\ 3,6} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.15 & 0.46 & 0 & 0.38 & 0 \\ 0.14 & 0 & 0 & 0.57 & 0.29 \\ 0.05 & 0.02 & 0.01 & 0.82 & 0.11 \\ 0.02 & 0 & 0.01 & 0.43 & 0.53 \end{pmatrix}$$

$$P_{Malawi\ 6,9} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.10 & 0.20 & 0 & 0.70 & 0 \\ 0 & 0 & 0 & 0.75 & 0.25 \\ 0.02 & 0.03 & 0.01 & 0.77 & 0.16 \\ 0.03 & 0 & 0.01 & 0.33 & 0.64 \end{pmatrix}$$

$$P_{Malawi\ 9,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.14 & 0 & 0.86 & 0 \\ 0 & 0.33 & 0.17 & 0.17 & 0.33 \\ 0.01 & 0.02 & 0 & 0.84 & 0.12 \\ 0.01 & 0.02 & 0 & 0.39 & 0.59 \end{pmatrix}$$

$$P_{Malawi\ 0,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.41 & 0.07 & 0 & 0.48 & 0.04 \\ 0.17 & 0.09 & 0 & 0.65 & 0.09 \\ 0.13 & 0.05 & 0 & 0.71 & 0.11 \\ 0.12 & 0.02 & 0 & 0.59 & 0.27 \end{pmatrix}$$

$$P_{Malawi\ 12,24} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.07 & 0 & 0.93 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0.04 & 0.02 & 0 & 0.85 & 0.09 \\ 0.05 & 0.01 & 0.01 & 0.49 & 0.45 \end{pmatrix}$$

$$P_{Urban\ 3,6} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.17 & 0.42 & 0 & 0.42 & 0 \\ 0.03 & 0.12 & 0.25 & 0.16 & 0.44 \\ 0.01 & 0.07 & 0.03 & 0.61 & 0.28 \\ 0 & 0 & 0.04 & 0.10 & 0.86 \end{pmatrix}$$

$$P_{Urban\ 6,9} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.05 & 0.35 & 0.15 & 0.35 & 0.10 \\ 0 & 0.04 & 0.48 & 0.17 & 0.30 \\ 0.01 & 0 & 0 & 0.76 & 0.23 \\ 0 & 0.01 & 0.02 & 0.09 & 0.89 \end{pmatrix}$$

$$P_{Urban\ 9,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.40 & 0 & 0.50 & 0.10 \\ 0 & 0 & 0.32 & 0.05 & 0.64 \\ 0 & 0.02 & 0.01 & 0.82 & 0.16 \\ 0 & 0 & 0 & 0.07 & 0.93 \end{pmatrix}$$

$$P_{Urban\ 0,12} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.37 & 0.05 & 0 & 0.42 & 0.16 \\ 0.22 & 0.05 & 0.03 & 0.17 & 0.53 \\ 0.03 & 0.03 & 0 & 0.66 & 0.23 \\ 0.04 & 0.01 & 0.02 & 0.25 & 0.68 \end{pmatrix}$$

$$P_{Urban\ 12,24} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.12 & 0.12 & 0 & 0.62 & 0.12 \\ 0 & 0.50 & 0 & 0.17 & 0.33 \\ 0.01 & 0.03 & 0.02 & 0.72 & 0.23 \\ 0 & 0 & 0.03 & 0.14 & 0.83 \end{pmatrix}$$

5.6 Transitions of interest

	1: Dead	2: Wasted-stunted	3 :Wasted	4: Stunted	5:Healthy
1:Dead	Still dead	Still dead	Still dead	Still dead	Still dead
2:Wasted-stunted	Proportion of those wasted-stunted at baseline (P_{ws}) who then died	P_{ws} now still wasted-stunted	P_{ws} now only wasted	P_{ws} still stunted but no longer wasted	P_{ws} now healthy
3:Wasted	Proportion of those wasted at baseline (P_w) who then died	P_w now also stunted	P_w still wasted	P_w stunted but no longer wasted	P_w now healthy
4:Stunted	Proportion of those stunted at baseline (P_s) who then died	P_s now also wasted	P_s now wasted but not stunted	P_s still only stunted	P_s now healthy
5:Healthy	Proportion of those healthy at baseline (P_H) who then died	P_H now wasted-stunted	P_H now wasted	P_H now stunted, without having been wasted	P_H still healthy

Table 5.9: Transitions of interest (rows are t_1, columns are t_2)

Table 5.9 summarises each transition in more detail. The transition matrices will be used to determine if the hypotheses, listed below, are true, by interpreting the probabilities of moving states and calculating relative risks:

Hypothesis 1: those who are both wasted and stunted are at a higher risk of death than those who are wasted, who are in turn at a higher risk than those who are stunted

Column 1 is the proportion moving from each of the nutritional states to a dead state. Previous research suggests that those who are wasted are at a higher risk of death than those who are stunted, and those who are wasted-stunted are at an even higher risk than both of these states. We therefore expect to observe in descending order, the probability of death: wasted-stunted, wasted, stunted, healthy.

Hypothesis 2: wasting tends to precede stunting

Row 3 is the proportion moving from a wasted state. Previous research suggests that wasting precedes stunting. We should therefore observe children moving from wasted to stunted.

Hypothesis 3: children who become stunted are likely to stay stunted

Row 4 is the proportion of children moving from a stunted state. Past research suggests stunting is largely irreversible. Children should therefore tend to stay in a stunted state. These hypotheses will be compared with our findings within Section 5.7. The purpose of the section is to provide a general overview of the pathways children take through nutritional states by describing the probabilities of moving between them. Relative risks and statistical significance is discussed in more detail within the conclusions section. Note that RRs are calculated using the exact probability of moving between states, not the rounded probabilities shown in the transition matrices. Inference is largely made on the pooled dataset and should therefore be treated as a generalised result. Further inference on the individual sets can be made by referring to the individual probability transition matrices in Section 5.4 and Section 5.5. Additional analysis on the pooled South African/middle class, village/peri-urban slum, urban slum and Malawian sets can be found in Appendix B. Note however, that many RRs could not be calculated due to denominators with values of 0. The methodology used to compute probabilities and RRs can be found in Section 2.10.

5.7 Relationships between wasting, stunting and death

5.7.2 Moving into a dead state

5.7.2.1 Transition probabilities

Hypothesis 1: those who are both wasted and stunted are at a higher risk of death than those who are wasted, who are in turn at a higher risk than those who are stunted

On the whole, our findings agree with this hypothesis. Within all three timeframes, children who are wasted-stunted are much more likely to die than those who are wasted. In the 3-6m and 9-12m timeframes, those who are wasted are more likely to die than those who are stunted. Only in the 6-9m timeframe are stunted children more likely to die than wasted children.

Timeframe	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.14	0.07	0.04	0.01
6-9m	0.09	NA	0.02	0.02
9-12m	0.02	0.02	0.01	0

Table 5.10: Pooled probability of moving from all states to dead

5.7.2.2 Relative risks

Wasted-stunted children are significantly more likely to die than healthy children in all three timeframes. Wasted children are significantly more likely to die than healthy children in the 3-6m and 9-12m timeframes, as are those who are stunted. Note that the probability of death decreases with age but relative to the healthy children, children are less vulnerable within the middle period.

Timeframe	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	11.1 (6.4, 19.1)	5.3 (2.6, 10.5)	3.0 (1.7, 5.2)	1
6-9m	5.4 (2.6, 11)	NA	1.0 (0.5, 2.1)	1
9-12m	10.0 (1.7, 59.7)	11.5 (1.9, 68.1)	4.8 (1.2, 18.5)	1

Table 5.11: RR (95% CI) of moving from each state to dead (baseline healthy)

5.7.3 Moving from a wasted-stunted state

5.7.3.1 Transition probabilities

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.14	0.37	0.02	0.42	0.05
6-9m	0.09	0.46	0.04	0.39	0.03
9-12m	0.02	0.50	0.03	0.45	0.01

Table 5.12: Pooled probability of moving from wasted-stunted to each state

Children who were wasted-stunted are most likely to revert to a stunted (but not wasted) state in the 3-6m timeframe, or stay wasted-stunted (42% to stunted compared with 37% to wasted-stunted) but are most likely to stay in a wasted-stunted state in the 6-9m and 9-12m timeframe. However, large proportions of children die within both the 3-6m timeframe (14%) and 6-9m timeframe (9%). Only 2% of children die at 12m if they were wasted-stunted at 9m.

5.7.3.2 Relative risks

Those who are wasted-stunted are significantly more likely to die than healthy children and are significantly more likely to become stunted in all three timeframes. Furthermore, wasted-stunted children are significantly more likely to stay in this state, or move to a stunted state in all three timeframes, compared with healthy children. These children are also more likely to become wasted in the 9-12m timeframe compared with healthy children.

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	11.1(6.4, 19)	64.8 (36.3, 115.8)	0.6 (0.2, 1.9)	4.2 (3.3, 5.2)	0.06 (0.03, 0.13)
6-9m	5.4 (2.6, 10)	49.9 (29.1, 85.7)	1.5 (0.5, 4.3)	3.3 (2.5, 4.3)	0.03 (0.01, 0.09)
9-12m	10.1 (1.7, 59)	87.4 (44.3, 172.5)	4.1 (1.1, 14.5)	3.9 (3.0, 5.1)	0.01 (0.00, 0.07)

Table 5.13: RR (95% CI) of moving from wasted-stunted to other states (baseline healthy)

5.7.4 Moving from a wasted state

5.7.4.1 Transition probabilities

Hypothesis 2: wasting tends to precede stunting

40% of children move from a wasted state to a healthy state in all three timeframes. The second most likely final state is wasted, with about one-third of children staying in this state. About one-sixth of wasted children become stunted.

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.07	0.10	0.24	0.18	0.40
6-9m	0	0.08	0.36	0.17	0.39
9-12m	0.02	0.14	0.33	0.12	0.39

Table 5.14: Pooled probability of moving from wasted to each state

5.7.4.2 Relative risks

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	5.3 (2.6, 10.5)	17.7 (8.7, 35.8)	7.6 (5.3, 10.7)	1.8 (1.3, 2.5)	0.47 (0.39, 0.57)
6-9m	NA	9.0 (4.0, 20.1)	15.8 (10.5, 23.9)	1.4 (0.9, 2.2)	0.46 (0.37, 0.59)
9-12m	11.5 (1.9, 68.0)	24.9 (10.9, 56.8)	47.1 (24.3, 90.9)	1.0 (0.5, 1.8)	0.44 (0.34, 0.57)

Table 5.15: RR (95% CI) of moving from wasted to other states (baseline healthy)

In comparison with healthy children, wasted children are significantly more likely to become stunted in period 3-6m but not in periods 6-9m or 9-12m. They are also

significantly more likely to become wasted-stunted or stay wasted within all three timeframes. They are also significantly more likely to die within the 3-6m and 9-12m timeframes.

5.7.5 Moving from a stunted state

5.7.5.1 Transition probabilities

Hypothesis: children who become stunted are likely to stay stunted

Past research suggested that stunting is largely irreversible as children who become stunted will not recover. Our findings partially agree with this hypothesis. The vast majority of children who are stunted at baseline stay stunted, and the probability of remaining in this state increases with age – from 59% of children during 3-6m to 80% during 9-12m.

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.04	0.06	0.02	0.59	0.29
6-9m	0.02	0.04	0.01	0.72	0.21
9-12m	0.01	0.04	0	0.80	0.15

Table 5.16: Pooled probability of moving from stunted to each state

The probability of moving to a healthy state falls with age, with 29% of children for 3-6m decreasing to 21% for 6-9m and 15% for 9-12m.

5.7.5.2 Relative risks

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	3.0 (1.7, 5.2)	9.8 (5.2, 18.5)	0.68 (0.3, 1.2)	5.9 (5.1, 6.8)	0.33 (0.30, 0.38)
6-9m	1.0 (0.5, 2.1)	4.6 (2.5, 8.6)	0.48 (0.2, 1.0)	6.2 (5.3, 7.1)	0.25 (0.21, 0.29)
9-12m	4.8 (1.2, 18.5)	6.4 (3.0, 13.5)	0.56 (0.1, 2.0)	7.1 (6.1, 8.1)	0.17 (0.14, 0.20)

Table 5.17: RR (95% CI) of moving from stunted to other states (baseline healthy)

Stunted children are significantly more likely to be stunted at the final time than healthy children. Stunted children are also significantly more likely to become wasted-stunted in all three timeframes, and significantly more likely to die in timeframes 3-6m and 9-12m.

5.7.6 Moving from a healthy state

5.7.6.1 Transition probabilities

As past research has suggested that wasting precedes stunting, we expect that children will most likely move from a healthy state to a wasted state. Our findings do not agree with this hypothesis. The majority of children who are healthy will stay in a healthy state (around 85% in all timeframes). The second most likely state that children will move into is a stunted state, with 10-12% of children moving into this state from a healthy baseline state. Only 3% of children move into a wasted state within the 3-6m timeframe, 2% within the 6-9m timeframe and 1% within the 9-12m timeframe.

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.01	0.01	0.03	0.10	0.85
6-9m	0.02	0.01	0.02	0.12	0.84
9-12m	0	0.01	0.01	0.11	0.87

Table 5.18: Pooled probability of moving from healthy to each state

5.8 Chapter conclusions

The analysis in Section 5.3 allowed us to test hypotheses 1-3, which we developed from previous literature. On the whole, results tended to agree with the hypotheses outlined in Section 1.3.2. We adopted a stochastic modelling approach to test the hypotheses. This approach allowed us to determine the probabilities of moving between various states conditional on the previous state. In this conclusions section, we discuss our findings, comparing them with other published literature.

5.8.1 Pooling datasets

We utilised the methodology developed in Section 2.9.1 to determine whether transition matrices were significantly different from one another. This provided justification for pooling datasets. This gave us more statistical power for the analyses within Appendix B, and further analysis within the programme of work (Chapter 6).

Each row of the transition matrices represented samples from multinomial distributions, therefore a “two sample” test was needed to determine whether paired rows were significantly different (see Equations (2.34) and (2.35)). A problem with the form of the GLRT in expression (2.34) was that it required non-zero probabilities. We added 0.5 to every frequency so each probability was non-zero. However some frequencies were small

which drastically affected results. The form of (2.34) is asymptotically equivalent to Equation (2.35), Pearson's χ^2 test (Pearson 1900). This means (2.34) and (2.35) differ by quantities that tend to zero as $n \to \infty$. While development of the GLRT was not a complex likelihood argument, no reference within the literature comparing transition matrices in this way could be found, although, a similar "one sample" test was presented by Bickenbach & Bode (2001), applied to financial data. This GLRT can be used in any application to determine whether two transition matrices are significantly different, where samples used to calculate the probabilities within each row are multinomial distributed. The datasets that were pooled were: middle class area and South Africa, village and peri urban slum. This resulted in 4 new datasets listed below:

- South Africa/middle class
- Urban slum
- Malawi
- Peri-urban slum/village

5.8.2 What nutritional states are most likely to lead to death?

Previous research has focused solely on determining whether nutritional status is a predictor of mortality. We hypothesised that wasting-stunting, wasting and stunting were all significant predictors of mortality based on this research.

Vesel et al. (2010) used sensitivity and specificity to determine whether stunting, wasting and underweight were good predictors of mortality in Ghana, India and Peru. Underweight and wasting served as predictors in some sets over the timeframes. Vesel's approach identified which predictors were good, but not whether they were significant, nor did the method allow probabilities to be calculated.

More recent research has suggested that those who are wasted-stunted are at a much higher risk of death those who are wasted, who are in turn at a higher risk than those who are stunted. Olofin et al. (2013) and McDonald et al. (2013) used Cox proportional hazards models to calculate hazard ratios for children who were wasted, underweight and stunted within the same dataset (using healthy as the reference category). Olofin and McDonald categorised children based on their most recent anthropometric measurement. However, the proportions of children in different states vary by age and cohort (see Figure 5.1) meaning that depending on when the baseline state is determined, different proportions of children will be in the different states. Our approach

allowed us to specify the baseline state at different ages and allowed the transition period to vary over time.

Both Olofin's and McDonald's studies found that wasted children were at a higher risk of death than stunted children. Olofin et al. (2013) obtained HRs of 9.4 for severe underweight, 11.63 for severe wasting and 5.48 for severe stunting (severe: $Z < -3$, compared to reference of $Z \geq 1$).

McDonald found HRs of 1.47 for stunted (S), 2.3 for wasted (W), 2.49 for underweight (U), 3.36 S+U, 4.69 W+U, 12.25 W+S+U. O'Neill et al. (2012) also used Cox proportional hazards models to determine whether the categories ($Z < -3, -3 < Z < -2, Z > -2$) of HAZ and BMI-for-age were significant predictors. Categories 2 and 3 of BMI-for-age were significant compared to category 1, with odds of 3 and 8.6 respectively. Those in category 3 for HAZ were 2.3 times more likely ($p = 0.01$).

The research conducted by Vesel et al. (2010), Olofin et al. (2013), McDonald et al. (2013) and O'Neill et al. (2012) suggests that children who are wasted-stunted will be most likely to die, followed by wasted, then stunted (*Hypothesis 1*).

Generally, we found this to be true in our pooled dataset. Those who were wasted-stunted were most likely to die, and those who were wasted were more likely to die than those who were stunted.

In the 3-6m timeframe wasted-stunted children were 11.1 times more likely to die than healthy children, wasted children were 5.3 times more likely and stunted children were 3.0 times more likely. In the 6-9m timeframe, the RRs of death were 5.4 for wasted-stunted children and 1.0 for stunted children (no RR for wasting could be calculated) and in the 9-12m timeframe the RRs were 10.0, 11.5 and 4.8.

An important result is that while the probability of death decreases as children age, children in adverse nutritional states are less vulnerable in the middle period in comparison to healthy children.

Another interesting result is that within the middle class cohort, children were much more likely to revert to a healthy state from adverse states. Those who were wasted-stunted, wasted and stunted at 3 months had respectively a 33%, 75% and 50% chance of moving to a healthy state at 12 months. These proportions are much higher than for the less affluent areas. For example in the village area the proportions were 7%, 26% and 31% and in the peri-urban slum 2%, 23% and 15%. This indicates children from relatively

affluent backgrounds are much more likely to recover from undesirable states, and that socio-economic status may influence a child's ability to recover.

5.8.3 Does wasting lead to stunting?

It is generally assumed that wasting occurs early in the malnourished state and then progresses onto stunting. A number of studies have suggested that this is the case. These studies have assessed whether measures which identify children as short or thin are related. Richard et al. (2012) found that wasting at 6-11 and 12-17 months was associated with decreased height-for-age Z score at 18-24 months and Doherty et al. (2001) found that day 1 WHZ in children aged 6-36m was predictive of change in HAZ over 90 days. Costello (1989) found relationships between initial WHZ and height velocity and a negative relationship with weight velocity. Victora (1992) regressed the prevalence of wasting on stunting in regions around the world and Walker et al. (1996) found weight-for-height predicted linear growth.

These published results tend to indicate that measures of wasting and/or stunting are related in groups of children, and that over time, wasted children tend to gain weight rather than height after initial wasting. Calculating correlation coefficients and building linear models to assess average group relationships between measures of relatively thin and short children are adequate but they only allow us to quantify the dependence of one on the other. In contrast, our analysis, using a stochastic approach, allows us to model the conditional probability of how individual children move from healthy and wasted to stunted states.

We hypothesised (*Hypothesis 2*) that children who are wasted at baseline would tend to move to a stunted state. Furthermore, we expected to observe children moving from a healthy state straight to a wasted state (no papers could be found which investigate which states children are most likely to move to from a healthy baseline state).

Our findings support the hypothesis that wasting increases the likelihood of stunting, but only in the 3-6m period. Compared with healthy children, the RR of moving from a wasted to a stunted state was 1.8 over months 3-6, with 18% of children moving from wasted to stunted, and 10% from healthy to wasted. Within the 6-9m and 9-12m timeframes, the RRs were not statistically significant, with RRs of 1.4 (17% vs 12% moving to stunted respectively) and 1.0 (12% vs. 11%). This indicates healthy children are just as likely to enter a stunted state as wasted children in latter time periods.

However, the RRs of moving from a wasted state to a wasted-stunted state were 17.7, 9.0 and 24.9 within the 3-6m, 6-9m and 9-12m timeframe, which does imply that wasting leads to stunting. In fact, this implies that wasted children become wasted-stunted before becoming stunted as many children move from wasted-stunted to stunted.

When analysing transition probabilities within individual cohorts, those from more affluent backgrounds were more likely to move to a healthy state, and those from poorer backgrounds were more likely to move to a stunted state. In the middle class area, it was 75% likely that children would move from a wasted to healthy state within the 3-6m timeframe and 100% likely in the 6-9m timeframe (no probability could be calculated for the 9-12m timeframe). Similar proportions of children moved from wasted to healthy states in South Africa, 63% in the 3-6m timeframe and 70% in the 6-9m timeframe. However, in the 9-12m timeframe, almost all children within this cohort remained wasted.

Children were likely to move directly from a healthy state to a stunted state (around 10% likely over all three timeframes), much more likely than to a wasted state (3%, 2% and 1% in the 3-6m, 6-9m and 9-12m timeframes). Furthermore, children were much more likely to move either from healthy to healthy to stunted ($p_{HH}.p_{HS} = 0.85 \text{ x } 0.12 = 0.10$), or healthy to stunted to stunted ($p_{HS}.p_{SS} = 0.10 \text{ x } 0.72 = 0.07$) than from healthy to wasted to stunted ($p_{HW}.p_{WS} = 0.03 \text{ x } 0.17 = 0.0051$). This indicates children are more likely to move directly into a stunted state than from healthy to stunted via wasted.

5.8.4 Is stunting irreversible?

Previous research has suggested that stunting is largely irreversible (*Hypothesis 3*) and that children in developing world countries are shorter on average than children from affluent societies (Golden 1994). For example, Victora et al. (2010) found children in low to middle income countries' HAZ scores quickly fall behind the WHO standard, especially in South Asia and sub-Saharan Africa. It is hypothesised that catch up growth is possible with a change of environment (Beaton et al. 1990). Martorell (1992) found that children who were given energy supplementation in a Guatemalan randomised trial were half as likely to be stunted at 3 years (proportions of stunted were the same at the start), indicating that catch-up grown is possible with intervention, and Schumacher et al. (1987) found children who moved to the USA caught up with their US counterparts since they arrived. Potential for catch up is generally thought to be limited after 24 months,

although Adair (1999) found evidence of catch up growth after this period. 63% of her sample were stunted at age 2, of whom 30% were no longer stunted at age 8.5 and 32% were no longer stunted at age 12.5. Adair found those not stunted at age 2 were 4 times more likely to not be stunted at age 8.5. Vella et al. (1994) found that the best predictor of stunting is previous stunting, and that mother's education was also a significant predictor.

Our findings agree with the hypothesis that stunting is largely irreversible for those that remain in the same environment. Stunted children are 59%, 72% and 80% likely to stay stunted and 29%, 21% and 15% likely to recover in timeframes 3-6m, 6-9m and 9-12m respectively. This implies that with increasing age the probability the child will stay stunted increases.

The RRs of remaining stunted were 5.9, 6.2 and 7.1 within the evaluated timeframes. Furthermore, the RRs of moving to a wasted-stunted state were 9.8, 4.6 and 6.4, implying that children who are stunted will not only remain stunted, but are more likely to also become wasted (compared with those who are healthy). In comparison, the RRs of moving to a wasted state were not significant, implying children who are initially stunted are at no more risk of becoming wasted than healthy children. Stunting significantly increased the RR of death in both the 3-6m (3.0) and 9-12m timeframes (4.8) but not the 6-9m timeframe. These results suggest that stunting is largely irreversible, however children who are naturally small tend to stay small, so our results may partially reflect genetic stature.

5.8.5 Further inference

Inference within this chapter was made on the overall pooled dataset, and therefore the results quantify the general transitions within the developing world. However, from the results of the GLRT, it is clear that transition probabilities are both time - and cohort - dependent. The results must therefore be treated as a generalised description of how children grow on average within the developing world. Further inference on the 4 pooled datasets can be made by referring to Appendix B, allowing distinction as to how children from the different cohorts move between states. These were not summarised within the main body of the thesis due to the large number of transition probabilities and RRs. However, these can be interpreted in the same way as the general pooled set.

Chapter 6

Predicting mortality from anthropometry

6.1 Introduction

Past research has suggested that measures of size and growth are both significant predictors of mortality. However, these papers have published conflicting results in terms of which is the best predictor. This chapter aims to address this issue by adopting weighted survival analysis, a method which has not yet been applied to data of this form. The original aim was to use generalised conditional weight gain as a predictor within a Cox proportional hazards model to determine whether growth over a recent time period is a predictor of mortality. Another model would then be built including only weight-for-age Z score at the most recent time point and the models would be compared. However, including two weight-for-age Z scores at t_1 and t_2 is the same, algebraically, as using generalised conditional weight gain in terms of predictive ability. Furthermore, dropping one of the terms changes the model from using growth as a predictor to using size as a predictor. We therefore adopt this approach to test hypothesis 4.

Survival analysis is a form of analysis for time-to-event outcomes, in this case we predict 'time until death'. Standard analyses such as linear regression are not adequate while predicting mortality because analyses do not cope with censoring and survival times are not normally distributed. Therefore a regression model tailored to survival data is needed, such as a Cox proportional hazards model.

During the initial modelling phase, the supervisory team hypothesised that weight gain would be the most valuable predictor. In fact, quite the opposite was found. We found

that in almost all analyses the same single variable stood as the best predictor – most recent weight-for-age. However, the survival analysis conducted on the datasets as they stood was dominated by children within the centre of the population distribution. Therefore, models identified the best predictors for the majority of children within each dataset. However, we are interested in children within the tails of the distribution.

Our initial thoughts were that it is children outwith the healthy range within the extreme regions (very low weight Z scores) that clinicians are more interested in, where weight gain is likely to be a discriminating factor. To develop a model which described the best predictors outwith the healthy range, an initial thought was to take a subsample of children who are considered extreme, determined by some threshold value (perhaps outside the normal range), and restrict analysis to them. However, this analysis would have little power as the sample size would be small.

We decided to use weightings to over represent those within the extremes, which is a way of correcting for lack of representativeness within datasets. Within our datasets, children within the extremes are not unfairly represented, but conducting analysis on the datasets as they stood would not allow us to determine significant predictors for children within the extremes. By increasing the numbers of children who have extreme measurements, we in effect increase the sample size so that children who are at the extremes are overrepresented. By conducting an analysis on these, effectively larger datasets, we can determine whether certain predictors might be significant for these now overrepresented children.

We set out to determine: to which extent do weightings need to be applied for the form of the model to change, and more specifically, how much do we need to over represent these children for growth to stand as the best predictor.

The previous chapter investigated the likely pathways children take through nutritional states. Part of Chapter 5 provided justification for pooling datasets. Results from Chapter 5 suggested that datasets could be pooled in the following way:

- South Africa/middle class
- Urban slum
- Malawi
- Peri-urban slum/village

These datasets will be used within this chapter since using pooled sets increases sample size and power.

Figure 6.1 displays the survival curves for the 6 separate datasets, as well as the 4 pooled datasets. Analyses will be carried out on the four new pooled datasets. In our analyses within this chapter, we set out to determine the effect on the best fitting model of upweighting children within the extremes.

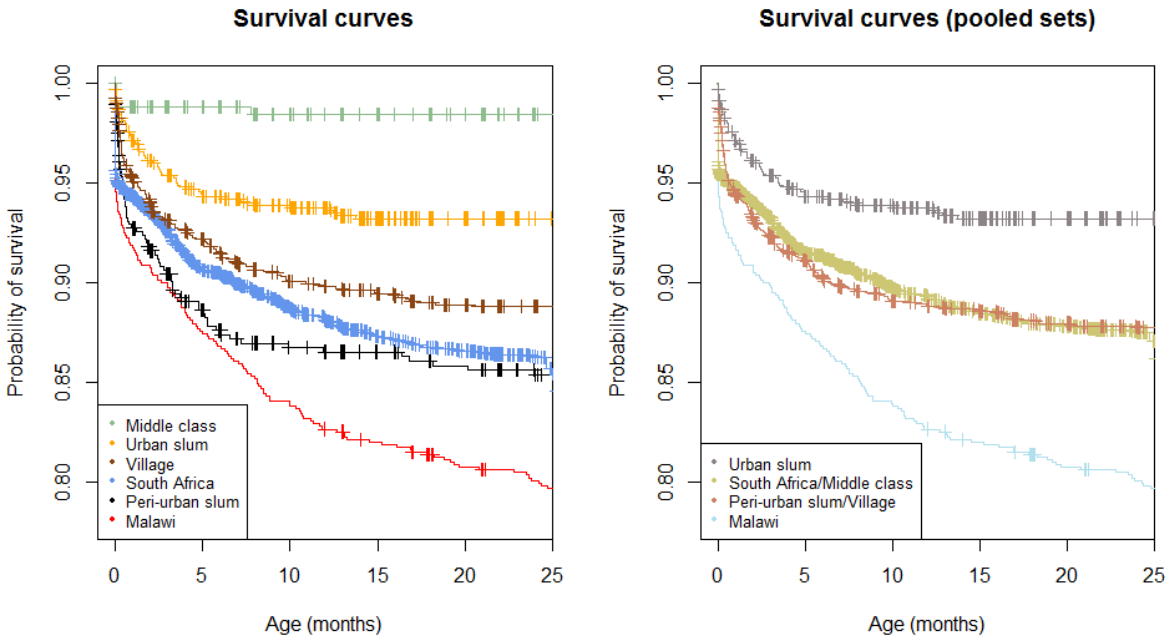

Figure 6.1: Survival curves: (a) original datasets, (b) pooled datasets

6.2 Weighting children within the extremes

As discussed in the introduction, it is children outside the healthy range that we are most interested in. Weights in survival analysis can be viewed simply as replicates of the rows. We will refer to the number of replicates per child as the weightings to avoid any confusion with weight measurements or weight-for-age Z score.

A function was developed which determines how many times each child's weight-for-age Z score should be replicated. The closer a child's weight Z score to the WHO median, on average, the smaller the data weight they receive.

The data weights used were as follows. $\overline{Z}_i$ is the average weight-for-age Z score, calculated as $\overline{Z}_i = (Z_{W_0} + Z_{W_3})/2$ while predicting from 3 months onwards and, $\overline{Z}_i = (Z_{W_0} + Z_{W_3} + Z_{W_6})/3$ while predicting from 6 months onwards:

$$w_i = \begin{cases} \left(\dfrac{\log\left(1 - \phi(|\overline{Z}_i|)\right)}{\log(0.5)}\right)^{\lambda} & \overline{Z}_i < 0 \\ 1 & \overline{Z}_i \geq 0 \end{cases} \qquad (6.1)$$

where the power λ varies from 0 to 1.3. By varying λ a range of different data weights can be assessed. Only children with negative Z scores are upweighted, by an amount that

increases the further they are below the median. The rationale for dividing by $\log(0.5)$ is that when $\overline{Z}_\iota = 0$, $\log\big(1 - \phi(|\overline{Z}_\iota|)\big) = \log(0.5)$. So when $\lambda = 0$ or $\overline{Z}_\iota \geq 0$ the weighting is 1, corresponding to an unweighted model.

A plot of the weighting function for a number of different λ values can be seen in Figure 6.2. The exact values were picked so that a range of applicable weights could be applied to observe the effects of using extreme weights.

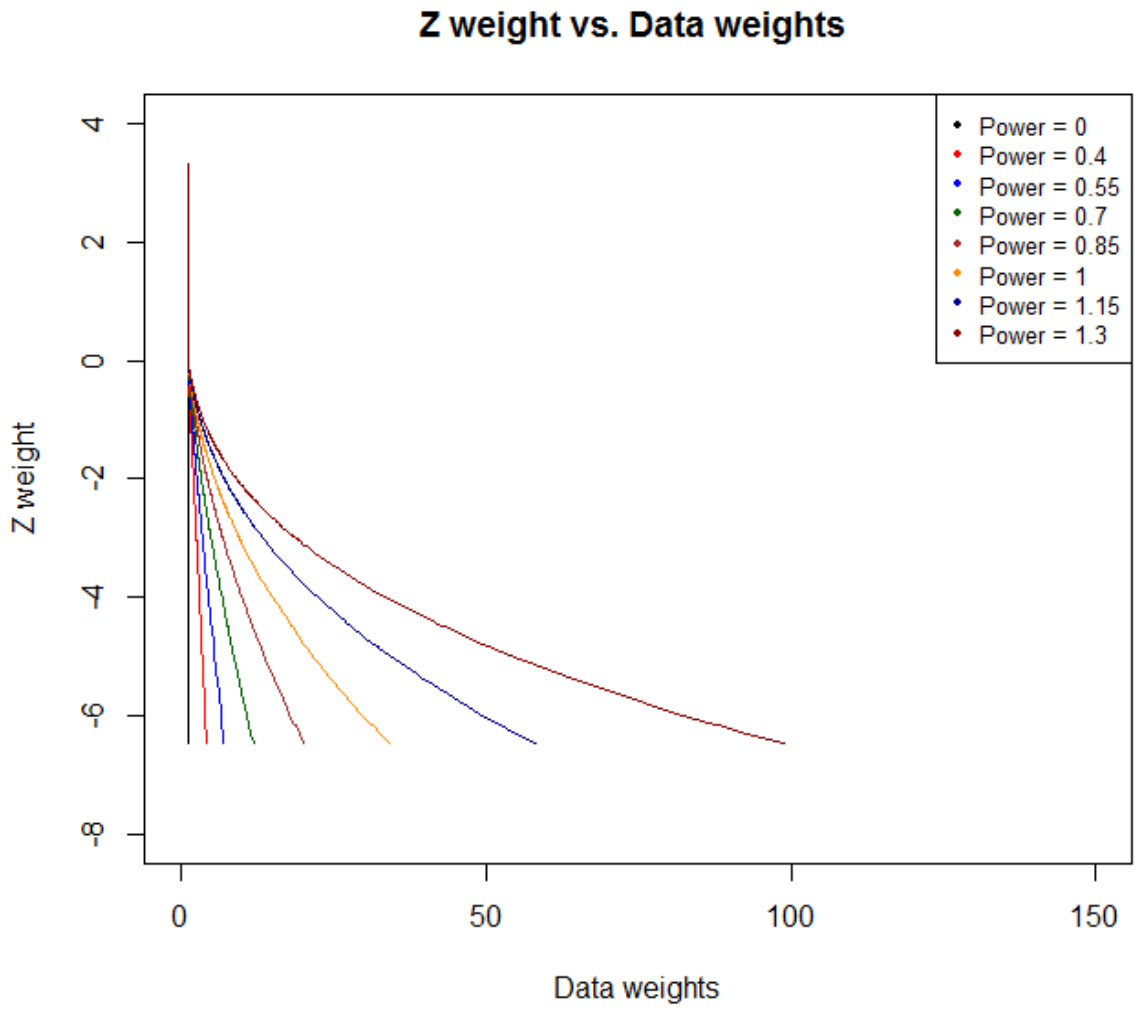

Figure 6.2: Weight function with various values of λ

6.4 Predictions after 3 months

The original proposed analysis was to use Equation (2.19), generalised conditional weight gain, as a covariate since it is a measure of growth. However, using Equation (2.19) is the algebraic equivalent of using weight-for-age as Z scores at t_1 and t_2. Furthermore, if either term is dropped the model simplifies to weight size rather than weight gain.

Model	Covariates
1	W0
2	W3
3	H0
4	H3
5	W0,W3
6	W0,H0
7	W0,H3
8	W3,H0
9	W3,H3
10	H0,H3
11	W0,W3,H0
12	W0,W3,H3
13	W0,H0,H3
14	W3,H0,H3
15	W0,W3,H0,H3

Table 6.1: Covariates used in Cox proportional hazards models (predicting after 3 months)

Models predicting time until death after 3 months were developed using combinations of four covariates: weight-for-age at 0 months (W0), weight-for-age at 3 months (W3), height-for-age at 0 months (H0) and height-for-age at 3 months (H3), all as Z scores relative to the WHO standard. In total 15 models were tested, as seen in Table 6.1. The notation W0, W3, H0 and H3 is used instead of $Z_{W_0}, Z_{W_3}, Z_{H_0}, Z_{H_3}$ as these are easier to read. Since the effective sample size changes as λ increases, BICs are not comparable across different λs. Children are only included within the analysis if all variables for that child are available, otherwise they are excluded.

λ	South Africa/middle class	Urban slum	Malawi	Peri-urban/Village
0	1794	615	585	976
0.4	2646	999	916	1763
0.55	3154	1228	1109	2258
0.7	3734	1532	1360	2934
0.85	4524	1941	1691	3869
1	5565	2499	2133	5176
1.15	6956	3271	2730	7027
1.3	8846	4356	3548	9678

Table 6.2: Effective sample size for each λ (predicting after 3 months)

		South Africa/middle		Urban slum		Malawi		Peri-urban/village	
Model rank	Power (λ)	Model	BIC	Model	BIC	Model	BIC	Model	BIC
1	0	W3	1395.9	W3	164.6	W3	819.8	W3	640.8
2	0	W3,H3	1401.4	W3,H0	170.7	H0	821.0	W3,H0	646.2
3	0	W0,W3	1402.9	W3,H3	170.8	W0	824.3	W0,W3	647.2
1	0.4	W3	2133.8	W3	363.5	W3	1355.1	W3	1303.1
2	0.4	W0,W3	2137.5	W0,W3	367.7	W3,H3	1360.4	W3,H0	1309.1
3	0.4	W3,H3	2138.8	W3,H0	368.4	H0	1360.6	W0,W3	1309.5
1	0.55	W3	2594.8	W3	512.1	W3	1699.5	W3	1767.9
2	0.55	W0,W3	2595.6	W0,W3	514.6	W3,H3	1704.7	W3,H0	1774.3
3	0.55	W3,H3	2599.6	W3,H0	515.9	W3,H0	1705.6	W0,W3	1774.4
1	0.7	W0,W3	3220.1	W0,W3	735.8	W3	2179.1	W3	2450.2
2	0.7	W0,W3,H0	3221.1	W3	735.8	W3,H3	2184.3	W0,W3	2456.8
3	0.7	W3	3224.3	W3,H0	738.1	W0,W3	2184.7	W3,H0	2456.9
1	0.85	W0,W3,H0	4079.0	W0,W3	1071.2	W3	2858.1	W3	3467.8
2	0.85	W0,W3	4083.9	W3,H0	1075.0	W0,W3	2862.6	W0,W3	3474.7
3	0.85	W0,W3,H0,H3	4085.7	W3	1075.0	W3,H3	2863.3	W3,H3	3474.7
1	1	W0,W3,H0	5280.3	W0,W3	1582.5	W3	3834.5	W3	5009.0
2	1	W0,W3,H0,H3	5287.5	W0,W3,H3	1588.0	W0,W3	3836.9	W3,H0	5015.6
3	1	W0,W3	5295.1	W3,H0	1588.5	W0,W3,H0	3839.5	W3,H3	5015.8
1	1.15	W0,W3,H0	6984.6	W0,W3	2366.1	W0,W3	5259.0	W3	7376.8
2	1.15	W0,W3,H0,H3	6992.1	W0,W3,H3	2370.9	W0,W3,H0	5259.3	W3,H0	7382.3
3	1.15	W0,W3	7016.4	W0,W3,H0	2372.2	W3	5260.8	W0,W3	7383.5
1	1.3	W0,W3,H0	9434.1	W0,W3	3571.4	W0,W3,H0	7362.2	W3	11063.9
2	1.3	W0,W3,H0,H3	9440.7	W0,W3,H3	3575.5	W0,W3	7365.8	W3,H0	11066.3
3	1.3	W0,W3,H3	9488.0	W0,W3,H0	3577.2	W0,W3,H0,H3	7368.1	W0,W3	11069.7

Table 6.3: Models ranked by BIC (predicting after 3 months)

		South Africa/Middle		Urban slum		Malawi		Peri-urban/Village	
Model rank	Power (λ)	Model	BIC	Model	BIC	Model	BIC	Model	BIC
1	0	W3	1395.9	W3	164.6	W3	819.8	W3	640.8
1	0.4	W3	2133.8	W3	363.5	W3	1355.1	W3	1303.1
1	0.55	W3	2594.8	W3	512.1	W3	1699.5	W3	1767.9
1	0.7	W0,W3	3220.1	W0,W3	735.8	W3	2179.1	W3	2450.2
1	0.85	W0,W3,H0	4079.0	W0,W3	1071.2	W3	2858.1	W3	3467.8
1	1	W0,W3,H0	5280.3	W0,W3	1582.5	W3	3834.5	W3	5009.0
1	1.15	W0,W3,H0	6984.6	W0,W3	2366.1	W0,W3	5259.0	W3	7376.8
1	1.3	W0,W3,H0	9434.1	W0,W3	3571.4	W0,W3,H0	7362.2	W3	11063.9

Table 6.4: Best fitting models (predicting after 3 months)

Table 6.3 displays the best three models with each set of weightings. Table 6.4 is a subset of this table, displaying only the best fitting model per λ. Note that the cohorts presented in Table 6.4 are in order of affluence, from left to right.

It is clear from Table 6.4 that in all four analyses, weight-for-age Z score at 3 months (most recent size) stands as the best predictor. Height-for-age does not serve as a valuable predictor while predicting after 3 months, inferring that the measure adds little value in children of this age. Velocity enters as the best predictor with weightings of 0.7 within the South African/middle and urban slum cohorts whereas velocity entered as the best predictor with weightings of 1.15 within the Malawi models. Finally, velocity never enters as the best predictor for the peri-urban slum/village. The results show that velocity enters as a significant predictor within more affluent cohorts before entering less affluent cohorts.

6.5 Predictions after 6 months

Model	Covariates	Model	Covariates
1	W0	33	W3,W6,H3
2	W3	34	W3,W6,H6
3	W6	35	W3,H0,H3
4	H0	36	W3,H0,H6
5	H3	37	W3,H3,H6
6	H6	38	W6,H0,H3
7	W0,W3	39	W6,H0,H6
8	W0,W6	40	W6,H3,H6
9	W0,H0	41	H0,H3,H6
10	W0,H3	42	W0,W3,W6,H0
11	W0,H6	43	W0,W3,W6,H3
12	W3,W6	44	W0,W3,W6,H6
13	W3,H0	45	W0,W3,H0,H3
14	W3,H3	46	W0,W3,H0,H6
15	W3,H6	47	W0,W3,H3,H6
16	W6,H0	48	W0,W6,H0,H3
17	W6,H3	49	W0,W6,H0,H6
18	W6,H6	50	W0,W6,H3,H6
19	H0,H3	51	W0,H0,H3,H6
20	H0,H6	52	W3,W6,H0,H3
21	H3,H6	53	W3,W6,H0,H6
22	W0,W3,W6	54	W3,W6,H3,H6
23	W0,W3,H0	55	W3,H0,H3,H6
24	W0,W3,H3	56	W6,H0,H3,H6
25	W0,W3,H6	57	W0,W3,W6,H0,H3
26	W0,W6,H0	58	W0,W3,W6,H0,H6
27	W0,W6,H3	59	W0,W3,W6,H3,H6
28	W0,W6,H6	60	W0,W3,H0,H3,H6
29	W0,H0,H3	61	W0,W6,H0,H3,H6
30	W0,H0,H6	62	W3,W6,H0,H3,H6
31	W0,H3,H6	63	W0,W3,W6,H0,H3,H6
32	W3,W6,H0		

Table 6.5: Covariates used in Cox proportional hazards models (predicting after 6 months)

The 63 models which were applied while predicting after 6 months can be seen in Table 6.5. These models include weight-for-age and height-for-age Z scores at 0, 3 and 6 months.

λ	South Africa/Middle class	Urban slum	Malawi	Peri-urban/Village
0	1429	471	496	679
0.4	2167	746	823	1237
0.55	2582	908	1019	1592
0.7	3113	1122	1278	2081
0.85	3800	1406	1626	2761
1	4700	1790	2098	3719
1.15	5893	2313	2746	5082
1.3	7497	3035	3645	7043

Table 6.6: Effective sample size for each λ (predicting after 6 months)

The effective sample sizes can be seen in Table 5.6, again, BICs are not comparable across different weights as the dataset effectively changes as λ increases.

		South Africa/Middle		Urban slum		Malawi		Peri-urban/Village	
Model rank	Power (λ)	Model	BIC	Model	BIC	Model	BIC	Model	BIC
1	0	W6	887.9	W6	68.2	W6	458.7	W6,H6	165.6
2	0	W6,H3	888.5	H0	70.7	H6	460.9	W6	167.8
3	0	H6	888.5	W6,H0	70.7	W3	460.9	W3,W6	171.1
1	0.4	H3,H6	1300.0	W3,W6,H0	138.6	W6	771.6	W6,H6	338.7
2	0.4	W6,H3	1300.3	W0,W3,W6	139.9	W6,H0	777.2	W3,W6	344.1
3	0.4	W0,H6	1300.7	W6,H0	141.1	W3,W6	777.4	W6	344.3
1	0.55	W0,H6	1548.7	W3,W6,H0	187.3	W6	971.7	W6,H6	455.6
2	0.55	H3,H6	1549.2	W0,W3,W6	188.4	W0,W6	977.4	W0,W6,H6	460.3
3	0.55	W0,W6	1549.9	W0,W3,W6,H0	188.5	W6,H0	977.5	W3,W6	460.3
1	0.7	W0,H6	1881.0	W0,W3,W6,H0	258.4	W6	1250.0	W6,H6	621.2
2	0.7	W0,H3,H6	1882.1	W0,W3,W6,H0,H6	260.6	W0,W6	1255.0	W0,W6,H6	624.2
3	0.7	W0,W6	1882.8	W3,W6,H0	260.8	W6,H6	1255.8	W3,W6	624.7
1	0.85	W0,H6	2332.7	W0,W3,W6,H0,H6	363.8	W6	1644.0	W0,W3,W6,H6	854.2
2	0.85	W0,H3,H6	2333.6	W0,W3,W6,H0	364.8	W0,W6	1647.5	W0,W3,W6	856.5
3	0.85	W0,W6	2335.5	W0,W3,W6,H0,H3,H6	367.8	W6,H6	1649.2	W6,H6	857.8
1	1	W0,H0,H3,H6	2955.3	W0,W3,W6,H0,H6	522.0	W6	2211.8	W0,W3,W6,H6	1183.0
2	1	W0,H6	2955.4	W0,W3,W6,H0,H3,H6	525.4	W0,W6	2212.5	W0,W3,W6	1185.6
3	1	W0,H3,H6	2956.5	W0,W3,W6,H0	528.4	W0,W6,H6	2212.7	W0,W3,W6,H0,H6	1189.3
1	1.15	W0,H0,H3,H6	3820.1	W0,W3,W6,H0,H6	766.9	W0,W6,H6	3035.5	W0,W3,W6,H6	1657.3
2	1.15	W0,W6,H0,H3	3824.5	W0,W3,W6,H0,H3,H6	769.4	W0,W6,H0,H6	3038.5	W0,W3,W6	1659.9
3	1.15	W0,H6	3825.4	W0,W3,W6,H0	781.9	W0,W3,W6,H6	3038.6	W0,W3,W6,H0,H6	1663.0
1	1.3	W0,H0,H3,H6	5042.0	W0,W3,W6,H0,H6	1149.2	W0,W6,H6	4263.0	W0,W3,W6,H6	2346.1
2	1.3	W0,W6,H0,H3	5047.7	W0,W3,W6,H0,H3,H6	1150.6	W0,W3,W6,H6	4264.8	W0,W3,W6	2348.4
3	1.3	W0,W6,H0,H3,H6	5048.4	W0,W3,W6,H0	1177.6	W0,W6,H0,H6	4264.9	W0,W3,W6,H0,H6	2350.3

Table 6.7: Models ranked by BIC (predicting after 6 months)

		South Africa/Middle		Urban slum		Malawi		Peri-urban/Village	
Model rank	Power (λ)	Model	BIC	Model	BIC	Model	BIC	Model	BIC
1	0	W6	887.9	W6	68.2	W6	458.7	W6,H6	165.6
1	0.4	H3,H6	1300.0	W3,W6,H0	138.6	W6	771.6	W6,H6	338.7
1	0.55	W0,H6	1548.7	W3,W6,H0	187.3	W6	971.7	W6,H6	455.6
1	0.7	W0,H6	1881.0	W0,W3,W6,H0	258.4	W6	1250.0	W6,H6	621.2
1	0.85	W0,H6	2332.7	W0,W3,W6,H0,H6	363.8	W6	1644.0	W0,W3,W6,H6	854.2
1	1	W0,H0,H3,H6	2955.3	W0,W3,W6,H0,H6	522.0	W6	2211.8	W0,W3,W6,H6	1183.0
1	1.15	W0,H0,H3,H6	3820.1	W0,W3,W6,H0,H6	766.9	W0,W6,H6	3035.5	W0,W3,W6,H6	1657.3
1	1.3	W0,H0,H3,H6	5042.0	W0,W3,W6,H0,H6	1149.2	W0,W6,H6	4263.0	W0,W3,W6,H6	2346.1

Table 6.8: Best fitting models (predicting after 6 months)

Eight different data weights were applied to the 63 models. The best three of each data weight category in terms of BIC can be seen in Table 6.7. A subset of Table 6.7 can be seen in Table 6.8, which shows the best model chosen by BIC.

Again, as with our models predicting after 3 months, the most recent measure of size stands as the best predictor within the unweighted analysis, apart from within the peri-urban slum/village cohort as height-for-age is also included as a significant predictor, inferring weight-for-height is a valuable predictor for children of this age within the dataset.

As weighting increases, measures of velocity enter as the best predictor within the other 3 cohorts. With a weighting of 0.4, height velocity enters as the best predictor within the South Africa/middle class cohort (height velocity/weight-for-height is the best predictor with heavier weightings). Weight velocity/weight-for-height enters as the best predictor with a weighting of 0.4 within the urban slum model. Not until a weighting of 1.15 is applied for the Malawian model does a measure of velocity enter as well as height-for-age.

Notably, as height-for-age tends to be included within the weighted models, this may imply that this measure may serve as a valuable predictor after 6 months in some populations. This is in contrast to the results predicting after 3 months (height-for-age was not found to be a valuable predictor within any of the >3m models).

6.6 Stratifying by HIV status

As mentioned in Section 3.5, HIV status is available within a number of the datasets. Analysing the datasets while including HIV status as a predictor was considered. However, HIV status was not available for the Malawi dataset, which made comparative analyses between datasets impossible. In addition, there were no cases of HIV within the Pakistani dataset.

Within the South Africa dataset, both the status and timing of HIV were available. Within the Malawi data set there were cases of HIV but they were not recorded. There was however information indicating whether the children's mothers were HIV positive at baseline.

In this section, two analyses are carried out using stratified variables – using child's HIV status in South Africa, and using mother's HIV status in Malawi. The procedure described in Sections 6.4 and 6.5 is carried out but HIV status is included as a stratified variable. Stratification allows us to fit a model that has a different baseline hazard for each stratum where the variable HIV status is a two level factor variable.

The first analysis was conducted on the South African dataset. The number of HIV cases per month can be seen in Table 6.10. Measurements for HIV status were taken whenever possible. However a child may have contracted the virus between the last time they tested negative and the first time they tested positive. Therefore, HIV status was defined in this analysis as the midpoint between the last negative test and the first positive.

The timing that children are infected with HIV can be categorised into three groups:

1.) In utero
2.) During delivery (most common)
3.) Postnatal during breastfeeding

Children who test positive at 6 weeks have either contracted the virus in utero or during delivery, children who have contracted the disease after 6 weeks have contracted the virus during breastfeeding.

Those who test positive at 6 weeks (in utero or during delivery) can be categorised into two further groups:

a.) Fast progressors
b.) Slow progressors

Half of those in group a who receive no treatment will die within 2 years (Becquet et al. 2012).

	>3 months		>6 months	
λ	Without HIV	With HIV (stratified)	Without HIV	With HIV (stratified)
0	W3	W3	W6,H3	W6
0.4	W3	W3	W6,H3	W6
0.55	W0,W3	W3	W0,W6	W6
0.7	W0,W3	W3	W0,W6,H6	W6
0.85	W0,W3	W3	W0,W6,H6	W6,H6
1	W0,W3,H0	W3	W0,W3,W6,H6	W0,W3,W6,H6
1.15	W0,W3,H0	W3	W0,W3,W6,H6	W0,W3,W6,H0,H6
1.3	W0,W3,H0	W0,W3,H0	W0,W3,W6,H0,H6	W0,W3,W6,H0,H6

Table 6.9: Stratified South Africa models

Since our analyses predict mortality after 3 and 6 months, the analysis will take into account all of those within the 'in utero' group, the 'during delivery' group and part of the postnatal group. Therefore, the stratification variable does not separate those who do and do not have HIV; it separates those who do and do not have HIV at 3/6 months.

Month	0	1	2	3	4	5	6	7	8	9	10	11	12
Cases	53	162	187	201	210	223	227	232	236	238	238	240	244
Month	13	14	15	16	17	18	19	20	21	22	23	24	
Cases	246	250	254	255	255	255	256	256	265	258	258	268	

Table 6.10: Number of HIV cases per month in South Africa

For the Malawian dataset, the same procedure is carried out but using the mothers HIV status at baseline (birth of child) as a stratified variable. 149/813 (18.3%) of mothers had HIV at baseline.

Within Table 6.10, the best Cox proportional hazards models with variable weights λ can be seen, as can the stratified Cox models. It is clear that as λ increases, the stratified models stay relatively simple compared to the unstratified models. Only with powers of 0 and 0.4 are the two different models the same, while predicting after 3 months.

When including mothers HIV status at baseline, significant predictors only change while using large data weights. While predicting after 3 months, only models using powers of 1 and 1.15 change. While predicting after 6 months, only using data weights where λ is equal to 1.3 does the best fitting model change.

	3 months		6 months	
λ	Without HIV	With HIV (stratified)	Without HIV	With HIV (stratified)
0	W3	W3	W6	W6
0.4	W3	W3	W6	W6
0.55	W3	W3	W6	W6
0.7	W3	W3	W6	W6
0.85	W3	W3	W6	W6
1	W3	W0,W3,H0	W0,W6,H6	W0,W6,H6
1.15	W0,W3	W0,W3,H0	W0,W6,H6	W0,W6,H6
1.3	W0,W3,H0	W0,W3,H0	W0,W6,H3,H6	W0,W6,H6

Table 6.11: Stratified Malawi models

6.7 Chapter conclusions

Research has indicated that measures of size are significant predictors of mortality. Papers by Vesel et al. (2010), Olofin et al. (2013), McDonald et al. (2013) and O'Neill et al. (2012) indicate that nutritional status is a predictor of mortality. Olofin and O'Neill's work suggests that magnitude of the Z scores representing nutritional status hugely influences the odds of death, which indicates that relative size on a continuous scale is a significant predictor. Furthermore, their research suggests that low WHZ and WAZ are more predictive and tend to lead to lower odds of survival than low HAZ. Many more papers were found that investigated measures of size as predictors of mortality. Low birth weight was shown to predict mortality by McIntire et al. (1999) and Bairagi & Chowdhury (1994) who found that weight-for-age, height-for-age and arm circumference were significant predictors. Fawzi et al. (1997) found weight-for-height and height-for-age were predictors.

Four papers were found which compared size and growth as predictors of mortality, those by Bairagi et al. (1985), Briend & Bari (1989), Bairagi et al. (1993) and O'Neill et al. (2012). Bairagi et al. (1985), Briend & Bari (1989) and Bairagi et al. (1993) all conclude that recent size is the best predictor of mortality, using weight-for-age % (NCHS reference) to represent size and the raw untransformed weight velocity to represent growth. In fact, Bairagi et al. (1985) found that velocity was not even a significant predictor.

O'Neill et al. (2012) on the other hand, found that velocity was a better predictor than size. However, O'Neill used BMI-for-age Z score to represent size, and 3 month growth velocity Z score to represent growth. O'Neill speculates that the difference in results may

be down to the use of the WHO standard to compute Z scores. Furthermore, Briend & Bari (1989) note that young children in their cohort were more likely to die, indicating a confounding effect may explain the poor performance of velocity. However, including age as a confounding factor made no difference to the relationship between velocity and mortality.

We applied weighted Cox proportional hazards models to the four pooled datasets to investigate this problem further. We used weight-for-age and height-for-age Z scores relative to the WHO standard as predictors at multiple times in our analysis. We originally planned on using Equation (2.19), generalised conditional weight gain, as a covariate. However, using Equation (2.19) to calculate weight and height gain is the algebraic equivalent as including weight-for-age and height-for-age as Z scores at t_1 and t_2. Furthermore, dropping terms allows assessment of size as a predictor.

We used Cox proportional hazards models to address this aim. A problem with standard analyses is that they are dominated by children within the centre of the population distribution. We are interested in children within the tails of the distribution, therefore applying Cox proportional hazards models similar to work by Olofin et al. (2013), McDonald et al. (2013) and O'Neill et al. (2012) is not appropriate for our analysis. We adopted a weighted Cox proportional hazard model so that we could over represent children with lower average weight-for-age Z scores.

Average weight-for-age Z scores were calculated as the average Z score between months 0 and 3 while predicting after 3 months, and the average over months 0, 3 and 6 while predicting after 6 months. The average Z score was then used in Equation (6.1) to determine the weighting. The weighting function was based on the inverse probability of a normal distribution so that children with low average Z scores received more weight, but scaled so that the minimum weight applied was 1 (any positive Z scores also received a weight of 1). Data weights in Cox proportional hazards models are replicates of the rows and a number of different data weights were used on the data while effects on the best fitting model by BIC were observed.

This weighted approach proved to be a very valuable method of analysing data of this kind. Results of the unweighted analyses agree with research by Bairagi et al. (1985), Briend & Bari (1989), Bairagi et al. (1993), Bairagi & Chowdhury (1994) and Fawzi et al. (1997) in that weight-for-age Z score is a significant predictor of mortality and partly with

Olofin et al. (2013), McDonald et al. (2013) and Vesel et al. (2010) as their research found groupings of weight-for-age Z scores were predictors.

In our weighted analysis, we set out to determine the extent to which do weightings need to be applied for the form of the model to change, specifically, what magnitude of weightings need to be applied before measures of velocity enter as the best predictors. While predicting after 3 months we found that as weighting increased, velocity entered as a significant predictor earlier for more affluent cohorts, whereas heavier weightings were needed for velocity to enter as the best predictor for less affluent cohorts. We hypothesise that since the majority of children within the less affluent cohorts are undernourished, and tend to fall from birth relative to the WHO standard, size best describes the way that children have arrived at their final state as the majority of children arrived there in the same way. Weighting the extreme children more heavily does not differentiate children who are at higher risk as most children are already at risk. The model therefore identifies that the majority of children in the weighted set are still those who have fallen. Refer to Figure 6.3 for a visual representation of how weighting impacts the proportion of children: in the healthy range, who fall, and who are naturally small. The width of the bar represents the proportion of children in each group. In Figure 6.3a, it is those who a have fallen who make up the majority of children. In Figure 6.3b, it is still those who have fallen who make up the majority.

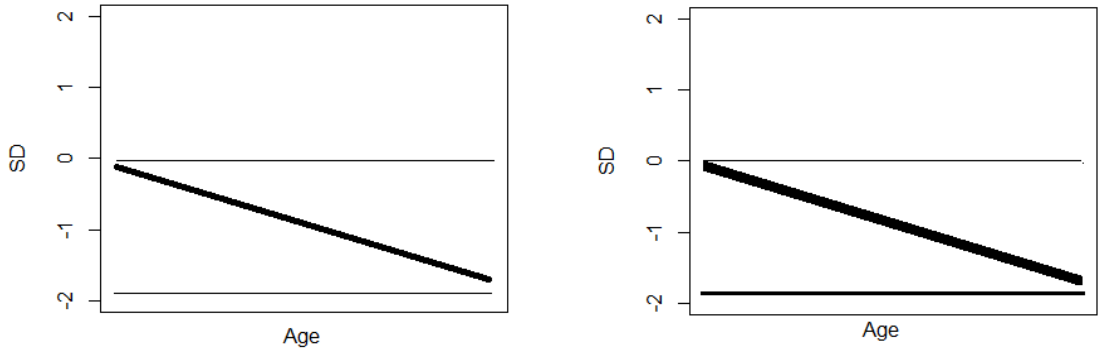

Figure 6.3: High prevalence cohort: (a) unweighted, (b) weighted

In unweighted analyses within more affluent populations, the majority of children are in the healthy range. The model identifies size as the best predictor as the majority of children have arrived at their size in the same way – tracking along the same centile. Therefore, the model does not use growth to differentiate between those who are and are not at risk, as it is dominated by those within the normal range. In the weighted

analysis, those who have fallen and those who are naturally small are now overrepresented. Notice the bars in Figure 6.4b have equal width, in contrast to Figure 6.4a, where the bar at 0SD (representing healthy children tracking the 50th centile) is thicker than the others. Velocity differentiates between those who are at risk (falling centiles) relative to those who are naturally small.

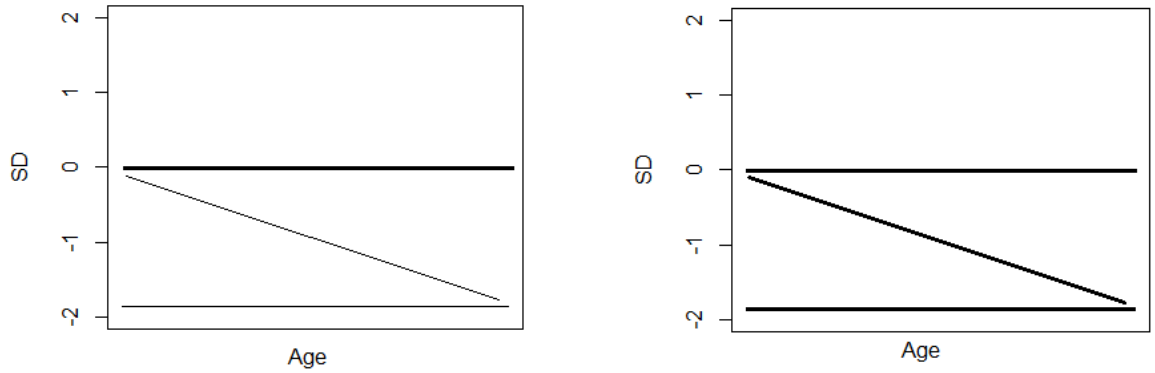

Figure 6.4: Low prevalence cohort: (a) unweighted, (b) weighted

In practice, HAZ was not as valuable a predictor as WAZ, although this was expected from previous research. For velocity to be a significant predictor, children within the extremes had to be very heavily weighted. In fact, O'Neill et al. (2012) found that only HAZ at 3m less than -3SD was predictive of 2 year mortality and McDonald (2013) found stunting was less predictive than wasting or underweight.

After 6 months the unweighted analysis showed that WAZ is the best predictor, while HAZ entered as a significant predictor alongside WAZ with higher weightings. This implies that while making predictions from later ages, HAZ may add value to WAZ. However, again, in reality, the recent size tends to be the best predictor for the majority of children, and other measures only enter after artificially adapting the datasets.

In summary the results suggest that to predict mortality weight-for-age is better than height-for-age, and recent weight-for-age is better than earlier weights. Growth adds little value over size for the majority of children. Stratifying by HIV status (South African dataset) had an effect on the best fitting models. Models became much simpler, implying that recent HIV status should be included within any further modelling approaches. HIV status of the mother at baseline (Malawi dataset) did not affect the form of the model unless very large weights were applied, suggesting that the mother's HIV status is a poor proxy for the child's HIV status. As HIV transmission rates are generally around 15-45% without intervention, and as low as 5% with intervention, our results are not surprising.

Chapter 7
Conclusions, limitations and further work

7.1 Overview

This thesis set out to devise statistical solutions to problems in undernutrition, identifying anthropometric measurements which are useful predictors of adverse future outcomes. While each chapter contains its respective conclusions, it is helpful to reflect on our findings in this final chapter. Here, the study aims will be discussed, as well as how they were developed. The chapter is split into sections for each aim, describing them in detail while discussing the findings for each. Prior hypotheses based on past scientific research will be compared with our results, limitations will be discussed, as will proposed future work.

7.2 Data Formatting and modelling

Chapter 1 was an introductory chapter, briefly reviewing undernutrition as well as its causes and effects. Measures of undernutrition were introduced as were their uses for the diagnosis of undernutrition. Investigating the background of undernutrition led us to the concept of using anthropometry to predict future outcomes. Literature was reviewed in the general area and as a result, we developed four hypotheses and 6 research aims. Aims 1-3 focused on data formatting and modelling, which this section of this chapter describes.

7.2.1 The datasets

The initial aims focused on data formatting, modelling and summarising the sets. Reformatting the data into a common format prepared them for all further analysis within the thesis.

We imported all datasets into a common format and summarised them in a database. Three developing world datasets were available, cohorts from: Malawi, Pakistan and South Africa. The data were collected as a result of three prospective cohort studies: the Malawian LCSS, Pakistani LSS and South African VTS. The Malawian LCSS was set up to provide data on the health of both pregnant women and their children under 5 years old, the Pakistani LLS was collected to "characterise the determinants of child health in a rapidly urbanising community", and the South African VTS was set up to examine patterns of infant feeding.

In general, anthropometry is an inexpensive and valuable way of evaluating the nutritional status of children, but collecting data of this kind is hard and can be labour intensive. The process of gathering this type of data is prone to errors - health workers need to be adequately trained and there is variability between staff.

However, these datasets are of very high quality and are representative of their population. Data were collected by carefully trained and supervised research staff. Within the Malawi study, trained local research assistants visited children every month up to 18 months and every 3 months thereafter (Kulmala et al. 2000). Within the Pakistani study, two months were spent training health teams on how to interview families and take anthropometric measurements, and during the research period, the health teams were retrained every month, measuring child growth every month in the four study areas: a village, urban slum, peri-urban slum and middle class area (Jalil, Lindblad, Hanson, Khan, Ashraf, et al. 1993). In the South African study, nurses and clinical assistants aimed to see infants as soon as possible after delivery and clinical visits were planned so measurements could be taken at the scheduled measurement dates (Bland et al. 2010). The frequency of scheduled measurement dates provided enough data to carry out numerous statistical analyses.

All three datasets have high rates of undernutrition and mortality. The proportions of those within wasted-stunted, wasted and stunted states were variable (depending on age), but in general there were high rates of undernutrition. By 24 months, 20% and 14% of those in the Malawian and South African cohorts had died. In Pakistan 12%, 15%, 7%

and 2% of those in the village, per-urban slum, urban slum and middle class area had died by 24 months, respectively.

The high mortality rates provided the opportunity to test prediction models for adverse outcomes that would be underpowered using developed world datasets, as mortality rates are generally low in the developed world. The statistical models and tests we employed produced statistical significant results, which may not have been possible with smaller datasets with less frequent observations.

In total, there were 6897 children in the pooled dataset. It must be noted however that these datasets were collected only over three geographical locations, two of which are relatively close - in Africa below the equator. In this sense, inference can be made about the sets but due to the lack of geographical diversity care must be taken when making generalisations based on the conclusions. On the other hand, undernutrition is mostly prevalent in East Africa and south-central Asia (in terms of underweight) (Black et al. 2008), where two of these datasets were collected.

7.2.2 Issues with the datasets

As with all datasets of this nature, missing values were a problem. If missing values are random, then there is generally no cause for concern as long as there are enough children still in the study to make meaningful inference. Monotone missing pattern refers to those who miss a measurement occasion and measurements for that child are never observed again. Non-monotone missing pattern refers to missing data where the subject is measured again at a later time point. Problems arise when either monotone or non-monotone missing pattern is present but the probability of missing data is dependent on the unobserved data, called Missing Not At Random (MNAR) (Rubin 1976) - this is down to unobserved factors which result in missing values. A potential example of these missing data in our research are the proportions of children who moved away from the study area or dropped out, which could potentially be down to unobserved factors. For example, in the Pakistani set, 13% of children moved away from the study area and 6% refused to continue to take part which could potentially have been down to unobserved factors which were not random. Our analyses did not control for these potentially unobserved factors, which may result in false conclusions about the whole study population. If, for example, the 6% of children who dropped out did so because of illness, and later died, the published mortality rate for the dataset would be biased downwards.

Unfortunately it is hard to tell whether these data are MNAR without observed data which can be used to identify trends, and while this cannot be quantified, it is worth noting that it is possible that at least some data may be MNAR, which could possibly result in bias.

One other concern was that data for the three datasets were collected in different decades. However, for the Pakistani set (in which data was collected over ten years), we established that growth did not change by era within the sample.

In summary, the database we created comprised three high quality diverse developing world datasets, each with scheduled measurement dates every month or less within early months. Throughout the analysis period, the database provided enough power not only to make general inferences, but also to make meaningful comparisons across the populations. This database is a resource for future studies and is currently being used for research in stunting at the University of Glasgow.

7.2.3 Developing growth reference charts

Once the database was created, we aimed to develop growth reference charts for the three datasets. The main reason for developing the charts was to demonstrate the use of current growth chart modelling methodology, though the charts we developed could potentially be used as internal references for the study areas.

With increasing age, the distribution of children's weights increase on average, spread out and tends to be skewed. The LMS method is a flexible approach, allowing users to obtain normalised growth centiles while taking both skewness and increasing variability into account. We utilised the method to model the datasets, creating weight-for-age references. Three curves were produced using the method: the Box-Cox power (L), the median (M) and coefficient of variation (S). Using the Box-Cox normal distribution, plugging in the values of L, M and S at the required age, the estimated distribution of weights can be obtained. Using the cumulative distribution function of the Box-Cox normal, the required values that lie on each centile (or any quantile for that matter) can be obtained. In fact, this method is the equivalent of using GAMLSS while assuming the three parameter Box-Cox Cole and Green (or Box-Cox normal) distribution, which we used to model the datasets. This allowed us to accurately represent growth while adjusting for non-normality, using penalised beta splines to generate the curves (Eilers & Marx 1996).

In any further research, sex could be included as a categorical variable and one model

could describe the whole pooled population. Furthermore, by including sex as a categorical variable, the t-value for sex could be compared with the t-distribution allowing researchers wishing to determine if sex is a significant variable for each curve. If the p-value is not significant, the data could be pooled for that curve, increasing the sample size and precision. GAMLSS provides the framework which allows researchers to produce density estimates across the whole age range, not just at the scheduled measurement dates, and it has been found to perform better than other methods which are used for this kind of work, such as quantile regression. Quantile regression is a non-parametric method which produces centiles based on the empirical data rather than producing a probability distribution. While its use has been shown to produce accurate and useful results, it can lead to overlapping centiles, implying negative probability (Rigby et al. 2013). GAMLSS on the other hand is a parametric method, producing density estimates where the whole shape of the distribution can vary according to the explanatory variables.

To summarise, the GAMLSS framework is a flexible way to estimate centile curves by assuming either the Box-Cox Cole and Green or Box-Cox power exponential distribution. Its use in the field of paediatrics has been shown to be valuable for centile estimation. It is recommended for researchers to quantify distributions which change smoothly over time. The charts we developed are locally representative internal references which describe how local children grow. Children's measurements can be plotted on these charts allowing practitioners to assess child growth relative to the distributions and can be used to diagnose undernutrition. Ideally children should be compared with references which describe the growth of healthy children in optimal conditions, such as those who make up the WHO standard. There are, of course, pros to comparing children with locally based references - not all children grow at the same rate under suboptimal conditions, and conditions that children live in are variable from dataset to dataset. Therefore, comparing children with those who live in the same geographical location provides a local reference allowing relative comparisons. However, justification in using an international standard itself lies in the fact that local references are variable - comparisons cannot be made across datasets unless one single reference is chosen as a baseline. Furthermore, children should not be compared with others that live in impoverished conditions as no real inference can be made to whether that child is growing at an adequate rate. The only inference that can be made is that they grow well relative to unhealthy counterparts

whose trajectory, on average, is not optimal growth. We therefore used the WHO standard to assess the relative growth of the children within our database.

7.2.4 Comparing real populations of pre-school children with the WHO standard

The WHO standard, published in 2006, was designed to quantify optimal growth. The standard was developed from diverse populations of healthy, breast fed infants living in optimal circumstances, including children from six different countries: Brazil, Ghana, India, Norway, Oman and USA (de Onis & WHO Multicentre Growth Reference Study Group 2006). We aimed to determine how populations of pre-school children from the developing world fit this standard using the database we developed.

Since the WHO charts represent optimal rather than average growth and were designed to be used in a wide range of settings, variations in fit were expected. Roelants (2013), Hui et al. (2008) and Wright et al. (2008) identified that children from the developed world tend to grow on par or outgrow the WHO standard in terms of weight. They did this by converting measurements to Z scores relative to the standard, presenting the mean Z scores. This approach was not found to have been applied to any developing world datasets, although many studies have published statistics on proportions of children below certain threshold values relative to the standard. We used GAMLSS to model the mean and standard deviation of weight-for-age Z scores within our datasets relative to the WHO standard. We found that children from the most affluent areas tend to grow on par with the standard whereas children from the more deprived areas tend to track along a low centile after around 6 months, or 12 months for those in Malawi. As the WHO standard reflects optimal growth, this implies that children living in affluent areas also live under optimal conditions, provided, on average, with adequate nutrition. As those living in more deprived area tend to track along low centiles, this implies these children lack adequate nutrition.

In the Pakistani cohort, growth is relatively slow with children falling away from the standard during months 0-3. However, they exhibit catch-up and track the standard from 6 months onwards, with children from the more deprived cohorts tracking at lower centiles. This implies that the less affluent the area, the less chance there is that children will meet their nutritional needs. Growth is also relatively slow for those in the Malawi cohort as children consistently fall from 0.5SD to 1.5SD by month 12, on average. These children then also tend track along a low centile. These children track above the Pakistani

village and peri-urban slum cohorts, but below the urban slum and middle class cohorts. This implies these children are receiving, on average, more nourishment than those in the village and peri-urban slum but again, not enough to keep up with the international standard. Children in South Africa are lighter than the WHO standard at birth, but quickly catch up. Notably, females outgrow the standard from month 6 onwards. This implies these children, on average, at least meet their nutritional needs. However, it must be noted that other factors influence growth other than nourishment, such as environment. A problem with this analysis was the effect of HIV status, which was not initially adjusted for. It was not until the end of the research period that we identified that HIV status (for the South African cohort) and maternal HIV status (for the Malawian cohort) could have been taken into account as confounding variables, by which time it was too late to re-analyse the data. Poor growth is a sensitive indicator of disease progression in children, and in HIV infected children failure to thrive is reported in 20-70% (Hirschfeld 1996). As HIV status is available for the South African cohort, an interesting extension to this piece of work would be to split the dataset into categories of children with and without HIV to observe the difference in growth patterns. Inclusion of this variable would result in two separate curves, which would most likely result in one growth curve outperforming the standard and one tracking at a low centile. However, this is speculation and it cannot be concluded without further analysis. Furthermore, prospective cohort studies in developing countries indicate that intrauterine growth may also be affected by maternal HIV (Lepage et al. 1996), although studies in western countries are less consistent where maternal factors are more of a problem (Selwyn et al. 1989) indicating maternal HIV may not have an independent effect on intrauterine growth and birth weight. As maternal HIV data were available in both South Africa and Malawi, splitting these children into groups by maternal HIV status and comparing birthweights would be an interesting piece of work. However, maternal confounding variables would also need to be taken into account (Arpadi 2005).

7.2.5 Generalising conditional weight gain

Our third aim was to generalise the expression for conditional weight gain for use with external reference distributions. The original aim had been: *"Describe what impact using an external reference has on some common measures of weight gain and investigate how to modify the measures in order to allow for this".* Conditional weight gain assesses the

weight change of a child, taking into account the fact that children's weights are expected to regress towards the mean. Weight gain was of interest because latter aims were established to investigate whether growth over a recent time period is more valuable than recent size. We aimed to use conditional weight gain to identify this within statistical models, comparing its value against size. The logic behind modifying the measure of weight gain was that the original methodology, which defines the change in weight Z score as a velocity Z score, a function of the two Z scores and the correlation between them, assumed that a locally based internal reference distribution is used in the calculation. Furthermore comparisons cannot be made directly across populations since the baseline reference distribution changes for each distinct dataset.

We generalised the expression for conditional weight gain to apply to an external reference, in particular the WHO standard, with the intention of using it within prediction models to assess significance.

Generally, to assess child growth over time, serial measurements are plotted on the charts developed by the LMS method in Chapter 3, such as the WHO standard charts. Growth velocity can also be calculated by dividing the change in measurement over time by the corresponding time interval. If children's measurements deviate from a previously established growth trajectory they indicate growth faltering. However, as discussed, serial measurements also regress to the mean. To account for this expected movement, conditional measures of velocity can be used. Cameron et al. (1980) and Berkey et al. (1983) built models to predict future size based on previous size on the raw measurement scale whereas Cole (1994), Cole (1995) and Cole (1997), aimed to provide a simple method that would allow practitioners to assess centile change between two time points t_1 and t_2 on the SD scale, called conditional weight gain. With an internal reference the means and standard deviations of the two Z scores are 0 and 1 respectively, and the expression simplifies accordingly. We adapted the expression developed by Cole for use with an external reference, which requires the means and standard deviations to be estimated.

The expressions for conditional weight gain and generalised conditional weight gain both include an adjustment for RTTM, related to their correlation (Cole 1995). We estimated the correlation between Z scores at t_1 and t_2. Correlations can be calculated between groups of measurements at scheduled measurement dates as data are available, providing correlation coefficients for all combinations of ages within the vector of

scheduled measurement dates. However, statistical models are needed to interpolate correlations between pairs of ages for which no data are available. We modelled correlation as a function of the initial age t_1 and latter age t_2, allowing any t_1/t_2 combination to be obtained. We adopted 4 modelling approaches: a fractional polynomial model developed by Cole (Cole 1995), the Argyle model (Argyle et al. 2008), fractional polynomials (Royston & Altman 1994) and GAMs (Wood 2006). The Argyle model used the log transformation and the others used the Fisher transformation to stabilise the variance. To assess fit, we used the AIC, BIC, RMS and the coefficient of determination. The AIC and BIC had to be adapted since the models were on different scales, using the Jacobian adjustment described by Box & Cox (1964).

Argyle (2008) and Roelants (2013) both state that the Argyle model is not always a satisfactory fit. However, the model was developed to be a simple general model, which we also attempted by combining unique terms within the fractional polynomial models. However this model did not perform very well on some of the datasets and was abandoned. We found that our fractional polynomial model provided the best fit, performing best in the village, urban slum, middle class area and South Africa.

GAMs performed best on the Malawi and peri-urban slum matrices. In retrospect, there are a number of models which could have been applied which might have provided a better fit such as the extended AR(1) model presented by Wade & Ades (1998), kernel based smoothing methods such as local linear regression (Bowman & Azzalini 1997) or machine learning methods such as Gaussian Processes (Mackay 1998). However, these methods are more complex to implement and interpret, and may only slightly outperform fractional polynomials. We therefore recommend using fractional polynomials to model data of this type. Generalised models are not advised as correlation structures vary between populations.

In summary, the estimated correlation coefficient coupled with estimates of the mean and standard deviations of the Z scores allow us to calculate generalised conditional weight gain. This expression is recommended for use in this field, allowing researchers to compute weight gain with an external reference such as the WHO standard while taking RTTM into account. The resulting Z score describes growth relative to that reference, and the expression simplifies to conditional weight gain if an internal reference is used. One major issue with using the measure within the developing world is that because growth conditions are so variable, a single correlation model is unlikely to generalise to other

populations. This makes the use of the measure impractical. However, the measure may be of more use within the developed world, where correlation matrices may be less variable.

7.3 Testing hypotheses

The remainder of this discussion focuses on assessing how anthropometry can be used to predict adverse future outcomes. We also summarise how these aims were developed. In fact, the original study aim for this thesis was to investigate whether growth over a recent time period is a better predictor of adverse outcomes than recent size. We later added an investigation of the relationships between 'nutritional states' as described in Section 1.2.7, defined as: healthy, wasted, stunted, wasted-stunted and dead. This investigation was added since literature investigating the original study aim highlighted that not only are size and growth useful predictors of adverse future outcomes, but nutritional states are also useful predictors. The literature review also highlighted that the pathways children take through nutritional states are poorly understood. Furthermore, as we started developing models for our original study aim, we recognised that our findings would not add as much to the already existing scientific literature as we initially thought. Finally, as more models were developed for the original study aim, we identified that due to the low mortality rates within some populations, we needed a way to justify pooling datasets to increase power. Investigating the relationships between nutritional states was carried out after work on Chapter 6, but was presented within Chapter 5 as some of it provided justification for pooling datasets for use in Chapter 6.

7.3.1 Exploring the pathways children take through nutritional states over time

To investigate the pathways children take through nutritional states over time, we needed a framework to assess the probability that children move from one state to another. Furthermore, this framework had to control the initial and final ages that we wished to assess since trends were evidently dependent on age. Previous work in this field did not control for this.

The methodology we chose to investigate the pathways was stochastic in nature, calculating the conditional probability of children moving from one nutritional state to another, presented as a stochastic matrix, representing the probability of all possible transitions. We developed 3 hypotheses by investigating research in this area. Within the

next sections, we discuss each of these hypotheses and the papers that both support and oppose them. We discuss our findings and compare them with previous work. Finally, we discuss the methodology's strengths, limitations and possible extensions for the future.

7.3.1.1 Which states are most likely to lead to death?

Firstly, prior research suggested that wasted-stunted children were at the highest risk of death compared with wasting or stunting alone. Our first hypothesis stated:

> *Hypothesis 1: those who are both wasted and stunted are at a higher risk of death than those who are wasted, who are in turn at a higher risk than those who are stunted.*

We found that in the 3-6m and 9-12m periods, wasted-stunted, wasted and stunted children were all significantly more likely to die than healthy children. Wasted-stunted children were at the highest risk of death in the 3-9m and 6-9m timeframes and wasted children were at the highest risk of death in the 9-12m timeframe. A notable result is that the RR of death was lower across all transitions within the 6-9m timeframe, indicating children in adverse states are less vulnerable between these ages, relative to healthy children. We also found that those in the middle class cohort were much more likely to recover from undesirable states, indicating that socio-economic status may be a driver of recovery.

In previous analyses by Olofin et al. (2010) and McDonald et al. (2013), children were categorised into a baseline state based on their most recent anthropometric measurement, where the age of categorisation varied per child. Our analysis adds value to this research area by allowing control of the effect of age on mortality. Effect sizes between these studies and ours are not directly comparable since baseline states differ between children. Furthermore, the analyses by Olofin and McDonald were conducted on data in which children were aged 0-5 years, opposed to our analyses which analysed data of children up to 2 years old. However, on the whole, our results agree in terms of which nutritional states are the most valuable predictors. Notably, multiple anthropometric deficits are more predictive than single deficits, and wasting is a better predictor of mortality than stunting.

7.3.1.2 Does wasting precede stunting?

Secondly, prior research suggested that wasting tends to precede stunting. Our second hypothesis stated:

Hypothesis 2: wasting tends to precede stunting.

We found that wasting significantly increased the risk of stunting compared with healthy children only in the 3-6m period (RR = 1.84). Wasting did not significantly increase the risk of stunting in periods 6-9m and 9-12m, again, compared with healthy individuals. This implies that wasting does not always precede stunting, as healthy children are just as likely as wasted children to become stunted.

Since prior research indicated that wasting tends to precede stunting, we expected that initially healthy children would be more likely to become wasted than stunted. What was surprising was that in all three time frames, the reverse was true – healthy children were more likely to become stunted than wasted.

We also found wasted children were at significantly higher risk of later wasting compared with healthy children, and that wasting significantly increased the risk of becoming both stunted and wasted.

Our research clarifies how wasting and stunting develop. Prior research utilised linear models and calculated correlation coefficients to quantify group relationships to assess whether measures of thinness and/or shortness are related. Our analysis, using a stochastic approach, allowed us to model the conditional probability of individuals moving from state to state while controlling for initial and final time points. To summarise, wasting can lead directly to stunting, but children are only at a higher risk in very early life. Wasting does, however, lead to later a wasted-stunted state.

7.3.1.3 Is stunting irreversible?

Finally, prior research suggested that stunting is largely irreversible after 2 years. Hypothesis 3 stated:

Hypothesis 3: children who become stunted are likely to stay stunted.

We found that stunted children were at a significantly higher risk of stunting in the future compared with healthy children in all timeframes. We also found that as the age that the probability of transition is calculated over shifts to higher ages, the probability of remaining stunted increases as the probability of recovery decreases, indicating that as children get older, the probability that they stay stunted increases. Children who were from the more affluent datasets were more likely to recover from stunting than their poorer counterparts, indicating socioeconomic status may play a part in recovery. Furthermore, children who were stunted were at no more risk of later wasting compared

to healthy children. In fact, very few children made the transition from stunting to wasting. Finally, stunted children were at much higher risk of later wasting and stunting compared to those who are healthy. Our analysis, however, did not take genetics into account as naturally small children are not necessarily malnourished. Hence the results with respect to stunting are not conclusive. However, we would expect the growth trajectory of naturally small children to parallel a reference distribution rather than to fall away from it, which would allow us to discriminate between the two.

7.3.1.4 Discussion of methodology

The stochastic modelling approach we adopted, coupled with the use of relative risks, proved to be very useful and is recommended for researchers working in this field.
It allowed us to model the conditional probability of moving between states whereas other methods which have been employed in this field simply aim to quantify average group relationships. The methodology also allowed us to control for the initial baseline state, and in effect, allowed us to control the dependence of adverse outcomes on age. Furthermore, transitions from one state to another could be compared with the proportion of healthy children moving to the same state, allowing us to calculate the increased risk of moving between states given an initial adverse state. The method does have its drawbacks, notably that there need to be enough children to compute the transition probabilities reliably.
This approach could be extended further. Other variables could be introduced into the framework by applying a GLM, specifying a multinomial family and logit link function for the mean. Each child could be categorised into one of the 5 nutritional states at each time point. The probabilities of moving state using the previous state and area as covariates could be used to estimate probabilities for the transition matrices. This method would allow researchers to include any extra variables within the models such as HIV status, family income or socio-economic status as discrete or continuous covariates. Since the transition probabilities are time dependent, the number of transition times could be extended so that a grid of probabilities for all combinations of times could be observed, for every set of transitions. By focusing on one transition, say wasting to death, the probability of children moving between all combinations of possible time points could be observed. This was investigated briefly at the end of the research period and was very informative, with clear trends over the 3D surface, highlighting that the probability of a

child moving from state to state was highly dependent on initial and final age. For example, the 3D surface showed that the higher the age of initial wasting, the lower the probability of death. This is what would be expected since younger children are intrinsically more vulnerable. This analysis could be carried out for any transition, not just the example mentioned above. The resulting probabilities could be plotted and modelled using similar techniques employed within the correlation modelling section in Chapter 4. Unfortunately, due to time constraints, the research could not be taken any further. In summary, the stochastic approach was useful, allowing us to control the baseline and final ages of the children when calculating transition probabilities. This method of analysis has not been applied in any research within this area but is recommended for use along with the extensions mentioned above.

7.3.2 Could our datasets be pooled?

Pooling data is appropriate when data are homogeneous in the parameters of interest. We developed a statistical test in Chapter 2 to identify whether datasets could be pooled based on the pathways children took through nutritional states, which were summarised in the previous section. This is more appropriate than simply pooling data in terms of say mortality rates or rates of undernutrition since it is the way that children die that we wish to quantify and analyse, not the number who die.

We developed a GLRT to assess whether probability transition matrices, which quantified the pathways in which children moved through states, could be pooled. No reference to a similar test could be found within the literature, only a “one sample” test was found, by Bickenback & Bode (2001), which was applied to financial data. This “two sample” test we developed is a future resource, not only for those who work in paediatrics, but for any researchers working with probability transition matrices. Our results indicated that children could not be pooled into a single group, but could be pooled into 4 sub-groups. The four sub-groups were: South Africa/middle class, village/per-urban slum, Malawi and the urban slum. Mortality rates were low in some of the datasets, and pooling the datasets increased statistical power.

7.3.3 Is growth over a recent time period a better predictor of mortality than most recent size?

Initially we planned on using generalised conditional weight gain as a predictor within survival models to assess growth as a predictor of mortality. However, in terms of

predictive power, using the combination of initial weight-for-age Z score and generalised conditional weight gain as predictors in a survival model is algebraically equivalent to including the Z scores at t_1 and t_2, i.e. Z_1 plus Z_2 provide the same information as Z_1 plus conditional weight gain. Dropping the initial term from the models allow us to compare growth over a recent time period with most recent size. Obviously this loses the benefit of working with conditional weight gain, but the predictive ability is the same.
Investigation of the literature in this area provided us with a final hypothesis:

> *Hypothesis 4: recent size is a more valuable predictor of mortality than growth over a recent time period.*

We assumed that the majority of papers were correct and hypothesised that size is a more valuable predictor of mortality than growth over a recent time period. We tested this hypothesis using Cox proportional hazards models, applied to the four pooled datasets resulting from the analysis within Chapter 5. In fact, without pooling the datasets in this way, the analysis for the middle class area would have been unreliable due to the low mortality rate.
Initially, the standard Cox proportional hazards models we applied were dominated by children within the centre of the population distribution. We therefore adopted a weighted approach, allowing us to over-represent children who are outwith the normal range so that significant predictors could be determined for those at the tails of the distribution. These children are the ones who are at the highest risk of mortality and the analysis design allowed us to control the degree of representation.
We developed a weighting function (Equation (6.1)) which was used to increase the weighting a child received. The lower their average weight-for-age Z score over the 3 and 6 month periods, the higher the weighting, where only those with negative Z scores received the weighting function. This approach strengthened our analysis as children within the tails were over-represented, hence, had greater influence on our statistical models. The weighting function was developed to be variable based on a power parameter i.e. the weighting that children within the extremes received was controlled, providing us with a valuable framework to evaluate the effect of the degree of weighting. This allowed us to investigate the effects of increasing the effective number of children within the analyses that lay in the extremes, which allowed us to identify the magnitude of weighting needed for growth to enter as an important predictor.

In the unweighted analysis predicting time until death after both 3 and 6 months it was clear that final size stood as the best predictor.

In the weighted analyses predicting after 3 months, we found that effect of weighting on the best fitting model differed by dataset. The less affluent the dataset (and the higher the prevalence of undernutrition), weight gain was less likely to enter as a significant predictor as the weightings increased. We hypothesise that the majority of children within the less affluent datasets tend to fall away from the WHO median, hence the majority are at risk. Size tends to contain most information about current state of the children and over-representing the children who are at higher risk does not change the form of the model as the majority are already at risk, and these children fall from the WHO median.

In more affluent populations, the majority of children are within the normal healthy range, so size is the best predictor in the unweighted analyses as the analysis is dominated by these children. However, fewer fall away, so that those who are a genetically small make up a larger proportion of all small children, so that over-representing those in the extremes results in the model identifying growth as a predictor. The analysis is no longer dominated by children within the normal, healthy range. Growth serves as a discriminator between those who fall, those who are naturally small and those who are tracking.

In conclusion, while predicting after 3 months, we found growth adds no more value over size for the majority of children. For other measures to enter as significant predictors, the weightings applied were so high that the datasets had more artificially replicated rows than there were real data. However this does imply its usefulness in assessing small children in more affluent populations. While predicting after 6 months, we again found that size stood as the best predictor, indicating that for the majority of children, this measure will provide enough information to identify those at risk. With higher weightings, height-for-age entered our models as well as weight-for-age, indicating that measures of height-for-age may be valuable while making predictions for children who are slightly older. This is in contrast to our analysis predicting after 3 months, as height-for-age did not enter our best fitting models at any weighting.

A limitation of this analysis was, again, the effect of HIV status, which we did not account for. Newell et al. (2004) estimated that children testing positively for HIV before 4 weeks of age had around a 50% chance of dying by age 2 years. The post-hoc analysis within

Chapter 6 indicated that stratifying by HIV status impacted the South African models. When including maternal HIV status within the Malawi analysis, models stayed relatively similar. This indicates that HIV status of the mother should not be used as a proxy for HIV status of the child. The results of the post-hoc analysis indicate that if available, HIV status should be included as a stratified variable. An interesting piece of further work would be to include HIV status as a time-varying covariate, since we expect the probability of survival to be a function of the time that HIV is contracted (Miller 2011).

In summary, the use of the weighting function increased the amount of control we had over our statistical analysis and it is recommended for those working in paediatrics. Using this approach, we were able to identify children who benefit from the extra value of growth in predicting mortality. Our findings, therefore, agree in part with research which has concluded that size is the best predictor, and in part with research that has concluded that growth is the best predictor. In high malnutrition prevalence populations, a large proportion of children fall away from the WHO median, making growth pattern non-discriminating.

7.4 Concluding remark

To conclude this final chapter, and the dissertation, we summarise our work in a final paragraph. The analyses we carried out within the programme of work were multidimensional and time dependant, allowing us to demonstrate different methods that can be used to handle data of this nature. We created a database for further research, developed six developing world growth reference charts, described how children from the developing world fit the WHO standard, generalised the measure of conditional weight gain and gained new insights into how measures of undernutrition can be used to predict adverse future outcomes. The inability to carry out analyses on the full pooled dataset provided us with information about the diversity and heterogeneity between datasets, highlighting that pooling data of this kind is not always appropriate. The methodology used throughout the thesis is recommended, as are the extensions suggested in this final discussion chapter.

Bibliography

Adair, L. S. (1999). Community and International Nutrition Filipino Children Exhibit Catch-Up Growth from Age 2 to 12 Years. *Children*, (February), 1140–1148.

Adair, L. S. (1999). Filipino children exhibit catch-Up growth from age 2 to 12 years. *The Journal of Nutrition*, *129*(February), 1140–1148.

Africa Centre. (2007). Africa Centre Vertical Transmission Study. Retrieved February 23, 2015, from http://www.africacentre.ac.za/Default.aspx?tabid=347

Akaike, H. (1974). "A new look at the statistical model identification." *IEEE Transactions on Automatic Control*, *19*(6), 716–723.

Akantziliotou, C., Rigby, R., & Stasinopoulos, D. (2002). "The R implementation of generalized additive models for location, scale and shape." *Statistical Modelling in Society: Proceedings of the 17th International Workshop on Statistical Modelling. Statistical Modelling Society*, 75–83.

Albert, R. (2009). *The Merck manual home health handbook.* (B. Porter, RS; Kaplan, JL; Homeier, Ed.). Merck & Company.

Alderman, M. H., Wise, P. H., Ferguson, R. P., Laverde, H. T., & D'Souza, a J. (1978). Reduction of young child malnutrition and mortality in rural Jamaica. *The Journal of Tropical Pediatrics and Environmental Child Health*, *24*(1), 7–11.

Ambler, G. (2013). Package "mfp" Retrieved 05/12/13. from R-project: http://cran.r-project.org/web/packages/mfp/mfp.pdfle.

Ambler, G., & Royston, P. (2001). Fractional polynomial model selection procedures: investigation of type i error rate. *Journal of Statistical Computation and Simulation*, *69*(1), 89–108. http://doi.org/10.1080/00949650108812083

Argyle, J. (2002). Statistical analysis of child growth data. *Retrieved from Durham E-Theses Online: http//etheses.dur.ac.uk/4113/*.

Argyle, J., Seheult, A. H., & Wooff, D. A. (2008). Correlation models for monitoring child growth. *Statistics in Medicine*, *27*(6), 888–904. http://doi.org/10.1002/sim.2973

Arifeen, S. E., Black, R. E., Caulfield, L. E., Antelman, G., & Baqui, a H. (2001). Determinants of infant growth in the slums of Dhaka: size and maturity at birth, breastfeeding and morbidity. *European Journal of Clinical Nutrition*, *55*(3), 167–178.

Arpadi, S. (2005). Growth failure in HIV-infected children. *Consultation on Nutrition and HIV/AIDS in Africa: …*, (April), 1–20.

Ashorn, P., Maleta, K., Espo, M., & Kulmala, T. (2002). Male biased mortality among 1-2 year old children in rural Malawi. *Archives of Disease in Childhood*, *87*(5), 386–387.

Ashraf, R. N., Jalil, F., Khan, S. R., Zaman, S., Karlberg, J., Lindblad, B. S., & Hanson, L. a. (1993). Early child health in Lahore, Pakistan: V. Feeding patterns. *Acta Paediatrica (Oslo, Norway : 1992). Supplement*, *82 Suppl 3*, 47–61.

Ashworth, A., & Millward, D. (1976). Catch-up growth in children. *Nutrition Reviews*, *44*(5), 157–163.

Ashworth, A., Shrimpton, R., & Jamil, K. (2008). Growth monitoring and promotion: review of evidence of impact. *Maternal & Child Nutrition*, *4 Suppl 1*, 86–117.

Babu, D. S., & Chuttani, C. S. (1979). Anthropometric indices independent of age for nutritional assessment in schoolchildren. *Journal of Epidemiology and Community Health*, *33*(33), 177–179.

Bairagi, R., & Chowdhury, M. K. (1994). Socioeconomic and anthropometric status, and mortality of young children in rural Bangladesh. *International Journal of Epidemiology*, *23*(6), 1179–1184.

Bairagi, R., Chowdhury, M. K., Kim, Y. J., & Curlin, G. T. (1985). Alternative anthropometric indicators of mortality. *American Journal of Clinical Nutrition*, *42*(2), 296–306.

Bairagi, R., Koenig, M. A., & Mazumder, K. A. (1993). Mortality-discriminating power of some nutritional, sociodemographic, and diarrheal disease indices. *American Journal of Epidemiology*, *138*(5), 310–317.

Barker, D. J. P. (2001). The malnourished baby and infant. *British Medical Bulletin*, *60*, 69–88.

Beaton, G., Kelly, A., Kevany, J., Martorell, R., & Mason, J. (1990). Appropriate Uses of Anthropometric Indices in Children. *ACC/SCN State-of-the-Art Series, Nutrition Policy Discussion Paper No. 7. ACC/SCN, Geneva.*

Beaumont Health System. (2015). How much will my child grow? Retrieved December 4, 2015, from http://www.beaumont.edu/childrens/health-and-safety/the-growing-child/school-age-6-to-12-years/

Becquet, R., Marston, M., Dabis, F., Moulton, L. H., Gray, G., Coovadia, H. M., … Newell, M. L. (2012). Children who acquire hiv infection perinatally are at higher risk of early death than those acquiring infection through breastmilk: A meta-analysis. *PLoS ONE*, *7*(2).

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

Bickenbach, F., & Bode, E. (2001). Markov or Not Markov? This Should Be a Question. *Regional Science and Urban Economics*, *29*, 257–281.

Black, R. E., Allen, L. H., Bhutta, Z. A., Caulfield, L. E., de Onis, M., Ezzati, M., Rivera, J. (2008a). Maternal and child undernutrition: global and regional exposures and health consequences. *The Lancet*, *371*(9608), 243–260.

Black, R. E., Allen, L. H., Bhutta, Z. A., Caulfield, L. E., de Onis, M., Ezzati, M., … Rivera, J. (2008b). Maternal and child undernutrition: global and regional exposures and health consequences. *The Lancet*.

Black, R. E., Victora, C. G., Walker, S. P., Bhutta, Z. a., Christian, P., De Onis, M., … Uauy, R. (2013). Maternal and child undernutrition and overweight in low-income and middle-income countries. *The Lancet*, *382*(9890), 427–451.

Blair, P. S., Nadin, P., Cole, T. J., Fleming, P. J., Smith, I. J., Platt, M. W., … Golding, J. (2000). Weight gain and sudden infant death syndrome: changes in weight z scores may identify infants at increased risk. *Archives of Disease in Childhood*, *82*(6), 462–9.

Bland, R. M., Coovadia, H. M., Coutsoudis, a., Rollins, N. C., & Newell, M. L. (2010). Cohort profile: Mamanengane or the Africa centre vertical transmission study. *International Journal of Epidemiology*, *39*(2), 351–360.

Bogin, B. (1978). Seasonal pattern in the rate of growth in height of children living in Guatemala. *American Journal of Physical Anthropology*, *49*(2), 205–210.

Bosch, A. M., Baqui, A. H., & van Ginneken, J. K. (2008). Early-life determinants of stunted adolescent girls and boys in Matlab, Bangladesh. *Journal of Health, Population and Nutrition*, *26*(2), 189–199.

Botton, J., Heude, B., Maccario, J., Ducimetière, P., & Charles, M.-A. (2008). Postnatal weight and height growth velocities at different ages between birth and 5 y and body composition in adolescent boys and girls. *American Journal of Clinical Nutrition*, *87*, 1760–8.

Bowman, A. W., & Azzalini, A. (1997). *Applied Smoothing Techniques for Data Analysis*. Oxford University Press.

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, (ii), 211–252.

Briend, A., & Bari, A. (1989). Critical assessment of the use of growth monitoring for identifying high risk children in primary health care programmes. *British Medical Journal*, *298*, 1607–1611.

Briend, A., Khara, T., & Dolan, C. (2015). Wasting and stunting--similarities and differences: policy and programmatic implications. *Food and Nutrition Bulletin*, *36*(1 Suppl), S15–23.

Brown, K., Black, R., & Becker, S. (1982). Seasonal changes in nutritional status and the prevalence of malnutrition in a longitudinal study of young children in rural Bangladesh. *The American Journal of Clinical Nutrition*, *36*(2), 303–313.

Brownie, C., Habicht, J. P., & Cogill, B. (1986). Comparing indicators of health or nutritional status. *American Journal of Epidemiology*, *124*(6), 1031–44.

Burgess, A. (2008). Mother and Child Undernutrition - Vitamin A Deficiency. *South Sudan Medical Journal*.

Burgess, A., Danga, L. (2008). Undernutrition in Adults and Children: causes, consequences and what we can do. Retrieved from http://www.southsudanmedicaljournal.com/archive/2008-05/undernutrition-in-adults-and-children-causes-consequences-and-what-we-can-do.html

Canadian Paediatric Society. (2004). A health professional's guide to using growth charts. *Paediatrics & Child Health.*, *9*(3), 174–176.

Carruth, B. R., Ziegler, P. J., Gordon, A., & Barr, S. I. (2004). Prevalence of picky eaters among infants and toddlers and their caregivers' decisions about offering a new food. *Journal of the American Dietetic Association*, *104*(suppl. 1), 57–64.

CDC. (2013). Sudden Infant Death. Retrieved from http://www.cdc.gov/sids/

Cheung, Y. B., Yip, P. S., & Karlberg, J. P. (2001). Fetal growth, early postnatal growth and motor development in Pakistani infants. *Int.J.Epidemiol.*, *30*(0300-5771 (Print) LA - eng PT - Journal Article PT - Research Support, Non-U.S. Gov't SB - IM), 66–72.

Cheung, Y., Yip, P., & Karlberg, J. (2001). Fetal growth, early postnatal growth and motor development in Pakistani infants. *International Journal of Epidemiology*.

Cole, T. J. (1985). A critique of the NCHS Weight for Height Standard. *Human Biology*, *57*(2), 183–209.

Cole, T. J. (1986). Weight/height^p compared to weight/height^2 for assessing adiposity in childhood: influence of age and bone age on p during puberty. *Annals of Human Biology*, *13*(5), 433–51.

Cole, T. J. (1989). Relating growth rate to environmental factors--methodological problems in the study of growth-infection interaction. *Acta Paediatr.Scand.Suppl*, *350* (0300-8843 LA - eng PT - Journal Article SB - IM), 14–20.

Cole, T. J. (1990). The LMS method for constructing normalized growth standards. *European Journal of Clinical Nutrition*, *44*(1), 45–60.

Cole, T. J. (1993). The use and construction of anthropometric growth reference standards. *Nutrition Research Reviews*, *6*(1), 19–50.

Cole, T. J. (1994). Growth charts for both cross-sectional and longitudinal data. *Statistics in Medicine*, *13*(23-24), 2477–92.

Cole, T. J. (1995). Conditional reference charts to assess weight gain in British infants. *Arch Dis Child*, *73*(1), 8–16.

Cole, T. J. (1997). Growth monitoring with the British 1990 growth reference. *Archives of Disease in Childhood*, *76*(1), 47–9.

Cole, T. J. (2002). A chart to link child centiles of body mass index, weight and height. *European Journal of Clinical Nutrition*, *56*, 1194–1199.

Cole, T. J. (2002). Assessment of growth. *Best Practice & Research. Clinical Endocrinology & Metabolism*, *16*(3), 383–398.

Cole, T. J., Donnet, M. L., & Stanfield, J. P. (1981). Weight-for-height indices to assess nutritional status: A new index on a slide-rule. *American Journal of Clinical Nutrition*, *34*(9), 1935–1943.

Cole, T. J., Freeman, J. V, & Preece, M. A. (1995). Body mass index reference curves for the UK, 1990. *Archives of Disease in Childhood*, *73*(1), 25–9.

Cole, T. J., Freeman, J. V., & Preece, M. a. (1998). British 1990 growth reference centiles for weight, height, body mass index and head circumference fitted by maximum penalized likelihood. *Statistics in Medicine*, *17*(4), 407–429.

Cole, T. J., & Green, P. J. (1992). Smoothing reference centile curves: the LMS method and penalized likelihood. *Statistics in Medicine*, *11*(10), 1305–19.

Costello, A. M. (1989). Growth velocity and stunting in rural Nepal. *Archives of Disease in Childhood*, *64*(10), 1478–1482.

Cox, D. R. (1972). Regression models and life tables. *Journal of the Royal Statistical Society. Series B:*, *34*(2), 187–220.

Craven, P., & Wahba, G. (1979). Smoothing noisy data with spline functions - Estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik*, *31*, 377–403.

Cunningham, N. (1978). The under-fives clinic--what difference does it make? *The Journal of Tropical Pediatrics and Environmental Child Health*, *24*(6), 239–334.

Dale, N. M., Grais, R. F., Minetti, A., Miettola, J., & Barengo, N. C. (2009). Comparison of the new World Health Organization growth standards and the National Center for Health Statistics growth reference regarding mortality of malnourished children treated in a 2006 nutrition program in Niger.tle. *Archives of Pediatrics and Adolescent Medicine*, *163*(2), 126–130.

de Onis, M., Garza, C., Onyango, A. W., & Borghi, E. (2007). Comparison of the WHO child growth standards and the CDC 2000 growth charts. *The Journal of Nutrition*, *137*(1), 144–148.

de Onis, M., Garza, C., Victora, C. G., Onyango, A. W., Frongillo, E. a., & Martines, J. (2004). The WHO Multicentre Growth Reference Study: Planning, study design, and methodology. *Food and Nutrition Bulletin*, *25*(1 SUPPL. 1), 1–18.

de Onis, M., Onyango, A., Van den Broeck, J., Chumlea, W., & Martorell, R. (2006). Measurement and standardization protocols for anthropometry used in the construction of a new international growth reference. *Food and Nutrition Bulletin*, *25*(1), S27–S36.

de Onis, M., & WHO Multicentre Growth Reference Study Group. (2006). WHO Child Growth Standards based on length/height, weight and age. *Acta Paediatr Suppl*, *450*, 76–85.

de Onis, M., Wijnhoven, T. M. a, & Onyango, A. W. (2004). Worldwide practices in child growth monitoring. *Journal of Pediatrics*, *144*, 461–465.

Doherty, C. P., Sarkar, M. a, Shakur, M. S., Ling, S. C., Elton, R. a, & Cutting, W. a. (2001). *Linear and knemometric growth in the early phase of rehabilitation from severe malnutrition. The British journal of nutrition* (Vol. 85).

Drewett, R., Wolke, D., Asefa, M., Kaba, M., & Tessema, F. (2001). Malnutrition and mental development: is there a sensitive period? A nested case-control study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *42*(2), 181–7.

Eilers, P. H. C., & Marx, B. D. (1996). Flexible smoothing with B-splines and penalties. *Statistical Science*, *11*(2), 89–121.

Ekelund, U., Ong, K., Linné, Y., Neovius, M., Brage, S., Dunger, D. B., … Rössner, S. (2006). Upward weight percentile crossing in infancy and early childhood independently predicts fat mass in young adults: The Stockholm Weight Development Study (SWEDES). *American Journal of Clinical Nutrition*, *83*(2), 324–330.

Emond, A. M., Blair, P. S., Emmett, P. M., & Drewett, R. F. (2007). Weight faltering in infancy and IQ levels at 8 years in the Avon Longitudinal Study of Parents and Children. *Pediatrics*, *120*(4), e1051–8.

Espo, M., Kulmala, T., Maleta, K., Cullinan, T., Salin, M. L., & Ashorn, P. (2002). Determinants of linear growth and predictors of severe stunting during infancy in rural Malawi. *Acta Paediatrica*, *91*(12), 1364–1370.

Fawzi, W. W., Herrera, M. G., Spiegelman, D. L., el Amin, A., Nestel, P., & Mohamed, K. A. (1997). A prospective study of malnutrition in relation to child mortality in the Sudan. *The American Journal of Clinical Nutrition*, *65*, 1062–1069.

Fisher, R. A. (1921). On the “Probable Error” of a Coefficient of Correlation Deduced from a Small Sample. *Metron*, *1*(4), 3–32.

Flegal, K. M., & Cole, T. J. (2013). Construction of LMS parameters for the Centers for Disease Control and Prevention 2000 growth charts. *National Health Statistics Reports*, *9*(63), 1–3.

Food And Agriculture Organization Of The United Nations. (2015). *The state of food insecurity in the world 2016: strengthening the enabling environment for food…security and nutrition*. Food and Agriculture Org.

Freeman, J. V, Cole, T. J., Chinn, S., Jones, P. R., White, E. M., & Preece, M. a. (1995). Cross sectional stature and weight reference curves for the UK, 1990. *Archives of Disease in Childhood*, *73*(1), 17–24.

Gahagan, S. (2006). Failure to Thrive: A Consequence of. *Pediatrics in Review*, *27*(1), e1–e11.

Gairdner, D., & Pearson, J. (1971). A Growth Chart for Premature and Other Infants. *Archives of Disease in Childhood*, *46*(250), 783–787.

Gale, C. R., O'Callaghan, F. J., Bredow, M., & Martyn, C. N. (2006). The influence of head growth in fetal life, infancy, and childhood on intelligence at the ages of 4 and 8 years. *Pediatrics*, *118*, 1486–1492.

Garner, P., Panpanich, R., & Logan, S. (2000). Is routine growth monitoring effective? A systematic review of trials. *Archives of Disease in Childhood*, *82*(3), 197–201.

Garza, C., & de Onis, M. (2004). Rationale for developing a new international growth reference. *Food and Nutrition Bulletin*, *25*(1 Suppl), S5–14.

Garza, C., de Onis, M., Martorell, R., Lartey, A., Dewey, K. G., & Reference, W. H. O. M. G. (2006). Relationship between physical growth and motor development in the WHO Child Growth Standards. *Acta Paediatrica*, *95*, 96–101.

Gibbons, T., & Fuchs, G. J. (2009). Malnutrition: a hidden problem in hospitalized children. *Clinical Pediatrics*, *48*(4), 356.

Golden, M. H. (1994). Is complete catch-up possible for stunted malnourished children? *Eur J Clin Nutr*.

Goldenberg, R. L., & Culhane, J. F. (2007). Low birth weight in the United States. *The American Journal of Clinical Nutrition*, *85*(2), 584S–590S.

Gorstein, J., Sullivan, K., Yip, R., de Onís, M., Trowbridge, F., Fajans, P., & Clugston, G. (1994). Issues in the assessment of nutritional status using anthropometry. *Bulletin of the World Health Organization*, *72*(2), 273–283.

Grantham-McGregor, S. (1995). A review of studies of the effect of severe malnutrition on mental development. *Journal of Nutrition*, *125*(8s), 2233S–2238S.

Gwatkin, D. R., Wilcox, J. R., & Wray, J. D. (1980). Can health and nutrition interventions make a difference? *Monograph, Overseas Development Council*, *13*.

Hack, M., & Fanaroff, A. A. (2000). Outcomes of children of extremely low birthweight and gestational age in the 1990s. *Seminars in Neonatology : SN*, *5*, 89–106.

Hamill, P. V, Drizd, T. A., Johnson, C. L., Reed, R. B., Roche, A. F., & Moore, W. M. (1979). Physical growth: National Center for Health Statistics percentiles. *The American Journal of Clinical Nutrition*, *32*(3), 607–29.

Hanson, J. W., Streissguth, A. P., & Smith, D. W. (1978). The effects of moderate alcohol consumption during pregnancy on fetal growth and morphogenesis. *Journal of Pediatrics*, *92*, 457–460.

Hastie, T., & Tibshirani, R. (1986). Generalized Additive Models. *Statistical Science*, *1*(3), 297–310.

Heimendinger, J., & Laird, N. (1983). Measuring the Effect of an Intervention. *Evaluation Review 7.1*, *7*(1), 80–95.

Hirschfeld, S. (1996). Dysregulation of growth and development in HIV-infected children. *Journal of Nutrition*, *126*, S2641–S2650.

Holme, A., Blair, P., & Emond, A. (2013). Psychosocial and educational outcomes of weight faltering in infancy in ALSPAC. *BMJ Open*, *3*(7), 1–9.

Hui, L., Schooling, C., Cowling, B., Leung, S., Lam, T., & Leung, G. (2008). Are universal standards for optimal infant growth appropriate? Evidence from a Hong Kong Chinese birth cohort. *Archives of Disease in Childhood*, *93*(7), 561–565.

Iannotti, L. L., Zavaleta, N., Huasaquiche, C., Leon, Z., & Caulfield, L. E. (2015). Early growth velocities and weight gain plasticity improve linear growth in Peruvian infants. *Maternal & Child Nutrition*, *11*(1), 127–37.

Iwaniec, D. (2004). *Children who Fail to Thrive: A Practice Guide*. Wiley.

Jalil, F., Lindblad, B. S., Hanson, L. A., Khan, S. R., Ashraf, R. N., Carlsson, B., … Karlberg, J. (1993). Early child health in Lahore, Pakistan: I. Study design. *Acta Paediatr Suppl*, *82 Suppl 3*, 3–16.

Jalil, F., Lindblad, B. S., Hanson, L. A., Khan, S. R., Yaqoob, M., & Karlberg, J. (1993). Early child health in Lahore, Pakistan: IX. Perinatal events. *Acta Paediatrica (Oslo, Norway : 1992). Supplement*, *82 Suppl 3*, 95–107.

Jamison, D., Feachem, R., & Makgoba, M. (2006). *Disease and Mortality in Sub-Saharan Africa.* (2nd editio). Washington (DC): World Bank.

Johnson, N. L. (1949). Systems of frequency curves generated by methods of translation. *Biometrika*, *36*(Pt. 1-2), 149–176.

Júlíusson, P. B., Roelants, M., Hoppenbrouwers, K., Hauspie, R., & Bjerknes, R. (2011). Growth of Belgian and Norwegian children compared to the WHO growth standards: prevalence below -2 and above +2 SD and the effect of breastfeeding. *Archives of Disease in Childhood*, *96*(10), 916–921.

Karlberg, J., Ashraf, R. N., Saleemi, M., Yaqoob, M., & Jalil, F. (1993). Early child health in Lahore, Pakistan: XI. Growth. *Acta Paediatrica. Supplementum*, *82 Suppl 3*, 119–149.

Kendall, M., & Stuart, A. (1979). *The advanced theory of statistics* ((Vol 1)). New York: Hafner.

Khalil, K., Lindblom, G., Mazhar, K., Khan, S., & Kajiser, B. (1993). Early child health in Lahore, Pakistan: VIII. Microbiology. *Acta Paediatrica, International Journal of Paediatrics, Supplement*.

Khan, S. R., Jalil, F., Zaman, S., Lindblad, B. S., & Karlberg, J. (1993). Early child health in Lahore, Pakistan: X. Mortality. *Acta Paediatr Suppl*, *82 Suppl 3*, 109–117.

Khara, T., & Dolan, C. (2014). Technical briefing paper: Associations between wasting and stunting, policy, programming and research implications. Oxford: Emergency Nutrition Network.

Kliegman, R. (2012). *Nelson Textbook of Pediatrics*. Saunders Elsevier.

Kramer, M. S., Matush, L., Vanilovich, I., Platt, R. W., Bogdanovich, N., Sevkovskaya, Z., … Shapiro, S. (2007). Effects of prolonged and exclusive breastfeeding on child height, weight, adiposity, and blood pressure at age 6.5 y: evidence from a large randomized trial. *The American Journal of Clinical Nutrition*, *86*(6), 1717–21.

Krieger, I. (1970). Growth failure and congenital heart disease: energy and nitrogen balance in infants. *American Journal of Diseases of Children*, *120*(6), 497–502.

Kuklina, E. V, Ramakrishnan, U., Stein, A. D., Barnhart, H. H., & Martorell, R. (2004). Growth and diet quality are associated with the attainment of walking in rural Guatemalan infants. *The Journal of Nutrition*, *134*(12), 3296–3300.

Kulmala, T., Vaahtera, M., Ndekha, M., Cullinan, T., Salin, M. L., Koivisto, A. M., & Ashorn, P. (2000). Socio-economic support for good health in rural Malawi. *East African Medical Journal*, *77*(3), 168–171.

Lepage, P., Msellati, P., Hitimana, D.-G., Bazubagira, A., Van Goethem, C., Simonon, A., … Dabis, F. (1996). Growth of human immunodeficiency type 1-infected and uninfected children: A prospective cohort study in Kigali, Rwanda, 1988 to 1993. *Pediatric Infectious Disease Journal*, *15*(6), 479–485.

Li, H., Stein, A. D., Barnhart, H. X., Ramakrishnan, U., & Martorell, R. (2003). Associations between prenatal and postnatal growth and adult body size and composition. *The American Journal of Clinical Nutrition*, *77*(6), 1498–1505.

Lutter, C., Mora, J., Habicht, J., Rasmussen, K., Sellers, S. G., Super, C. M., & Herrera, M. (1988). Nutritional supplementation: effects on child stunting because of diarrhea. *The American Journal of Clinical Nutrition*, *50*(1), 1–8.

Mackay, D. J. C. (1998). Introduction to Gaussian processes. *Neural Networks and Machine Learning*, *168*(1996), 133–165.

Mahmud, A., Jalil, F., Karlberg, J., & Lindblad, B. S. (1993). Early child health in Lahore, Pakistan: VII. Diarrhoea. *Acta Paediatrica, International Journal of Paediatrics, Supplement*, *82*(390), 79–85.

Maleta, K., Virtanen, S., Espo, M., Kulmala, T., & Ashorn, P. (2003a). Timing of growth faltering in rural Malawi. *Archives of Disease in Childhood*, *88*(7), 574–8.

Maleta, K., Virtanen, S. M., Espo, M., Kulmala, T., & Ashorn, P. (2003). Childhood malnutrition and its predictors in rural Malawi. *Paediatric and Perinatal Epidemiology*, *17*(4), 384–390.

Maleta, K., Virtanen, S. M., Espo, M., Kulmala, T., & Ashorn, P. (2003b). Seasonality of growth and the relationship between weight and height gain in children under three years of age in rural Malawi. *Acta Pædiatrica*, *92*(4), 491–497.

Marcus, R., Peritz, E., & Gabriel, K. E. (1976). On closed testing procedures with special reference to ordered analysis of variance. *Biometrika*, *63*(3), 655–660.

Martorell, R. (1992). Overview of long-term nutrition intervention studies in Guatemala, 1968-1989. *Food & Nutrition Bulletin*, *14*(3), 270–277.

Martorell, R., Khan, L. K., & Schroeder, D. G. (1994). Reversibility of stunting: epidemiological findings in children from developing countries. *European Journal of Clinical Nutrition*, *48 Suppl 1*, S45–S57.

McDonald, C. M., Olofin, I., Flaxman, S., Fawzi, W. W., Spiegelman, D., Caulfield, L. E., … Danaei, G. (2013). The effect of multiple anthropometric deficits on child mortality: meta-analysis of individual data in 10 prospective studies from developing countries. *The American Journal of Clinical Nutrition*, *97*(4), 896–901.

McIntire, D. D., Bloom, S. L., Casey, B. M., & Leveno, K. J. (1999). Birth weight in relation to morbidity and mortality among newborn infants. *The New England Journal of Medicine*, *340*(16), 1234–1238.

Mclaren, D., & Read, W. C. (1972). Clasification of Nutritional Status in Early Childhood. *The Lancet*, *300*(7769), 146–148.

Mclaren, D., & Read, W. C. (1975). WEIGHT/LENGTH CLASSIFICATION OF NUTRITIONAL STATUS. *The Lancet*, *306*(7927), 219–221.

Miller, R. (2011). *Survival Analysis* (2nd ed.). Wiley.

Mjönes, S. (1987). Growth in Turkish children in Stockholm. *Annals of Human Biology*, *14*(4), 337–47.

Moy, R., & Wright, C. (2014). Using the new UK-WHO growth charts. *Paediatrics and Child Health (United Kingdom)*.

Multicentre, W. H. O., Reference, G., & Group, S. (2006). WHO Child Growth Standards based on length/height, weight and age. *Acta Paediatrica. Supplementum*, *450*, 76–85.

Naeye, R. L., Burt, L. S., Wright, D. L., Blanc, W. A., & Tatter, D. (1971). NEONATAL MORTALITY, THE MALE DISADVANTAGE. *Pediatrics*, *48*(6), 902–906.

Neinstein, L. (2002). *Adolescent Health Care: A Practical Guide* (4th ed.). Lippincott Williams and Wilkins.

Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized Linear Models. *Journal of the Royal Statistical Society. Series A (General)*, *135*(3), pp. 370–384.

Nikolaeva, R., Bhatnagar, A., & Ghose, S. (2015). Exploring curvilinearity through fractional polynomials in management research. *Organizational Research Methods*, *18*(4), 738–760.

O'Neill, S., Fitzgerald, A., Briend, A., & Van den Broeck, J. (2012). Child mortality as predicted by nutritional status and recent weight velocity in children under two in rural Africa. *Journal of Nutrition*, *142*(3), 520–525.

Olofin, I., McDonald, C. M., Ezzati, M., Flaxman, S., Black, R. E., Fawzi, W. W., … Danaei, G. (2013). Associations of suboptimal growth with all-cause and cause-specific mortality in children under five years: a pooled analysis of ten prospective studies. *PloS One*, *8*(5), e64636.

Olsen, E. M., Petersen, J., Skovgaard, a M., Weile, B., Jørgensen, T., & Wright, C. M. (2007). Failure to thrive: the prevalence and concurrence of anthropometric criteria in a general infant population. *Archives of Disease in Childhood*, *92*(2), 109–114.

Ong, K. K., Ahmed, M. L., Emmett, P. M., Preece, M. A., & Dunger, D. B. (2000). Association between postnatal catch-up growth and obesity in childhood: prospective cohort study. *BMJ (Clinical Research Ed.)*, *320*(7240), 967–71.

Painter, R., Osmond, C., Gluckman, P., Hanson, M., Phillips, D., & Roseboom, T. (2008). Transgenerational effects of prenatal exposure to the Dutch famine on neonatal adiposity and health in later life. *BJOG*, *115*(10), 1243–1249.

Pan, H., & Cole, T. (2011). LMS growth: a Microsoft Excel add-in to access growth references based on the LMS method. Version 2.74.

Panpanich, R., & Garner, P. (1999). Growth monitoring in children. *Cochrane Database of Systematic Reviews*, (4).

Patel, D., Bland, R., Coovadia, H., Rollins, N., Coutsoudis, A., & Newell, M.-L. (2010). Breastfeeding, HIV status and weights in South African children: a comparison of HIV-exposed and unexposed children. *AIDS (London, England)*, *24*(3), 437–45.

Pearson, K. (1900). "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling." *Philosophical Magazine Series 5*, *50*(302), 157–175.

Pelletier, D. L., Frongillo, E. A., Schroeder, D. G., & Habicht, J. P. (1995). The effects of malnutrition on child mortality in developing countries. *Bulletin of the World Health Organization*, *73*(4), 443–448.

Pringle, P. J., Geary, M. P. P., Rodeck, C. H., Kingdom, J. C. P., Kayamba-Kay'S, S., & Hindmarsh, P. C. (2005). The influence of cigarette smoking on antenatal growth, birth size, and the insulin-like growth factor axis. *Journal of Clinical Endocrinology and Metabolism*, *90*(February), 2556–2562.

R Development Core Team. (2013). R Development Core Team. *R: A Language and Environment for Statistical Computing*.

Rao, K. V., & Singh, D. (1970). An Evaluation of the Relationship between Nutritional Status and Anthropometric Measurements. *The American Journal for Clinical Nutrition*, *83-93*(23), 83–93.

Relman, D. (2013). Undernutrition - Looking within for answers. *Science*, *339*(6119), 530–532.

Richard, S. A., Black, R. E., & Checkley, W. (2014). Revisiting the relationship of weight and height in early childhood. *World Review of Nutrition and Dietetics*, *109*(6), 93.

Richard, S. A., Black, R. E., Gilman, R. H., Guerrant, R. L., Kang, G., Lanata, C. F., … Checkley, W. (2012). Wasting Is Associated with Stunting in Early Childhood. *Journal of Nutrition*, *142*(7), 1291–1296.

Rigby, R. A., & Stasinopoulos, D. M. (2004). Smooth centile curves for skew and kurtotic data modelled using the Box-Cox power exponential distribution. *Statistics in Medicine*, *23*(19), 3053–3076.

Rigby, R. A., & Stasinopoulos, D. M. (2013). Automatic smoothing parameter selection in GAMLSS with an application to centile estimation. *Statistical Methods in Medical Research*, *23*(4), 318–332.

Rigby, R., & Stasinopoulos, D. (2004). *Box-Cox t distribution for modelling skew and leptokurtotic data.*

Rigby, R., & Stasinopoulos, D. (2005). Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society. Series C: Applied Statistics*, *54*(3), 507–554.

Roberfroid, D., Kolsteren, P., Hoeree, T., & Maire, B. (2005). Do growth monitoring and promotion programs answer the performance criteria of a screening program? A critical analysis based on a systematic review. *Tropical Medicine & International Health*, *10*(11), 1121-1133.

Roelants, M. (2013). *Normal variation in human growth*. Vrije Universiteit Brussel.

Rollins, N. C., Becquet, R., Bland, R. M., Coutsoudis, A., Coovadia, H. M., & Newell, M.-L. (2008). Infant feeding, HIV transmission and mortality at 18 months: the need for appropriate choices by mothers and prioritization within programmes. *AIDS*, *22*, 2349–2357.

Rowland, M. G., Cole, T. J., & Whitehead, R. G. (1977). A quantitative study into the role of infection in determining nutritional status in Gambian village children. *The British Journal of Nutrition*, *37*(3), 441–50.

Rowland, M. G., & McCollum, J. P. (1977). Malnutrition and gastroenteritis in The Gambia. *Trans.R.Soc.Trop.Med.Hyg.*, *71*(0035-9203 LA - eng PT - Journal Article RN - 0 (Bile Acids and Salts) SB - IM), 199–203.

Rowland, M., Rowland, S., & Cole, T. (1988). Impact of infection on the growth of children from 0 to 2 years in an urban West African community. *The American Journal of Clinical Nutrition*, *47*(1), 134–138.

Royston, P., & Altman, D. G. (1994). Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling. *Appl Statist*, *16*(2), 165–166.

Royston, P., & Wright, E. M. (1998). A method for estimating age-specific reference intervals (’normal ranges') based on fractional polynomials and exponential transformation. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *161*(1), 79–101.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, *63*(3), 581–592.

Saleemi, M., Ashraf, R., Mellander, L., & Zaman, S. (2001). Determinants of stunting at 6, 12, 24 and 60 months and postnatal linear growth in Pakistani children. *Acta Paediatrica*, *90*(11), 1304–1308.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147–177.

Schroeder, D. G., & Brown, K. H. (1994). Nutritional status as a predictor of child survival: summarizing the association and quantifying its global impact. *Bulletin of the World Health Organization*, *72*(4), 569–79.

Schumacher, L. B., Pawson, I. G., & Kretchmer, N. (1987). Growth of immigrant children in the newcomer schools of San Francisco. *Pediatrics*, *80*(6), 861–868.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464.

Selwyn, P. A., Schoenbaum, E. E., Davenny, K., Robertson, V. J., Feingold, A. R., Shulman, J. F., Rogers, M. F. (1989). Prospective study of human immunodeficiency virus infection and pregnancy outcomes in intravenous drug users. *Jama*, *261*(9), 1289–1294.

Sheskin, D. J. (1997). *Handbook of Parametric and Nonparametric Statistical Procedures*. *American Statistician* (Vol. 51).

Shields, B., Wacogne, I., & Wright, C. M. (2012a). Weight faltering and failure to thrive in infancy and early childhood. *BMJ (Clinical Research Ed.)*, *345*(September), e5931.

Shields, B., Wacogne, I., & Wright, C. M. (2012b). Weight faltering and failure to thrive in infancy and early childhood. *BMJ (Clinical Research Ed.)*, *345*(September), e5931.

Sistrom, C. L., & Garvan, C. W. (2004). Proportions, odds, and risk. *Radiology*, *230*, 12–19.

Stein, A. D., Barnhart, H. X., Hickey, M., Ramakrishnan, U., Schroeder, D. G., & Martorell, R. (2003). Prospective study of protein-energy supplementation early in life and of growth in the subsequent generation in Guatemala. *The American Journal of Clinical Nutrition*, *78*(1), 162–167.

Stein, A. D., Barnhart, H. X., Wang, M., Hoshen, M. B., Ologoudou, K., Ramakrishnan, U., … Martorell, R. (2004). Comparison of linear growth patterns in the first three years of life across two generations in Guatemala. *Pediatrics*, *113*(3 Pt 1), e270–5.

Stigler, S. M. (1997). Regression towards the mean, historically considered. *Statistical Methods in Medical Research*, *6*(2), 103–114.

Stuebe, A. (2009). The risks of not breastfeeding for mothers and infants. *Reviews in Obstetrics and Gynecology*, *2*(4), 222–231.

Tanner, J., & Buckler, J. (1997). Revision and update of Tanner-Whitehouse clinical longitudinal charts for height and weight. *European Journal of Pediatrics*, *156*(3), 248–249.

Tanner, J. M. (1990). *Fetus into Man: Physical Growth from Conception to Maturity* (2nd editio). Cambridge, MA: Harvard University Press.

Tanner, J. M., & Whitehouse, R. H. (1975). Revised standards for triceps and subscapular skinfolds in British children. *Archives of Disease in Childhood*, *50*(2), 142–145.

Tanner, J. M., & Whitehouse, R. H. (1976). Clinical longitudinal standards for height, weight, height velocity, weight velocity, and stages of puberty. *Archives of Disease in Childhood*, *51*(3), 170–179.

Tanner, J. M., Whitehouse, R. H., & Takaishi, M. (1966a). Standards from birth to maturity for height, weight, height velocity, and weight velocity: British children, 1965. I. *Archives of Disease in Childhood*, *41*(219), 454–471.

Tanner, J. M., Whitehouse, R. H., & Takaishi, M. (1966b). Standards from birth to maturity for height, weight, height velocity, and weight velocity: British children, 1965. II. *Archives of Disease in Childhood*, *41*(220), 613–635.

Unicef. (2015). PROGRESS FOR CHILDREN: A WORLD FIT FOR CHILDREN STATISTICAL REVIEW. Retrieved December 7, 2015, from http://www.unicef.org/progressforchildren/2007n6/index_41505.htm

Vaahtera, M., Kulmala, T., Ndekha, M., Koivisto, a M., Cullinan, T., Salin, M. L., & Ashorn, P. (2000). Antenatal and perinatal predictors of infant mortality in rural Malawi. *Archives of Disease in Childhood. Fetal and Neonatal Edition*, *82*, F200–F204.

Van Den Broeck, J., Eeckels, R., & Vuylsteke, J. (1993). Influence of nutritional status on child mortality in rural Zaire. *Lancet*, *341*(8859), 1491–5.

Van Wieringen, J. C. (1978). Secular growth changes. In *Human growth* (pp. 445–473). Springer US.

Vella, V., Tomkins, A., Borghesi, A., Migliori, G. B., & Oryem, V. Y. (1994). Determinants of stunting and recovery from stunting in northwest uganda. *International Journal of Epidemiology*, *23*(4), 782–786.

Vesel, L., Bahl, R., Martines, J., Penny, M., Bhandari, N., Kirkwoodd, B. R., … Underwood, B. (2010). Use of new World Health Organization child growth standards to assess how infant malnutrition relates to breastfeeding and mortality. *Bulletin of the World Health Organization*, *88*(1), 39–48.

Victora, C. G. (1992). The association between wasting and stunting: an international perspective. *The Journal of Nutrition*, *122*(5), 1105–1110.

Victora, C. G., Adair, L., Fall, C., Hallal, P. C., Martorell, R., Richter, L., & Sachdev, H. S. (2008). Maternal and child undernutrition: consequences for adult health and human capital. *The Lancet*, *371*(9609), 340–357.

Victora, C. G., de Onis, M., Hallal, P. C., Blössner, M., & Shrimpton, R. (2010). Worldwide timing of growth faltering: revisiting implications for interventions. *Pediatrics*, *125*(3), e473–e480.

Waldron, I. (1987). [Patterns and causes of excess female mortality among children in developing countries]. *World health statistics quarterly. Rapport trimestriel de statistiques sanitaires mondiales*, *40*(3), 194-210.

Wade, A. M., & Ades, A. E. (1998). Incorporating correlations between measurements into the estimation of age-related reference ranges. *Statistics in Medicine*, *17*(17), 1989–2002.

Walker, S. P., & Golden, M. (1988). Growth in length of children recovering from severe malnutrition. *European Journal of Clinical Nutrition*, *42*(5), 395–404.

Walker, S. P., Grantham-McGregor, S. M., Himes, J. H., & Powell, C. A. (1996). Relationships between wasting and linear growth in stunted children. *Acta Paediatr*, *85*(6), 666–9.

Webb, A. L., Conlisk, A. J., Barnhart, H. X., Martorell, R., Grajeda, R., & Stein, A. D. (2005). Maternal and childhood nutrition and later blood pressure levels in young Guatemalan adults. *International Journal of Epidemiology*, *34*(4), 898–904.

Whitehead, R. G., Paul, A. A., & Cole, T. J. (1989). Diet and the growth of healthy infants. *Journal of Human Nutrition and Dietetics*, *2*, 73–84.

WHO Multicentre Growth Reference Study Group. (2009). WHO Child Growth Standards: Growth velocity based on weight, length and head circumference: Methods and development. *Geneva: World Health Organization*.

WHO Multicentre Growth Research Centre. (2006). WHO Motor Development Study: Relationship between physical growth and motor development in the WHO Child Growth Standards. *Acta Pædiatrica*, *405*, 96–101.

Wood, S. (2006). *Generalized additive models: an introduction with R.* CRC press.

World Health Organisation. (2015). Moderate malnutrition. Retrieved December 7, 2015, from http://www.who.int/nutrition/topics/moderate_malnutrition/en/

Wright, C., Lakshman, R., Emmett, P., & Ong, K. K. (2008). Implications of adopting the WHO 2006 Child Growth Standard in the UK: two prospective cohort studies. *Archives of Disease in Childhood*, *93*(7), 566–569.

Wright, C. M. (2000). Identification and management of failure to thrive: a community perspective. *Archives of Disease in Childhood*, *82*(1), 5–9.

Wright, C. M., Booth, I. W., Buckler, J. M. H., Cameron, N., Cole, T. J., Healy, M. J. R., … Williams, a F. (2002). Growth reference charts for use in the United Kingdom. *Archives of Disease in Childhood*, *86*(1), 11–14.

Wright, C. M., & Cheetham, T. D. (1999). The strengths and limitations of parental heights as a predictor of attained height. *Archives of Disease in Childhood*, *81*(3), 257–260.

Wright, C. M., Matthews, J. N., Waterston, A., & Aynsley-Green, A. (1994). What is a normal rate of weight gain in infancy? *Acta Paediatr*, *83*(4), 351–356.

Wright, C. M., & Parkinson, K. N. (2004). Postnatal weight loss in term infants: what is normal and do growth charts allow for it? *Archives of Disease in Childhood. Fetal and Neonatal Edition*, *89*, F254–F257.

Yaqoob, M., Ferngren, H., Jalil, F., Nazir, R., & Karlberg, J. (1993). Early child health in Lahore, Pakistan: XII. Milestones. *Acta Paediatrica (Oslo, Norway : 1992). Supplement*, *82 Suppl 3*(16), 151–7.

Zaman, S., Jalil, F., Karlberg, J., & Hanson, L. A. (1993). Early child health in Lahore, Pakistan: VI. Morbidity. *Acta Paediatrica. Supplement*, (390), 63–78.

Zumrawi, F., Dimond, H., & Waterlow, J. (1987). Faltering in infant growth in Khartoum province, Sudan. *Human Nutrition. Clinical Nutrition*, *41*(5), 383–395.

Appendix A

A1 Proportion of children in each state (including NA values)

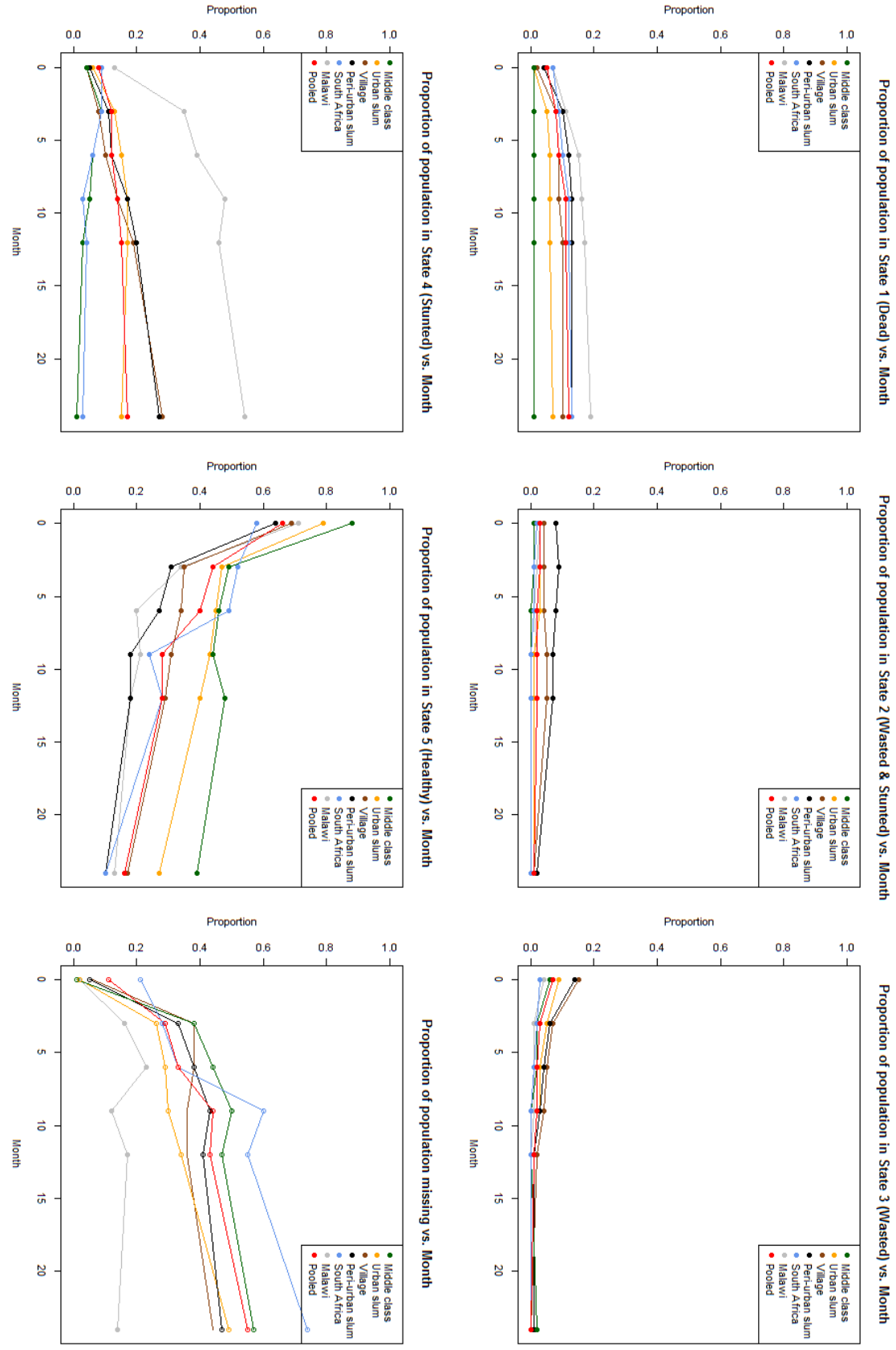

Figure A1: Proportion of population in states (including NA values): (a) dead, (b) wasted - stunted, (c) wasted, (d) stunted, (e) healthy, (f) missing

Malawi	Month					
State	0	3	6	9	12	24
Dead (State1)	0.07	0.11	0.15	0.16	0.17	0.19
Wasted & Stunted (State 2)	0.03	0.02	0.01	0.02	0.02	0.01
Wasted (State 3)	0.04	0.01	0.01	0.01	0.00	0
Stunted (State 4)	0.13	0.35	0.39	0.48	0.46	0.54
Healthy (State 5)	0.71	0.34	0.20	0.21	0.18	0.13
NA	0.02	0.16	0.23	0.12	0.17	0.14

Table A1: Proportion of children in each state, including NAs (Malawi)

Middle Class	Month					
State	0	3	6	9	12	24
Dead (State1)	0.01	0.01	0.01	0.01	0.01	0.01
Wasted & Stunted (State 2)	0.01	0.01	0	0	0	0
Wasted (State 3)	0.06	0.02	0.02	0	0	0.02
Stunted (State 4)	0.04	0.09	0.06	0.05	0.03	0.01
Healthy (State 5)	0.88	0.49	0.46	0.44	0.48	0.39
NA	0.01	0.38	0.44	0.50	0.47	0.57

Table A2: Proportion of children in each state, including NAs (middle class)

Village	Month					
State	0	3	6	9	12	24
Dead (State1)	0.02	0.08	0.09	0.09	0.10	0.10
Wasted & Stunted (State 2)	0.04	0.04	0.04	0.05	0.05	0.01
Wasted (State 3)	0.15	0.07	0.05	0.04	0.02	0
Stunted (State 4)	0.04	0.08	0.10	0.14	0.19	0.28
Healthy (State 5)	0.69	0.35	0.34	0.31	0.29	0.17
NA	0.06	0.38	0.38	0.36	0.36	0.44

Table A3: Proportion of children in each state, including NAs (village)

Urban slum	Month					
State	0	3	6	9	12	24
Dead (State1)	0.01	0.05	0.06	0.06	0.06	0.07
Wasted & Stunted (State 2)	0.03	0.03	0.03	0.01	0.01	0.01
Wasted (State 3)	0.09	0.05	0.03	0.03	0.01	0.01
Stunted (State 4)	0.06	0.13	0.15	0.17	0.17	0.15
Healthy (State 5)	0.79	0.47	0.45	0.43	0.40	0.27
NA	0.02	0.26	0.29	0.30	0.34	0.49

Table A4: Proportion of children in each state, including NAs (urban slum)

Peri-urban slum	Month					
State	0	3	6	9	12	24
Dead (State1)	0.04	0.10	0.12	0.13	0.13	0.13
Wasted & Stunted (State 2)	0.08	0.09	0.08	0.07	0.07	0.02
Wasted (State 3)	0.14	0.06	0.04	0.03	0.01	0.01
Stunted (State 4)	0.05	0.11	0.12	0.17	0.20	0.27
Healthy (State 5)	0.64	0.31	0.27	0.18	0.18	0.10
NA	0.05	0.33	0.38	0.43	0.41	0.47

Table A5: Proportion of children in each state, including NAs (peri-urban slum)

South Africa	Month					
State	0	3	6	9	12	24
Dead (State1)	0.07	0.09	0.10	0.12	0.12	0.13
Wasted & Stunted (State 2)	0.02	0.01	0.01	0	0	0
Wasted (State 3)	0.03	0.02	0.01	0	0	0
Stunted (State 4)	0.09	0.09	0.06	0.03	0.04	0.03
Healthy (State 5)	0.58	0.52	0.49	0.24	0.28	0.1
NA	0.21	0.28	0.33	0.60	0.55	0.74

Table A6: Proportion of children in each state, including NAs (South Africa)

Pooled	Month					
State	0	3	6	9	12	24
Dead (State1)	0.05	0.08	0.09	0.11	0.11	0.12
Wasted & Stunted (State 2)	0.03	0.03	0.02	0.02	0.02	0.01
Wasted (State 3)	0.07	0.03	0.02	0.02	0.01	0
Stunted (State 4)	0.08	0.12	0.12	0.14	0.15	0.17
Healthy (State 5)	0.66	0.44	0.40	0.28	0.28	0.16
NA	0.11	0.29	0.33	0.44	0.43	0.55

Table A7: Proportion of children in each state, including NAs (pooled)

Appendix B

B1 Moving into a dead state

B1.1 South Africa/Middle class

Timeframe	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.15	0.06	0.03	0.02
6-9m	0.37	0	0.02	0.03
9-12m	0	0	0	0

Table B1: Probability of moving from all states to dead

Timeframe	Wasted-stunted	Wasted	Stunted
3-6m	9.48 (3.78, 23.77)	3.87 (1.18, 12.59)	2.11 (0.90, 4.95)
6-9m	14.13 (5.23, 38.18)	NA	0.90 (0.21, 3.81)
9-12m	NA	NA	NA

Table B2: RR (95% CI) of moving from states to dead (baseline healthy)

B1.2 Village/Peri-urban slum

Timeframe	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.11	0.08	0.05	0.01
6-9m	0.07	0	0.01	0
9-12m	0.03	0.04	0.01	0

Table B3: Probability of moving from all states to dead

Timeframe	Wasted-stunted	Wasted	Stunted
3-6m	13.10 (4.20, 40.78)	8.26 (2.38, 28.60)	4.98 (1.36, 18.27)
6-9m	13.22 (2.61, 66.91)	NA	1.31 (0.11, 14.38)
19-12m	4.78 (0.68, 33.42)	6.68 (0.96, 46.41)	1.63 (0.23, 11.55)

Table B4: RR (95% CI) of moving from states to dead (baseline healthy)

B1.3 Malawi

Timeframe	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.15	0.14	0.05	0.02
6-9m	0.10	0	0.02	0.03
9-12m	0	0	0.01	0.01

Table B5: Probability of moving from all states to dead

Timeframe	Wasted-stunted	Wasted	Stunted
3-6m	8.73 (1.76, 43.36)	8.10 (1.03, 63.49)	2.86 (0.92, 8.86)
6-9m	3.6 (0.44, 29.26)	NA	0.84 (0.25, 2.85)
9-12m	NA	NA	2.19 (0.26, 18.61)

Table B6: RR (95% CI) of moving from states to dead (baseline healthy)

B1.2 Urban slum

Timeframe	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.17	0.03	0.01	0
6-9m	0.05	0	0.01	0
9-12m	0	0	0	0

Table B7: Probability of moving from all states to dead

Timeframe	Wasted-stunted	Wasted	Stunted
3-6m	NA	NA	NA
6-9m	NA	NA	NA
9-12m	NA	NA	NA

Table B8: RR (95% CI) of moving from states to dead (baseline healthy)

B2 Moving from a wasted-stunted state

B2.1 South Africa/Middle class

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.15	0.20	0.03	0.50	0.12
6-9m	0.37	0.13	0	0.05	0
9-12m	0	0.86	0.14	0	0

Table B9: Probability of moving from wasted-stunted

Timeframe	Dead	Wasted-stunted	Wasted	Stunted
3-6m	9.48 (3.78, 23.77)	265.58 (33.59, 2099.42)	1.80 (0.25, 13.04)	10.93 (7.19, 16.61)
6-9m	14.13 (5.23, 38.18)	47.12 (4.73, 468.71)	NA	8.19 (3.88, 17.30)
9-12m	NA	NA	NA	NA

Table B10: RR (95% CI) of moving from wasted-stunted (baseline healthy)

B2.2 Village/Peri-urban slum

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.11	0.40	0.02	0.40	0.05
6-9m	0.07	0.57	0.01	0.34	0.01
9-12m	0.03	0.54	0.03	0.40	0

Table B11: Probability of moving from wasted-stunted

Timeframe	Dead	Wasted-stunted	Wasted	Stunted
3-6m	13.10 (4.21, 40.78)	17.81 (9.16, 33.41)	0.28 (0.06, 1.15)	4.67 (3.11, 7.01)
6-9m	13.22 (2.61, 66.91)	20.67 (11.17, 38.25)	0.22 (0.03, 1.67)	2.08 (1.42, 3.05)
9-12m	4.78 (0.68, 33.41)	38.27 (15.53, 93.69)	0.86 (0.19, 3.84)	2.92 (2.00, 4.27)

Table B12: RR (95% CI) of moving from wasted-stunted (baseline healthy)

B2.3 Malawi

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.15	0.46	0	0.38	0
6-9m	0.10	0.20	0	0.70	0
9-12m	0	0.14	0	0.86	0

Table B13: Probability of moving from wasted-stunted

Timeframe	Dead	Wasted-stunted	Wasted	Stunted
3-6m	8.73 (1.75, 43.36)	104.77 (13.59, 807.28)	NA	0.89 (0.44, 1.80)
6-9m	3.6 (0.4,4 29.26)	NA	NA	2.14 (1.34, 3.42)
9-12m	NA	7.00 (1.27, 38.43)	NA	2.21 (1.64, 2.96)

Table B14: RR (95% CI) of moving from wasted-stunted (baseline healthy)

B2.2 Urban slum

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.17	0.42	0	0.42	0
6-9m	0.05	0.35	0.15	0.35	0.10
9-12m	0	0.40	0	0.50	0.10

Table B15: Probability of moving from wasted-stunted

Timeframe	Dead	Wasted-stunted	Wasted	Stunted
3-6m	NA	NA	NA	3.98 (2.25, 7.04)
6-9m	NA	135.41 (18.08, 1014.12)	7.75 (2.09, 28.81)	4.03 (2.00, 8.09)
9-12m	NA	54.42 (12.08, 245.15)	NA	7.02 (3.31, 14.87)

Table B16: RR (95% CI) of moving from wasted-stunted (baseline healthy)

B3 Moving from a wasted state

B3.1 South Africa/Middle class

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.06	0.02	0.12	0.16	0.64
6-9m	0	0	0.15	0.07	0.78
9-12m	0	0.10	0.90	0	0

Table B17: Probability of moving from wasted

Timeframe	Dead	Wasted-stunted	Wasted	Stunted
3-6m	3.87 (1.18, 12.59)	25.80 (1.63, 406.58)	7.37 (3.11, 17.45)	3.49 (1.76, 6.92)
6-9m	NA	NA	17.95 (3.96, 81.30)	1.17 (0.17, 7.90)
9-12m	NA	NA	NA	NA

Table B18: RR (95% CI) of moving from wasted (baseline healthy)

B3.2 Village/Peri-urban slum

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.08	0.15	0.34	0.18	0.25
6-9m	0	0.12	0.38	0.15	0.35
9-12m	0.04	0.18	0.23	0.17	0.38

Table B19: Probability of moving from wasted

Timeframe	Dead	Wasted-stunted	Wasted	Stunted
3-6m	8.25 (2.38, 28.60)	6.60 (2.95, 14.76)	4.01 (2.60, 6.20)	2.08 (1.18, 3.67)
6-9m	NA	4.42 (1.85, 10.60)	6.62 (4.00, 10.94)	0.92 (0.50, 1.70)
9-12m	6.67 (0.96, 46.41)	13.35 (4.75, 37.56)	7.28 (3.38, 15.66)	1.22 (0.64, 2.34)

Table B20: RR (95% CI) of moving from wasted (baseline healthy)

B3.3 Malawi

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.14	0	0	0.57	0.29
6-9m	0	0	0	0.75	0.25
9-12m	0	0.33	0.17	0.17	0.33

Table B21: Probability of moving from wasted

Timeframe	Dead	Wasted-stunted	Wasted	Stunted
3-6m	8.10 (1.03, 63.49)	NA	NA	1.32 (2.29, 0.42)
6-9m	NA	NA	NA	2.29 (1.24, 4.23)
9-12m	NA	16.33 (3.32, 80.26)	NA	0.42 (0.07, 2.60)

Table B22: RR (95% CI) of moving from wasted (baseline healthy)

B3.2 Urban slum

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.03	0.12	0.25	0.16	0.44
6-9m	0	0.04	0.48	0.17	0.30
9-12m	0	0	0.32	0.05	0.64

Table B23: Probability of moving from wasted

Timeframe	Dead	Wasted-stunted	Wasted	Stunted
3-6m	NA	40.62 (4.68, 352.63)	6.77 (2.98, 15.33)	1.49 (0.62, 3.54)
6-9m	NA	6.76 (0.63, 71.80)	24.78 (10.07, 60.97)	2.00 (0.76, 5.23)
9-12m	NA	NA	NA	0.63 (0.08, 4.53)

Table B24: RR (95% CI) of moving from wasted (baseline healthy)

B4 Moving from a stunted state

B4.1 South Africa/Middle class

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.04	0.06	0.02	0.38	0.51
6-9m	0.03	0.03	0.04	0.41	0.49
9-12m	0	0	0	0.63	0.37

Table B25: Probability of moving from stunted

Timeframe	Dead	Wasted-stunted	Wasted	Stunted
3-6m	2.11 (0.90, 4.95)	66.61 (8.64, 513.37)	1.15 (0.39, 3.32)	8.51 (6.30, 11.50)
6-9m	0.90 (0.21, 3.81)	13.62 (2.30, 80.38)	4.54 (1.15, 17.82)	7.10 (4.89, 10.31)
9-12m	NA	NA	NA	9.55 (7.05, 12.94)

Table B26: RR (95% CI) of moving from stunted (baseline healthy)

B4.2 Village/Peri-urban slum

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.05	0.12	0.05	0.54	0.24
6-9m	0.01	0.10	0.01	0.72	0.16
9-12m	0.01	0.09	0	0.80	0.10

Table B27: Probability of moving from stunted

Timeframe	Dead	Wasted-stunted	Wasted	Stunted
3-6m	4.98 (1.36, 18.26)	5.58 (2.55, 12.23)	0.53 (0.21, 1.33)	6.36 (4.46, 9.06)
6-9m	1.31 (0.11, 14.38)	3.58 (1.68, 7.62)	0.11 (0.01, 0.83)	4.41 (3.47, 5.62)
9-12m	1.63 (0.23, 11.55)	6.22 (2.35, 16.43)	NA	5.81 (4.45, 7.60)

Table B28: RR (95% CI) of moving from stunted (baseline healthy)

B4.3 Malawi

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.05	0.02	0.01	0.82	0.11
6-9m	0.02	0.03	0.01	0.77	0.16
9-12m	0.01	0.02	0	0.84	0.12

Table B29: Probability of moving from stunted

Timeframe	Dead	Wasted-stunted	Wasted	Stunted
3-6m	2.96 (0.92, 8.85)	4.16 (0.47, 36.97)	0.69 (0.11, 4.11)	1.89 (1.60, 2.22)
6-9m	0.84 (0.25, 2.85)	NA	1.45 (0.15, 13.86)	2.37 (1.86, 3.02)
9-12m	2.19 (0.26, 18.61)	1.02 (0.26, 3.90)	NA	2.16 (1.75, 2.66)

Table B30: RR (95% CI) of moving from stunted (baseline healthy)

B4.2 Urban slum

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.01	0.07	0.03	0.61	0.28
6-9m	0.01	0	0	0.76	0.23
9-12m	0	0.02	0.01	0.82	0.16

Table B31: Probability of moving from stunted

Timeframe	Dead	Wasted-stunted	Wasted	Stunted
3-6m	NA	23.45 (2.92, 188.30	0.83 (0.24, 2.90)	5.81 (4.07, 8.29)
6-9m	NA	NA	NA	8.79 (6.04, 12.79)
9-12m	NA	4.49 (0.41, 0.49)	NA	11.46 (7.45, 17.63)

Table B32: RR (95% CI) of moving from stunted (baseline healthy)

B5 Moving from a healthy state

B5.1 South Africa/Middle class

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.02	0.01	0.02	0.05	0.90
6-9m	0.03	0	0.01	0.06	0.90
9-12m	0	0	0	0.07	0.93

Table B33: Pooled probability of moving from healthy

B5.2 Village/Peri-urban slum

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.01	0.02	0.08	0.09	0.80
6-9m	0.01	0.03	0.06	0.16	0.74
9-12m	0	0.01	0.03	0.14	0.81

Table B34: Pooled probability of moving from healthy

B5.3 Malawi

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0.02	0	0.01	0.43	0.53
6-9m	0.03	0	0.01	0.33	0.64
9-12m	0.01	0.02	0	0.39	0.59

Table B35: Pooled probability of moving from healthy

B5.2 Urban slum

Timeframe	Dead	Wasted-stunted	Wasted	Stunted	Healthy
3-6m	0	0	0.04	0.10	0.86
6-9m	0	0.01	0.02	0.09	0.89
9-12m	0	0	0	0.07	0.93

Table B36: Pooled probability of moving from healthy