University of Southampton

Faculty of Medicine

School of Clinical and Experimental Sciences

Clinical, Nutritional, Genomic and Metabolomic Influences on Growth and Body Composition in Very Preterm Infants

Volume 1 of 1

By:

Aneurin Young

ORCID: 0000-0003-3549-3813

Thesis for the degree of Doctor of Philosophy

April 2023

University of Southampton

Abstract

Faculty of Medicine

School of Clinical and Experimental Sciences

Thesis for the degree of Doctor of Philosophy

Clinical, Nutritional, Genomic and Metabolomic Influences on Growth and Body Composition in Very Preterm Infants

by

Aneurin Young

Infants born before 32 weeks postmenstrual age (PMA) are at high risk of growth failure. Current guidelines recommend that the growth of preterm infants should match that of the equivalent fetus in utero, both in terms of weight gain and body composition, but that target is commonly missed. There is emerging evidence that nutrition and growth during the neonatal period is associated with neurodevelopmental outcome. However, factors influencing the growth of very preterm infants are incompletely understood, limiting the capability of clinicians to adjust nutritional care to the individual needs of each infant.

This research project aims to assess multiple clinical, nutritional, genomic and metabolomic influences on the growth of very preterm infants, working towards a toolkit to guide personalised nutritional care.

The work is centred on the formation of a comprehensive quality-assured relational database: the Southampton Preterm Nutritional Database. This contains prospectively gathered information for over 600 infants born before 32 weeks PMA and cared for in Southampton's neonatal unit. It combines nutritional intake data for over 33,000 care days with comprehensive demographic, clinical and biochemical information. Regression and machine learning techniques can be applied to these data to provide insights into the impact of these factors on growth. This research programme also includes total body water analysis of nine infants, using deuterium oxide dilution, allowing a marker of body composition to be tracked longitudinally in this subgroup. Whole exome sequencing has been performed for 13 infants, with these data being included in the database as a pilot for future analysis of genetic factors influencing growth. Metabolomic analysis of weekly urine samples from 14 infants also provides a model for the investigation of the effect of genomic, clinical and nutritional factors on the metabolic maturation of the very preterm infant. In addition to detailed analysis of these infants from a single centre, growth data for around 30,000 infants born across England were analysed to assess changes in growth patterns over time.

Random forest machine learning has been used to identify the key factors influencing growth. Growth charts were published based on the local cohort of infants for whom a detailed accompanying description of nutritional and clinical factors was available, setting out an expected growth pattern in response to a defined nutritional approach. Growth, nutritional and biochemical results have also allowed presentation of an exploration of the influence of protein intake on plasma urea. A changing national pattern of weight gain, with generally more rapid growth in the most preterm infants, has been identified.

Taken together, the findings presented in this thesis provide a guide to key modifiable factors influencing growth, a range of charts to monitor growth and pilot data on total body water, genomic analysis and metabolomic profiling which will be employed to provide further insights in the future. The work is underpinned by the formation of a comprehensive research database.

Table of Contents

Table of Tables

Table of Figures

List of Appendices

Research Thesis: Declaration of Authorship

Print name: Aneurin Young

Title of thesis: Clinical, Nutritional, Genomic and Metabolomic Influences on Growth and Body Composition in Very Preterm Infants

I declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:
    - Young A, Andrews ET, Ashton JJ, Pearson F, Beattie RM, Johnson MJ. *Generating longitudinal growth charts from preterm infants fed to current recommendations*. Arch Dis Child Fetal Neonatal Ed. 2020;105(6):646-51.[1] (Appendix 1)
    - Young A, Brown LK, Ennis S, Beattie RM, Johnson MJ. *Total body water in full-term and preterm newborns: systematic review and meta-analysis*. Arch Dis Child Fetal Neonatal Ed. 2021;106(5):542-8.[2] (Appendix 2)
    - Young A, Beattie RM, Johnson MJ. *Optimising growth in very preterm infants: reviewing the evidence*. Arch Dis Child Fetal Neonatal Ed. 2022.[3] (Appendix 3)
    - Young A, Cole TJ, Cheng G, Ennis S, Beattie RM, Johnson MJ. *Changes in the growth of very preterm infants in England 2006-2018*. Arch Dis Child Fetal Neonatal Ed. 2022.[4] (Appendix 4)

Signature: **[Redacted from online copy]** Date: 14th April 2023

Acknowledgements

I would like to acknowledge the careful guidance and steadfast support of my supervisors, Dr Mark Johnson, Prof R Mark Beattie and Prof Sarah Ennis. I also acknowledge Dr Hang Phan's assistance with database setup and Dr Guo Cheng's expert statistical and genomics advice. Prof Tim Cole (UCL Great Ormond Street Institute of Child Health) provided invaluable guidance with the use of the SITAR technique and assessment of growth patterns. Prof Simon Eaton (UCL Great Ormond Street Institute of Child Health) performed mass spectrometry for detection of deuterium oxide. Prof Jonathan Swann (University of Southampton) supported method development, sample analysis and interpretation of metabolomic samples.

I acknowledge the Neonatal Data Analysis Unit (Imperial College London) as the source of the National Neonatal Research Database (NNRD) data used in this thesis, the neonatal units contributing data to the NNRD and the patients included on the database.

I would especially like to thank the research nurses of the neonatal unit at University Hospital Southampton: Sister Pip Crowley, Sister Sarah McKay and Sister Jenny Pond, who gathered much of the nutritional data for this project and who enthusiastically supported acquisition of clinical samples. I am very grateful to the infants and families who received care in Southampton and who were included in the data used for this thesis.

Abbreviations Used

AAP	American Academy of Pediatrics
AGA	Appropriate Weight for Gestational Age
ANOVA	Analysis of Variance
AUC	Area Under the Receiver Operator Characteristic Curve
BAM	Binary Alignment Map file
BIA	Bioelectrical Impedance Analysis
BIC	Bayesian Information Criterion
BiPAP	Biphasic Positive Airway Pressure
BM	Breastmilk
BPD	Bronchopulmonary Dysplasia
BRC	Biomedical Research Centre
CADD	Combined Annotation–Dependent Depletion score
CCDS	Consensus Coding Sequence
COVID-19	Coronavirus disease 2019
CPAP	Continuous Positive Airway Pressure
CRP	C-reactive Protein
CSV	Comma-separated Values file
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleoside Triphosphate base
DXA	Dual-energy X-ray Absorptiometry
EDTA	Ethylenediaminetetraacetic Acid
eGFR	Estimated Glomerular Filtration Rate
ESPGHAN	European Society of Paediatric Gastroenterology, Hepatology and Nutrition
FFM	Fat Free Mass
FM	Fat Mass
GAP	Growth Assessment of Preterm Infants
GATK	Genome Analysis Toolkit
GFR	Glomerular Filtration Rate
GLP	Glucagon-Like Peptide
GLUT	Glucose Transporter
GVCF	Genomic Variant Call Format file
GWAS	Genome-wide Association Studies
HF	High Flow
HFOV	High Frequency Oscillatory Ventilation
IAEA	International Atomic Energy Agency

IGF-1	Insulin-like growth factor 1
IMD	Index of Multiple Deprivation
IQR	Interquartile Range
IR	Insulin Receptor
IUGR	Intrauterine Growth Restriction
IVH	Intraventricular Haemorrhage
LCPUFA	Long Chain Polyunsaturated Fatty Acid
LGA	Large for Gestational Age
LMS	Lambda, Mu, Skew method
LSOA	Lower Super Output Area
MAF	Minor Allele Frequency
MDT	Multidisciplinary Team
MODY	Maturity-Onset Diabetes of the Young
MRI	Magnetic Resonance Imaging
MS	Mass Spectrometry
NDAU	Neonatal Data Analysis Unit, Imperial College London
NEC	Necrotising Enterocolitis
NEON	Nutritional Evaluation and Optimisation in Neonates Study
NGS	Next Generation Sequencing
NHS	National Health Service of the UK
NICE	National Institute for Health and Care Excellence
NICU	Level 3 Neonatal Intensive Care Unit
NIHR	National Institute for Health and Care Research
NMR	Nuclear Magnetic Resonance
NNRD	National Neonatal Research Database
PCA	Principal Component Analysis
PDA	Patent Ductus Arteriosus
PLS	Partial Least Squares regression
PMA	Postmenstrual Age
PN	Parenteral Nutrition
PQN	Probabilistic Quotient Normalization
PreCES	Preterm Contraction of Extracellular Spaces
PRISMA	Preferred Reporting Items for Systematic Review and Meta-Analyses
QC	Quality Control
RDS	Respiratory Distress Syndrome
REC	Research Ethics Committee
RMSEP	Root Mean Square Error of Prediction
ROC	Receiver Operator Characteristic

rsID	Reference Single Nucleotide Polymorphism Identity
SAM	Sequence Alignment Map file
SCAMP	Standardised, Concentrated and Additional Macronutrients in neonatal Parenteral nutrition study
SD	Standard Deviation
SDS	Standard Deviation Score
SENNAT	Southampton Electronic Neonatal Nutrition Assessment Tool
SGA	Small for Gestational Age
SHAP	Shapley additive explanations
SITAR	SuperImposition by Translation And Rotation method
SPIN	Standardising Preterm Infant Nutrition Study
SPND	Southampton Preterm Nutrition Database
SQL	Structured Query Language
SSV6	Agilent SureSelect Human All Exon V6 probes
TBW	Total Body Water
TEA	Term-Equivalent Age
TeCES	Term Contraction of Extracellular Spaces
TSP	Trimethylsilylpropanoic Acid
UHS	University Hospital Southampton NHS Foundation Trust
UK-NICM	United Kingdom Newborn Infant Close Monitoring Growth Chart
USN	Unique Study Number
VCF	Variant Call Format file
VIF	Variance Inflation Factor
WHO	World Health Organisation

Chapter 1 Introduction and Literature Review

This chapter introduces the importance of growth in preterm infants and sets out the current research landscape addressing the targeting, monitoring and management of growth in preterm infants. Sections of this chapter are adapted from a review article written by me and published in Archives of Disease in Childhood: Fetal and Neonatal Edition in 2022 (Appendix 3).[3]

A significantly preterm infant enters the world profoundly unprepared for life outside the womb. During the months between birth and discharge home, he or she must survive in this hostile environment, is likely to contend with multiple medical complications and is expected to double, triple or quadruple in weight. Promoting healthy growth in the context of this adversity presents a challenge to those providing care to such infants. This thesis aims to explore the factors which influence the growth of preterm infants and seeks to identify clinical and nutritional measures which will enhance the monitoring and management of growth, not only in terms of weight gain but also bearing in mind healthy head growth and body composition. The later chapters explore the emerging role of metabolomics and genomics in personalising nutritional management. This introductory chapter makes the case for the importance of growth in the care of preterm infants and describes the current understanding of how growth should be monitored and managed on the neonatal unit.

1.1 The Importance of Growth for the Preterm Infant

Preterm infants are exceptionally prone to poor growth. The steady supply of nutrition provided by the placenta is abruptly interrupted at birth, the protective thermoneutral *in utero* environment is replaced by the harsh conditions of the neonatal intensive care unit and the infant is exposed to the energy demands of breathing and thermal regulation. In addition, preterm infants (especially the most premature) can be expected to experience significant medical complications of prematurity, with respiratory distress syndrome, chronic lung disease, necrotising enterocolitis and sepsis having especially profound impacts on growth.

These factors present significant challenges to the neonatal clinician. Achieving good growth during the neonatal period requires efforts to sustain adequate nutrition to promote healthy growth whilst mitigating other risks, especially the risk of necrotising enterocolitis, a devastating gastrointestinal disease which is a significant cause of mortality in very preterm infants. Despite a

significant body of research over at least half a century, the factors which influence the growth of preterm infants remain incompletely understood.

The Association of Growth with Neurodevelopmental Outcomes

Very preterm infants (i.e. infants born before 32 weeks completed postmenstrual age) are at high risk of adverse neurodevelopmental outcomes. The seminal EPICure study in the mid-1990s identified that more than a fifth of the most premature infants had severe neurodevelopmental delay at school age, with a majority exhibiting some delay and 12% meeting diagnostic criteria for 'disabling cerebral palsy'.[5] These problems persisted into adulthood.[6] Further studies into the outcomes of this cohort identified wide-ranging developmental problems, spanning behavioural,[7] psychiatric[8] and autistic[9] domains. During the intervening time period there have been significant improvements in survival after very preterm birth. However, improvements in neurodevelopmental outcomes have failed to keep pace. In fact, recent data from EPICure 2 (a study investigating a new cohort of infants born in the mid 2000s) has shown that there was no improvement in neurodevelopmental outcomes despite a decade of advances in neonatal care.[10] EPIPAGE-2 was a similar exercise in France, comparing outcomes for infants born in 1997 and those born in 2011.[11] These data were more hopeful, with improvements in mortality accompanied by modest improvements in neurodevelopmental outcomes.

The persistence of adverse neurodevelopmental outcomes for very preterm infants has led to a shift in focus in the neonatal clinical and research community towards improving neurological outcomes. Several studies have demonstrated an association between growth and later neurodevelopmental status. A study by Belfort and co-workers identified that more rapid weight gain was associated with improved mental and psychomotor development during the first 18 months of life.[12] This effect was most pronounced in the smallest infants. Interestingly, they found that the critical period of weight gain was during the time from birth to term corrected age, with weight gain after term corrected age demonstrating no association with neurodevelopmental outcome. These findings were consistent with those from other groups,[13, 14] resulting in an increasingly robust collection of observational data indicating an association between improved growth during the neonatal period and enhanced neurological outcomes. Other work has demonstrated an association between early protein intake and improved mental development scores.[15] It is important to note that these data are limited to observational studies. Despite careful attempts to adjust for known confounding factors, it is difficult to absolutely eliminate the effects of residual confounding. Specifically, growth is likely to be slower in infants who experience significant comorbidities or complications of prematurity, and it is difficult to accurately capture such medical acuity as a numeric value for adjustment of models.

Metabolic Consequences of Preterm Birth

Infants born preterm are at higher risk of developing some features of the metabolic syndrome during adulthood compared to infants born at term.[16, 17] In term-born infants, low birthweight and rapid weight gain has been consistently associated with development of elements of the metabolic syndrome.[18] Studies have identified that this effect is largely mediated by rapid weight gain during the first six months of life.[19] These findings have led to concerns that faster growth during early postnatal life may increase the risk of metabolic and cardiovascular ill-health during adulthood.

These concerns have led to recent efforts to define the period of growth which may influence the increased risk of metabolic disease seen in preterm infants. Analyses from the EPICure study[20] and from a group in Newcastle[21] suggest that, in contrast to term-born infants, rapid weight gain during the first year of life seems not to contribute to later risk of signs of the metabolic syndrome. Rapid weight gain during later childhood does increase the risk of markers of the metabolic syndrome.

Taken in combination, these data suggest that the period between preterm birth and term corrected age represents a key phase during which improved growth may improve neurodevelopmental outcome. Importantly, rapid growth during this period may not exacerbate risk of later adverse cardiometabolic outcomes. It has been suggested that enhanced early growth in the preterm population (prior to the infant reaching term equivalent age) may even improve metabolic outcomes by removing the impetus for later 'catch-up growth' (recovery to genetic growth potential after a period of poor growth).[17, 22] This doctoral project focuses on the critical growth period which takes place on the neonatal unit.

1.2 Definitions of Optimal Growth

Both in research and in clinical practice, it is necessary to have a target for growth which should be achieved. Longstanding international guidelines recommend that growth of preterm infants should mimic the growth of an equivalent fetus in utero. The European Society of Paediatric Gastroenterology, Hepatology and Nutrition (ESPGHAN) states the following:

> *The major goal of enteral nutrient supply to these [preterm] infants is to achieve growth similar to foetal growth coupled with satisfactory functional development.*[23]

Similarly, the American Academy of Pediatrics (AAP) recommends:

> *Current nutritional goals for the preterm infant are to provide nutrients to approximate the rate of growth and composition of weight gain for a normal fetus of the same postmenstrual age while maintaining normal concentrations of nutrients in blood and other tissues.*[24]

This further challenge to match not only the growth but also the body composition of the equivalent fetus presents important difficulties, as explored below.

In practice, it has proven difficult to achieve postnatal growth matching fetal growth rates. Cole and co-workers assessed growth of preterm infants in England and Wales between 2006 and 2011.[25] They showed that this cohort of around 5000 infants were consistently discharged at a weight lower than would be predicted if they had maintained a fetal rate of growth. This deviation from the fetal growth trajectory was mostly attributed to slow growth during the first three weeks of life.

Published data from Southampton suggest that careful application of quality improvement efforts can achieve weight gain rates which closely mimic fetal weight gain.[26] Nevertheless, several groups have questioned whether fetal rates of growth are achievable or desirable.

Formation of Growth Charts

Growth indices of preterm infants are routinely plotted on charts, with corrected gestational age along the x-axis and weight, length or head circumference along the y-axis. These charts are critically important to the management of nutrition and growth of preterm infants, as changes in clinical management are often made in response to an infant's trajectory. Chapter 6 of this thesis uses local and national data to propose new growth charts which may provide more useful targets for growth. This section provides an overview of the charts which are currently in use, along with the benefits and downsides of each.

Charts are marked with centile lines showing the expected growth trajectory for infants depending on their size and gestation at birth. The UK Newborn Infant Close Monitoring chart for girls is given as an example in Figure 1-1 overleaf.[27]

Centile lines on these charts are usually formed with the LMS method.[28] This approach uses a series of paired measurements and gestational ages. Three parameters are then calculated as age-specific smoothed cubic spline curves describing: a Box-Cox power to remove skewness (the L curve), the median (M curve) and the coefficient of variation (S curve). The age-specific L, M and S values can then be used to calculate the measurement value (C) at a defined standard deviation score (or z-score, Z):

$$C = M(1 + LSZ)^{\frac{1}{L}}$$

These points can then be connected and smoothed to provide lines at set z-scores. As there is a predictable relationship between z-score and centile values, centile lines can be plotted. By convention, lines are drawn at intervals of two thirds of a standard deviation (0.4th, 2nd, 9th, 25th,

50th, 75th, 91st, 98th and 99.6th centiles) as in Figure 1-1. It follows that such charts are critically dependent on the measurement-gestation pairs used to create them. The following subsections set out some contrasting approaches to defining these measurements and an alternative approach which integrates LMS-derived measurements with adjustments made to reflect proposed physiological effects in early life.

[Image removed from online version due to copyright. Newborn Infant Close Monitoring growth charts can be found at https://www.rcpch.ac.uk/resources/uk-who-growth-charts-neonatal-infant-close-monitoring-nicm]

Figure 1-1. UK Newborn Infant Close Monitoring Chart for girls born preterm (Department of Health, 2009)

Using Cross-Sectional Birthweight Data to form Growth Charts

Most growth charts which are currently in widespread use were created using the LMS method with birthweight data. This means that all growth data used for creating growth centiles is derived from infants at the point of birth.

The UK-NICM chart (Figure 1-1) is derived from birthweights of preterm infants born around 1990, which were reanalysed in 2009.[27] These charts illustrate nine centile lines ranging from 0.4th to 99.6th centile along with lines at -3SD and -4SD below the centile lines. These charts are primarily used in the UK. The Fenton growth chart is commonly used in the USA and originally used a simple method of defining centiles by directly analysing birthweight data and calculating the 3rd, 10th, 50th, 90th and 97th centiles.[29] This chart was subsequently revised using a large international dataset of preterm birthweights employing the LMS method.[30]

The simplest interpretation of these charts is that they reflect the *in utero* growth of the fetus, by taking measurements of infants at the point of birth, when their size is defined entirely by their *in utero* growth and before any postnatal factors have influenced growth velocities. However, both these charts are limited by sparse information about the pregnancies included in analysis. Preterm birth is frequently associated with problems during pregnancy and with placental insufficiency. Therefore, it is at least questionable whether these charts truly reflect normal *in utero* growth patterns. Furthermore, infants included in these studies were exclusively born in developed countries and white infants are likely to be very over-represented by global standards. This led the WHO to question whether these growth charts are appropriate to assess non-Caucasian infants or those born in low- and middle-income countries.

In an attempt to address these issues, the WHO initiated the INTERGROWTH consortium.[31] This project recruited pregnant women in eight geographical areas distributed across the world. It collected fetal growth data by ultrasound measurement along with birthweight and postnatal growth information. Tabulation of health status during pregnancy allowed the team to exclude pregnancies with significant problems from their subsequent analysis. These data allowed the INTERGROWTH team to generate preterm size at birth reference charts.[32] However, it is notable that most very preterm deliveries arose from pregnancies with some risk factors and so some high-risk pregnancies were ultimately included in analysis. They excluded infants with “ultrasound evidence of FGR [fetal growth restriction] before birth”, calling into question the reliability of their charts at lower birthweight centiles. It is also important to note that the INTERGROWTH approach led to only 408 very preterm infants being included in the analysis. The number of included infants was especially low at the most preterm gestations (for example, only six boys below 25 weeks of gestation were included). These shortcomings mean that it is difficult to assess the reliability of the resultant growth standards at the earlier gestations.

Growth charts derived from birthweights using the LMS methods can be considered to be accurate reflections of the relationship between birthweight and gestational age in the cohorts they use. This means that they are well-suited to the task of defining the growth status of an infant at the time of his or her birth. If one is to endorse the current guidelines recommending close mimicking of fetal growth in ex utero preterm infants, then postnatal growth can aim to progress along the centile line at birth. However, in practice this growth is rarely achieved, and there are some physiological arguments that it should not be expected, especially during early fluid shifts. Therefore, several groups have taken alternative approaches to growth chart creation, as discussed in the following subsections.

Using Longitudinal Measurements to Form Growth Charts

The INTERGROWTH group took the view that the assumption that postnatal growth should keep pace with fetal growth is not sound. Therefore, they aimed to define centiles by repeated measures of growth in preterm infants.[33] They limited their data analysis to "healthy preterm neonates" and used a fractional polynomial model, with patient identity as a random effect to account for repeated measures, to define the 3rd, 10th, 50th, 90th and 97th centile lines.

Growth charts for other phases of growth (i.e. from birth at term to adulthood) have often been created from repeated measurements of individuals over time. This is a simpler proposition than an equivalent exercise in the preterm infant as each infant shares a starting point (birth at term). Charts created in this way for preterm infants are difficult to conceptualise; values at each gestational age are derived from birthweight measurements for some infants and weight values during later phases of growth for other infants. The INTERGROWTH project partially addressed this by testing gestational age at birth as a determining factor but found this to be not significant. Again, this should be treated with caution in the case of very preterm infants, as such infants made up only a very small proportion of the infants included in these analyses.

There is ongoing scepticism about the expected growth patterns implied by these charts. For example, is it reasonable to expect that optimal growth of a newborn infant born weighing 1kg at 30 weeks of gestation is the same as a five-week-old infant who had been born at 25 weeks of gestation and had reached a weight of 1kg after five weeks of life? Whether derived from birthweights or from repeated measures during growth, current growth charts imply that their growth should be identical, whereas in practice they are usually different. These questions have led to interest in personalised growth charts which can integrate gestational age at birth and current corrected gestational age as separate predictors.

Personalised Growth Charts

With the increasing use of web-based resources and widening access to app-creation tools, the possibility of integrating clinical data and responsively producing growth targets has become feasible. A research group led by Christoph Fusch has carried out investigations into the physiological changes after birth and have combined these with conventional growth data to create individualised growth charts.[34] There is a well-recognised pattern of early weight loss expected in term-born infants, which is predominantly made up of loss of body water (termed the "term contraction of extracellular spaces" or TeCES). This group propose that an equivalent body water loss can also be expected in preterm infants (PreCES). This presumption relies on direct measurement of body water shifts carried out in the 1980s. More recent investigations into this phenomenon have been less conclusive. Cole and co-workers found that this effect could not be detected in extremely preterm infants (i.e. those born before 28 weeks of gestation).[25] Additionally, as conceded by the Fusch group, nutritional deficit is likely to contribute to early weight loss or slow growth. These observations led this group to create a predicted optimal growth trajectory which incorporates an initial downwards crossing of centile lines (the PreCES phase) followed by growth at the median day-specific growth velocity (based on a Fenton chart but with an adjustment factor), resulting in growth which initially drops through Fenton centiles before gradually crossing upwards, so that the weight at 42 weeks corrected gestational age is close to that of a full term infant born on a similar centile. Charts for individual infants can be generated using a web-based application. This approach elegantly sidesteps the difficulties of single charts which are designed to be used for preterm infants regardless of their gestation at birth. However, they rely on an assumption that the early centile drop observed in preterm infants is a desirable physiological adaptation, as opposed to a pathological response to malnutrition. The group partially addressed this issue by showing that growth along their predicted line led to consistent fat mass regardless of starting centile, suggesting that such growth promoted healthy body composition. However, there remains little independent evidence that this pattern of slow early growth followed by gradual catch up promotes optimal neurodevelopmental or cardiometabolic outcomes.

Growth Standard Summary

Studies examining the relationship between growth and clinical outcome have generally used the change in standard deviation score (equivalent to crossing of centiles) on birthweight-derived charts as the marker of growth. The EPICure group[20] and Belfort and co-workers used internal z-scores in their analysis.[12] Cordova and co-workers compared Fenton and INTERGROWTH z-score changes (along with some generated from earlier data) as predictors, finding that change in INTERGROWTH z-score was the best predictor of neurodevelopmental outcome.[14] In practice,

these charts are so similar that it is unlikely that the selection of one over another provides any great advantage. This doctoral project uses change in z-score based on UK-NICM values as the primary outcome measure.

1.3 Body Composition

As detailed above, the American Academy of Pediatrics explicitly states that the growth of the preterm infant should mimic that of the fetus, not only in term of weight gain but also in terms of body composition.[24] However, studies have consistently shown that the body composition of preterm infants at term equivalent age differs significantly from normal term infants.[35] Specifically, preterm infants are significantly lighter than their term counterparts but have a similar fat mass. The result of this is that their percentage fat mass (FM) is significantly higher and their lean mass is lower than a term-born infant of the same weight. There is some evidence that fat free mass (FFM) gains are associated with improved neurodevelopmental outcomes in preterm infants, whereas fat mass gains do not contribute to this improvement.[36] Additionally, subcutaneous fat is decreased compared to the term infant, whereas intra-abdominal fat mass is increased.[37] Taken in conjunction, these findings suggest that the neurodevelopmental benefits of weight gain may be mediated primarily (or exclusively) by gains in fat free mass, with the result that even greater emphasis should be placed on the goal of reproducing fetal composition changes so that the composition of the preterm infant matches more closely with a term-born infant at term equivalent age.

Furthermore, it is well-recognised that abnormal body composition contributes significantly to the risk of cardiometabolic disease in later life. The causes of excess cardiometabolic ill-health in ex-preterm infants remain ill-defined.[16] A wide range of candidate mechanisms have been postulated, ranging from effects of the hypothalamic-pituitary-adrenal axis to telomere shortening to epigenetic effects, but without substantial evidence for any.[17] To date there has been little investigation into whether body composition aberrations mediate the excess cardiometabolic illness of ex-preterm children and adults. This is likely due to the difficulty of assessing body composition in infants, combined with the very long follow-up periods required to detect cardiometabolic illness. Nevertheless, it may be hoped that normalising body composition may help to address the cardiometabolic consequences of preterm birth.

Measuring Body Composition

In practice, it remains very difficult to track body composition in preterm infants, especially during their first weeks when they require intensive or high-dependency care. Studies into the body composition of term infants (or preterm infants at term equivalent age) have used whole body

MRI,[37-39] air-displacement plethysmography[40] or dual-energy x-ray absorptiometry (DXA)[41, 42] with significant success. These methods are validated and incur a low-to-medium cost after initial outlay. However, there has been limited success in using these methods on preterm infants, with only plethysmography using the Pea Pod system showing promise in moderately preterm infants.[43] These limitations mean that such approaches are well-suited to research assessment of differences between body composition of term-born infants and preterm infants at term equivalent age. However, they cannot be easily applied to tracking body composition during the period from preterm birth to term equivalent age.

Two methods have been investigated for their use in preterm infants: stable isotope dilution and bioelectrical impedance analysis (BIA). Dilution studies aim to measure total body water. In essence, this technique involves the introduction of an exogenous substance at a known quantity into the pool of total body water. The ideal properties of this substance would include non-toxicity, not binding to protein and being freely filtered in the kidney. The concentration of this substance in the urine will then be inversely proportionate to the total body water volume. Body water percentage can then be calculated using the infant's total body weight. Normative values can be used to estimate fat free mass (and, by extension, fat mass and percentage fat mass). Early dilution studies used antipyrine, but this substance had several features rendering it a suboptimal choice for the dilution method.[44-48] It slightly binds to protein, has some diuretic effects and is sometimes altered by metabolic processes rather than being excreted in the urine unaltered. These shortcomings led to the adoption of stable isotopes of water (deuterium oxide, 18-oxygen water and tritiated water) as the preferred substances for such studies. These substances have no physiological effects at the relevant doses and are excreted unaltered in the urine. Guidance from the International Atomic Energy Agency confirms that they are safe for use in human studies.[49] The main downside of using these water isotopes is the complexity and expense of measuring their concentration in urine, along with the high cost of formulating products suitable for intravenous or oral delivery to preterm infants. Whilst other methods were used in the past, mass spectrometry is the simplest, most reliable and cheapest approach in the 21st century.

BIA has been used to assess preterm infants. However, efforts to validate the resultant data as a marker of body composition have proven troublesome, with models of fat mass including BIA values failing to outperform models only using simple anthropometric data.[50]

The effect of all these factors is that stable isotope dilution is the only feasible research method to repeatedly assess body composition (using total body water as a proxy) in the preterm infant.

Developing a Cotside Marker of Body Composition

The expense and complexity of stable isotope dilution renders it impractical as a routine cotside measure of body composition in clinical practice. Studies in older infants have shown that limb circumference measurements are a reliable proxy for body composition in this group.[51] These findings have led to interest in the use of limb circumference measurements as a convenient proxy for body composition in the preterm infant. Work in Southampton preceding this doctoral project has demonstrated that limb circumferences grow in a predictable pattern in this group and that they contain information which is distinct from simple measurements of weight, length and head circumference.[52] The body composition arm of this doctoral project aims to assess whether limb circumference measurements reliably predict body composition as measured by stable isotope dilution.

1.4 Nutritional Requirements

Notwithstanding difficulties in defining optimal growth, in practice it is necessary to set target nutrient intakes which may then be adjusted for individual infants based on their clinical status and growth trajectory. Chapter 4 of this thesis includes analysis of nutrient intakes associated with faster or slower growth Table 1-1 presents several international guidelines for nutrient intake along with the nutritional provision of some typical feeds and parenteral nutrition products. In a widely-used handbook of preterm nutrition, Koletzko and co-editors provide guidance for the requirements of enterally fed preterm infants.[53] These recommendations have been widely adopted, although they provide only sparse guidance regarding nutritional intakes during parenteral nutrition. ESPGHAN also sets out guidance for enteral nutrient intake, which are generally in good agreement with those of Koletzko.[54]

Specific guidance for nutrient intakes for parenteral nutrition is available from ESPGHAN[55] and from the National Institute for Health and Care Excellence (NICE).[54] Some of these recommendations are for very broad ranges (e.g. 75-120kcal/kg/day), limiting their use when setting nutritional strategies.

Table 1-1 also describes the nutrient content of some typical enteral and parenteral feeding products. Certain facts are immediately apparent from these data. They show that feeding with breastmilk alone will not meet enteral nutrition requirements at the standard feeding volume of 150ml/kg/day. This will be explored further in the Breastmilk Fortifier section below. There is only one licenced commercially available parenteral nutrition product available in the UK. Many neonatal units order premade parenteral nutrition products which have been formulated to their needs (for example, the Southampton Concentrated Preterm formulation in Table 1-1). Alternatively (or as a supplementary approach to these products) bespoke parenteral nutrition

can be made to meet the specific needs of individual infants. Practical delivery of PN will be considered below.

Table 1-1. Guideline recommendations for enteral and parenteral nutritional intakes for preterm infants along with the range nutritional intakes from some typical feeds (at 150ml/kg/day) and a commercially available parenteral nutrition product.

		ENTERAL FEEDING GUIDELINES		ENTERAL FEEDS					PARENTERAL FEEDING GUIDELINE		PARENTERAL FEEDING PRODUCTS	
		Koletzko et al.[53]	ESPGHAN[23]	Breastmilk (AAP[24])	Fortified Breastmilk	Preterm Formula	Preterm Follow-on Formula	Term Formula	NICE54	ESPGHAN[55]	Commercially Available Numeta Preterm PN[56]	Southampton Concentrated Preterm PN
Nutrient	**Unit**	(per kg/day)	(per kg/day)	(per 150ml)	(per 150ml)	(per 150ml)	(per 150ml)	(per 150ml)	(per kg/day)	(per kg/day)	(per 100ml)	(per 100ml)
Energy	kcal	110-130	110-135	~100	120-130	120	110	100	75-120	90-120	90	104.7
Macronutrients												
Protein / Amino Acids	g	3.5-4.5	4-4.5 (<1kg body weight) 3.5-4 (>1kg body weight)	~1.4	3-4.5	4-4.5	3	2-2.2	3-4	2.5-3.5	3.1	3.43
Lipid	g	4.8-6.6	4.8-6.6	~5	5-6	5.5-6	5.5-6	5	3-4	3-4	2.5	3.5
Carbohydrate	g	11.6-13.2	11.6-13.2	~12	13-16	12-13	10.5-11.5	11-12	9-16	11.5-14.4	13	14
Micronutrients												
Sodium	mg	69-115	69-115	~20-40	70-90	80-100	40-50	30-50	Responsive*	46-115	51	115
	mmol	1.2-2	1.2-3	~1-2	3-4	3.5-4.5	1.8-2.4	1-2		2-5	2.2	5
Potassium	mg	78-195	66-132	~60-80	120-150	120-170	120-130	100-140	Responsive*	78-117	78	78
	mmol	2-5	1.7-3.4	1.5-2	2.5-3.5	3-4.5	3-3.5	2.5-3.5		2-3	2	2
Calcium	mg	120-200	120-140	~30-40	130-150	150-180	125	70	60-80	64-140	50	40
	mmol	3-5	3-3.5	0.7-1	3.5-4	4-4.5	3.1	1.7	1.5-2	1.6-3.5	1.3	1
Phosphorus	mg	60-140	60-90	~20	80-90	100-120	70-75	40-50	60	50-108	40	78
	mmol	2-4.5	2-3	0.6	2.5-3	3-4	2-2.5	1.3-1.6	2	1.6-3.5	1.3	2.5
Iron	mg	2-3	2-3	~0.05-0.14	0-3**	2-2.5	1.1-1.8	0.5-0.8	0	0	0	0
Zinc	mg	1.4-2.5	1.1-2	~0.15-0.5	1-2	1.5-2	1.3	0.8	NG	0.4-0.5	†	0.5
Vitamin A	µg RE	400-1100	400-1000	~50-90	400-600	500-550	100-150	90	NG	700-1500	†	100
Vitamin D	IU	400-1000	800-1000	~0.05	250-300	180-200	100	90	NG	200-1000	†	87
Vitamin E	mg α-TE	2.2-11	2.2-11	~0.5-1.2	5-7	5.5-7	1.3-3	1.7-2.1	NG	2.8-3.5	†	NG
Vitamin K1	µg	4.4-28	4.4-28	~0.3-0.5	10-15	9-10	6-9	5-7	NG	10	†	NG
Choline	mg	8-55	8-55	NG	3-3.5	30-40	35	33	NG	NG	NG	NG
DHA	mg	55-60	12-30	NG	2-2.5	30-40	30	25	NG	NG	NG	NG

RE, retinol equivalents; α-TE, α-tocopherol equivalents; NG, no guidance given or not defined in product information

* Guidance recommends adjusting intake based on clinical sampling. ** Some fortifiers do not contain iron. † Must be added to product.

Methods to Determine Nutritional Requirements

A number of methods have been used to attempt to define the nutritional requirements of the preterm infant. They can be divided primarily into the physiology-based factorial method and recent experimental approaches.

The Factorial Method

The factorial method aims to define nutrient requirements by estimating three components: the growth component (i.e. the amount of the nutrient accreted by a normally growing fetus), losses or energy expenditure of the nutrient, and efficiency of nutrient absorption (in the case of enteral feeding).[57] This approach has been most comprehensively applied to protein (or amino acid) metabolism.

Studies of the composition of miscarried fetuses were summarised to form a "reference fetus" several decades ago.[58] Examination of the change in the body content of a nutrient can then be used to derive the growth component of the factorial method. These data can be combined with studies into nutrient losses, absorption efficiency and energy expenditure to estimate nutritional requirements.

Experimental Approaches to Macronutrient Intakes

Several recent trials have randomized preterm infants to receive differing nutritional regimens. The NEON (Nutritional Evaluation and Optimisation in Neonates) study assessed the impact of immediate or incremental increases in amino acid intakes and lipid emulsions during the first few days of life.[59] This study did not identify any differences in its primary outcomes of non-adipose mass and intrahepatocellular lipid at term equivalent age. The group with higher early provision of amino acids had lower growth of head circumference (8mm adjusted mean difference, $p=0.02$). The study was criticised for the minimal effect of allocation on actual macronutrient intake over the course of neonatal stay.[60]

The SCAMP (Standardised, Concentrated and Additional Macronutrients in neonatal Parenteral nutrition study) Trial randomized infants to a control parenteral nutrition product or an intervention product which delivered more amino acid, lipid and glucose with a resultant higher energy provision throughout the period of parenteral nutrition.[61] The group receiving more macronutrients had faster head circumference growth to day 28 of life (the primary outcome, 5mm mean difference, $p<0.001$) and this persisted to 36 weeks postmenstrual age. Weight was unaffected as were all other clinical outcomes.

Similarly, a Norwegian group demonstrated an increased growth velocity to 36 weeks postmenstrual age in response to an enhanced supply of amino acid, lipid and energy (17.4 vs 14.3 g/kg/day, p<0.001).[62] A Dutch group randomized infants to a range of intakes of amino acids and types of lipid, with one group (high amino acids and mixed lipid emulsion) demonstrating greater weight gain at two years corrected age, but without any differences in neurodevelopment (the primary outcome).[63]

Taken together, these studies suggest that higher rates of nutritional intake lead to faster early growth, although an impact on neurodevelopment remains unproven.

Experimental Approaches to Specific Micronutrients

Several micronutrients have recently come under scrutiny, either for a possible general effect on growth or as targets for improving other specific clinical outcomes.

Trials of high doses of vitamin D supplementation have been shown to improve radiological markers of bone mineralization and to increase weight gain (13.6 vs 16.4g/day, p<0.01) and length gain (0.69 vs 0.79cm/week, p=0.02)[64]. Vitamin A supplementation has been shown to improve a marker of retinal function.[65]

Choline and docosahexaenoic acid are implicated in phosphatidylcholine metabolism and are found in high concentrations in fetal plasma, falling rapidly after preterm birth. A small trial has shown that supplementation with choline is practical and can restore plasma choline to near fetal concentrations although further work is needed to assess any potential impact on growth or neurodevelopment.[66]

A Cochrane review of long chain polyunsaturated fatty acid (LCPUFA) supplementation in preterm infants found no proven effect (2260 subjects in 17 trials).[67] Zinc has also risen to prominence in recent years, with a Cochrane review suggesting that enteral supplementation with zinc is likely to improve growth and reduce mortality.[68] This is particularly significant given that commonly available parenteral and enteral nutrition products typically deliver markedly insufficient amounts of zinc (Table 1-1), and zinc deficiency is common in preterm infants.[69]

Metabolic Tolerance

Metabolic disturbance is more common in the most preterm infants and in those with intrauterine growth restriction.[70] The substantial energy needs of the preterm infant necessitate the delivery of a significant load of carbohydrate, protein and lipid. However, these infants are prone to hyperglycaemia in the early neonatal period, which has been associated with an increased risk of death, poor growth and most major morbidities associated with prematurity,

although it is difficult to prove a causal link given the presence of likely uncontrolled confounding factors.[71] Technological advancements in continuous glucose monitoring have been shown to improve glycaemic control but the impact on outcomes remains uncertain.[72] Similarly, hypertrigyceridaemia is common at intravenous lipid delivery levels meeting nutritional requirements and is associated with poorer clinical outcomes.[73, 74]

Parenteral Nutrition Limitations

Current formulations of parenteral nutrition often do not meet target or recommended nutrient requirements, especially for micronutrient minerals such as calcium and phosphate. In part, this is due to concerns about stability of these substances in solution and the possibility of precipitation. Studies continue in this area, especially as there is a pressing need to optimise calcium and phosphate delivery to prevent metabolic bone disease of prematurity.[75]

Sepsis and Inflammation

Preterm infants frequently experience episodes of inflammation, both from infections and from other causes, including surgical interventions. Infection is common, with around 10% of preterm infants experiencing late onset infection (Vermont Oxford Network VLBW cohort).[76] Acute inflammation profoundly alters the metabolic state of the preterm infant, driving catabolism, insulin resistance and suppression of growth factors such as IGF-1.[77] This is likely to lead to impaired nutrient metabolism with usual or increased nutrition in this context likely to drive hyperglycaemia and hypertriglyceridaemia without contributing to growth. This theoretical problem is reflected in well-established findings in critically ill adults and children, where early aggressive parenteral nutrition during acute illness is deleterious.[78, 79]

Fluid Restriction

Newborn infants have a limited capacity for diuresis and so fluid intake is often limited during the first few days of life. In addition, fluid restriction may be part of medical management, for example in the presence of patent ductus arteriosus. Even once total fluid restriction is relaxed, there is often a period during which breastmilk replaces much more energy-dense parenteral nutrition. These multiple restrictions of fluid intake inevitably limit delivery of nutrition. These difficulties may be addressed by strategies including increasing the concentration of parenteral nutrition (as recommended by NICE)[80] and by earlier initiation of breastmilk fortification.

Nutritional Content of Breastmilk

Mother's own breastmilk provides substantial benefits to the preterm infant and is recommended as the ideal basis for enteral feeding.[23, 24] Using breastmilk in preference to formula also significantly reduces the risk of necrotising enterocolitis. However, breastmilk alone cannot provide adequate nutritional intakes and hence multicomponent fortification has been widely adopted. A Cochrane review in 2004 recommended routine fortification as it improves short-term growth and identified no increase in adverse events related to its use, albeit with insufficient long-term follow-up data to reach a conclusion on neurodevelopmental outcomes.[81]

Breastmilk fortifier is typically formulated using extensively hydrolysed cow's milk protein. During the last decade, milk fortification products based on donated human milk have been developed. Initial studies establishing the use of the first of these products were troubled by design flaws and there is ongoing controversy surrounding its potential benefits and costs.[82] A recent systematic review and meta-analysis concluded that there is a suggestion of decreased risk of necrotising enterocolitis with human milk-based fortifier but that the overall quality of evidence is low and so its routine use cannot currently be recommended.[83]

Individualised Breastmilk Fortification

Maternal milk constitution and infant nutritional requirements are both highly variable. Therefore, attempts have been made to personalise breastmilk fortification to adjust breastmilk nutritional contents to prespecified values[84] or in response to infant blood urea level,[85] or both.[86] Altering fortification in response to blood urea has shown promise in improving growth,[85] although there was significantly higher protein provision to the intervention group, meaning that it is difficult to know whether increased protein or personalisation per se was the important factor. A Cochrane review identified that targeted fortification improved weight, length and head growth during initial neonatal stay but that there was insufficient evidence for other outcomes.[87]

Donor Breastmilk

Milk banking has increased the availability of donated breastmilk throughout North America and Europe during the last decade. A recent Cochrane review addressed many questions relating to the relative safety and efficacy of fortified donor breastmilk compared to preterm formula.[88] Weight and length gain were better in the formula-fed group, with no difference in head growth or neurodevelopmental outcomes. Necrotising enterocolitis was more common in the formula fed group (risk ratio 1.87, 95%CI 1.23-2.85).

1.5 Local Nutrition Practice

Nutritional practice on the Southampton Neonatal Unit is underpinned by a comprehensive nutrition guideline (Appendix 5). This guideline aims to meet published nutritional requirements using feeding products detailed in Table 1-1 and taking into account the challenges described above. This approach relies on an initial nutritional risk assessment, based primarily on gestation and weight at birth. For high risk very preterm infants, there is an initial period of relatively restricted parenteral nutrition, limited by fluid restriction and metabolic intolerance. This is followed by a period of increasing parenteral nutrition to achieve nutrient requirements within the first five to seven days of life. In parallel to this process, maternal or donor breastmilk is gradually introduced. As breastmilk feeds are increased, parenteral nutrition must be reduced to avoid excess fluid delivery. This transition phase is accompanied by a risk of decreased nutrient provision. To mitigate this risk, parenteral nutrition is provided in a concentrated form (see Table 1) and breastmilk fortifier is introduced when feeds reach 100ml/kg/day, prior to cessation of parenteral nutrition. Figure 1-2 shows the progression of feeds and the associated nutrition intake in a high-risk infant who is subjected to the nutritional guidance without any complications. Despite these measures, intervening illness, requirements for further fluid restriction, intolerance of glucose or fat and difficulties maintaining central venous access sometimes necessitate reduction in nutritional intake. Infants included in this doctoral project are managed according to these principles. The degree of success in achieving nutritional targets is discussed in the database formation chapter below.

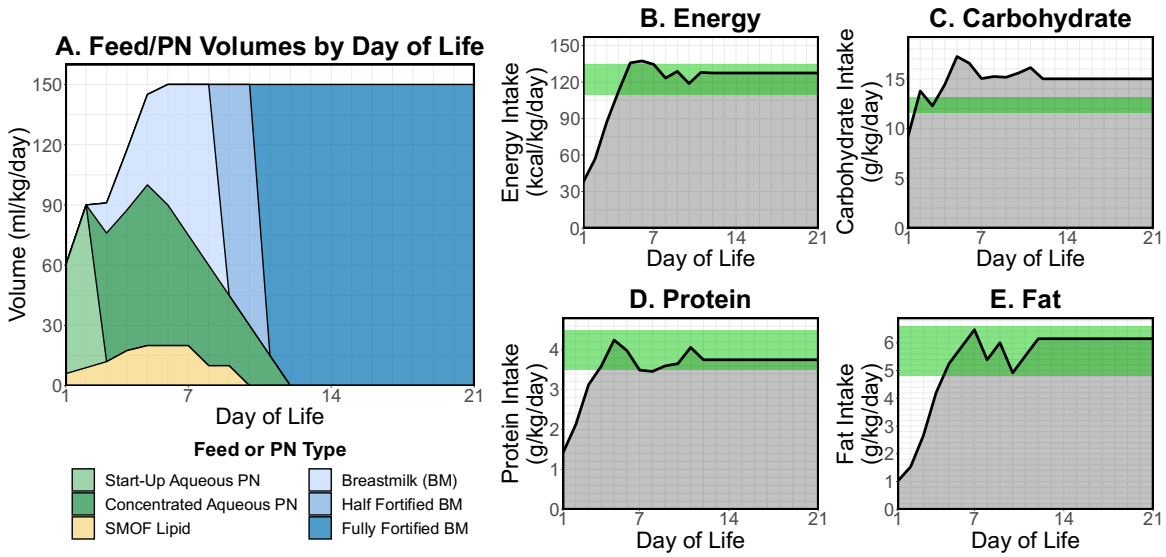


Figure 1-2. A. Changing volumes of feeds and parenteral nutrition products and B-E. Macronutrient intakes in infants following the Southampton nutritional guideline for high risk infants (green bands indicating target nutrient intake as defined by Koletzko[53], see table 1-1).

1.6 Protein and Urea

There has long been interest in exploring the usefulness of measurements of serum or urine urea in assessing the adequacy of nutrition, and especially the question of whether sufficient protein is being provided. A study by Polberger and co-workers established a correlation between protein intake and serum urea.[89] This study examined the response of 28 very low birthweight (<1500g) fully enterally fed infants to differing protein provisions. It found that there was a strong correlation between protein provision and serum urea at the end of a study period of around 28 days. Another more recent interventional study examined the response of 77 infants to two different formulas containing differing protein loads over a longer period of time and found a similar correlation.[90] A third study used observational data and multiple linear regression to investigate the association between amino acid intake and serum urea in over 600 infants receiving parenteral nutrition during their first week of life, and found a significant association.[91] These data have led to interest in using serum urea as a marker of protein intake adequacy. There are difficulties with this approach, especially as serum urea level may be influenced by other clinical factors, especially renal impairment. However, this topic provides an interesting line of enquiry in terms of forming a complete package of nutritional monitoring for the preterm infant. The potential for local data to provide further insights on the relationship between protein intake and urea levels is explored later in Chapter 7 of this thesis.

1.7 Genomics and Metabolomics

Genome-wide association studies (GWAS) have identified genetic loci which have a significant impact on birthweight (and presumably *in utero* growth).[92-94] Certain genes have known effects on metabolic processes relevant to growth, including ADCY5's implication in glucose regulation.[93] Some of these genes have been shown to influence birthweight in the preterm population.[95] However, considering the different environmental exposures of the preterm infant compared to the fetus *in utero*, it seems likely that genetic influences on growth will be distinct during postnatal growth. GWAS data have not been gathered to assess the effect of genetic influences on the growth of preterm infants *ex utero*. Exome sequencing has proven fruitful in identifying genetic determinants of other elements of preterm health, including respiratory pathology.[96] Therefore, collection of exome data may help to explain some of the variance in growth during the neonatal period, as explored in Chapter 10 of this thesis.

As set out above, preterm infants frequently exhibit disturbances of glucose metabolism at glucose delivery rates which are required to provide adequate energy intake. Insulin, c-peptide and glucagon-like peptide (GLP) have been implicated in this phenomenon and have been shown to display aberrant levels during the preterm period.[97] Animal models have confirmed changes in

expression of multiple factors linked to glucose metabolism in primate fetuses, including insulin receptor (IR)-β, Akt and glucose transporters (GLUT-1 and GLUT-4).[98] Targeted analysis of polymorphisms in these genes may reveal a role in the glucose intolerance seen in preterm infants.

In practice, it is not feasible to collect sufficient genomic information during this doctoral project to reach firm conclusions about these target genes (let alone GWAS approaches). Further background to the current state of genomic investigations into nutrition of preterm infants, along with the methods used in genomic analysis, is given at the beginning of the genomics chapter of this thesis.

There has recently been increasing interest in metabolic profiling of preterm infants. These infants undergo an abrupt change in their environment, from fetal life to *ex utero* life. It may be expected that adaptation to this environment requires significant changes in metabolism. Early published data suggest that there are measurable changes in the urinary metabolome over time in very preterm infants.[99] These data were gathered as part of an interventional study comparing different nutritional regimens. Metabolic differences could not be detected between the intervention and control groups but there were changes during the transition to postnatal life which were influenced by the gestational age of the infants. Interestingly, there were also significant differences between the metabolome of growth-restricted and appropriately grown infants.

Specific investigations into the metabolomic profiling of preterm infants are described in the metabolomics chapter of this thesis, along with an exploration of the methods used to perform metabolomic analysis.

1.8 Conclusion to the Literature Review

This literature review sets out the current understanding of the consequences of poor growth, the ideal patterns of growth in the preterm infant, body composition aberrations seen in this population and consensus nutritional requirements. It highlights the importance of good nutrition and growth to the developmental trajectory of infants born preterm and explores the current evidence for the critical growth periods associated with impacts on later outcomes of interest, both developmental and cardiometabolic. Additionally, some challenges in neonatal nutrition practice and brief outlines of the roles of genomic and metabolomic factors are described. This knowledge underpins the work of the doctoral project, providing the basis for the chosen outcome measures, justification for the methods used and the basis for considerations of the underlying causes of poor nutrition and growth.

Chapter 2 Aims and Hypothesis

This chapter sets out the hypothesis to be tested and details the aims of this thesis.

The over-arching aim of this thesis is to identify clinical and nutritional strategies which optimise the growth of preterm infants. I hypothesise that the growth and developing body composition of individual infants will depend upon their demographic features, their genomic make-up and metabolomic responses, the complications of prematurity which they suffer and the nutritional management to which they are subjected. The goal of this doctoral project is to provide scientific basis for management of growth in preterm infants, bearing in mind nutritional, clinical, genomic and metabolomic factors. This requires detailed analysis of their phenotype and genotype alongside growth tracking. The ultimate goal of this research programme is to provide a toolkit to guide personalised nutritional and growth management of preterm infants. The genomic and metabolomic influences on growth in preterms are not sufficiently well-understood at present to be used in clinical practice; the work in this thesis relating to those elements focuses on building capacity to answer those questions in the future.

Figure 2-1 illustrates an overview of the data sources for this thesis, the analyses to be undertaken and the contribution of each of these workstreams to the eventual conclusions of the thesis.

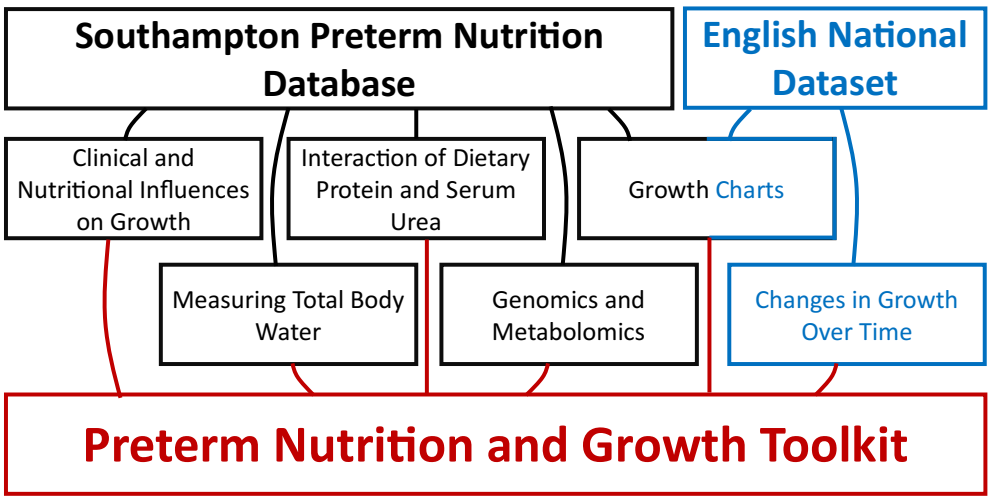


Figure 2-1. Overview of data sources, analyses and outcomes from this thesis.

The detailed aims of each element of this thesis are given within the individual chapters. The general aims of each of these elements are as follows:

1. To create a sustainable database containing detailed information on the demographic details, clinical care and nutritional intakes of preterm infants cared for at University Hospital Southampton (see chapter 3).
2. To use the data held in this database to discover the clinical features and management approaches which are associated with faster or slower growth of the preterm infant (chapter 4).
3. To determine the value of measuring serum urea as a marker of adequate protein intake (chapter 7).
4. To use national growth data to discover changes in the growth of preterm infants in England over the last decade (chapter 5).
5. To use national and local growth data to generate charts to track the adequacy of growth of the preterm infants in real time, in order to contextualise current and achievable patterns of growth (chapter 6).
6. To review the current knowledge of the body composition of preterm infants in terms of total body water and to take serial measurements of total body water in preterm infants to establish the normal pattern of total body water change during the growth of preterm infants (chapters 8 and 9).
7. To set up an infrastructure to enable the collection of specimens for genomic analysis of preterm infants and to begin to assess associations between genomic markers and responses to nutritional intake (chapter 10).
8. To establish an infrastructure to enable to collection of specimens for analysis of the metabolomic responses and to identify markers of metabolic maturation during the growth of preterm infants (chapter 11).

The following outputs are expected:

1. A detailed understanding of national and local growth of preterm infants (published as Young at al. *Changes in the growth of very preterm infants in England 2006-2018*. Arch Dis Child Fetal Neonatal Ed. 2022, Appendix 4).[4]
2. An overview of current best practice to manage the growth of preterm infants (published, in part, as Young et al. *Optimising growth in very preterm infants: reviewing the evidence*. Arch Dis Child Fetal Neonatal Ed. 2022, Appendix 3).[3]
3. Growth charts which can be used to track growth and identify poor growth (published as Young et al. *Generating longitudinal growth charts from preterm infants fed to current recommendations*. Arch Dis Child Fetal Neonatal Ed. 2020;105(6):646-51, Appendix 1).[1]
4. Guidance on the target range of macronutrients to optimise the growth preterm infants.
5. An assessment of the current understanding of body composition in growing preterm infants, along with measurements of total body water during growth (published as Young

et al. *Total body water in full-term and preterm newborns: systematic review and meta-analysis*. Arch Dis Child Fetal Neonatal Ed. 2021;106(5):542-8, Appendix 2).[2]

6. A pathway to generate genomic and metabolomic profiling of preterm infants.

Taken together, these outputs will provide the scientific basis to inform a suite of tools allowing clinicians and families to guide the nutritional care of preterm infants.

Chapter 3 Shared Methods

This is a generic methods chapter. It begins with a brief description of each of the cohorts of infants used in this project. There follows a report of my creation of the Southampton Preterm Nutrition Database which underpins much of the work of this thesis. It concludes with a description of the methods used to model growth in individuals and populations and then an outline of the regression and machine learning methods used.

The work in this chapter was completed by me and I am grateful to members of the Data Science cross-cutting theme of the NIHR Southampton Biomedical Research Centre and especially Dr Hang Phan for their advice. The involvement of Dr Phan is described in more detail later in the chapter.

3.1 Local and National Cohorts of Preterm infants

This project is primarily based on analysis of the detailed clinical, nutritional and growth data collected for infants born before 32 weeks of gestation and cared for in the neonatal unit of University Hospital Southampton NHS Foundation Trust. Data were obtained concerning the growth of a much larger group of infants born in England over a four year period from 2014 to 2018 in order to understand patterns of growth in this population of infants at a national level.

3.1.1 Southampton Cohort

Study Design and Recruitment

Detailed data for infants born in Southampton were initially gathered as part of the Standardising Preterm Infant Nutrition (SPIN) study (national research ethic committee (REC): 11/SC/0191) which started in 2009. Data collection continued as part of the Growth Assessment of Preterm Infants (GAP) Study. Both studies were initiated prior to the start of this doctoral research project and were approved by a national REC (Oxford A 14/SC/1275). The most recently approved protocol for the GAP Study is given as Appendix 6.

The latter study was initiated to gather detailed clinical data alongside daily nutrient intakes and growth parameters including limb circumference measurements. This study has been subject to amendments. I contributed to Substantial Amendment 5, which introduced consent for genomic

and metabolomic data. I managed Substantial Amendment 6 which was approved in July 2021 and facilitated MRI scans of preterm infants (not considered in this thesis).

Infants were recruited to the study after informed written consent was given by their parents. Infants were eligible if they were born in University Hospital Southampton (UHS) or were transferred to UHS for ongoing care. During the first years of the study, only infants born before 30 weeks of gestation were included. After Substantial Amendment 5 during the course of the doctoral research period, infants born before 32 weeks could be recruited. Infants transferred to UHS for specialist surgical care and those with identified genetic conditions affecting growth were excluded. A detailed description of the Southampton cohort of infants is given in Chapter 4.

3.2 Database Curation

The concept of a comprehensive neonatal nutrition database

Observational approaches to exploring the influences on the growth of preterm infants require the marshalling of a large amount of complex data. It is likely that a large number of variables will interact to influence growth. Broadly, these variables can be grouped into seven categories:

- Demographic details
- Perinatal factors, medications and disease states (such as diagnosis of chronic lung disease or necrotising enterocolitis)
- Nutritional intake
- Blood test results
- Metabolomic and genomic data
- Outcome growth data

Any investigation of these variables will be facilitated by inclusion of these data in a database, with measures to permit their easy analysis and use for modelling. The following subsections will explore the reasons for including each of these data types. This will be followed by a discussion of the concepts of a relational database and an exploration of the data sources used in this project.

Demographic details

Certain demographic features are required to understand the cohort in question, especially the distribution of gestational age and birthweight (along with the derived birthweight z-score to identify the prevalence of SGA). The growth pattern of preterm infants is different (both quantitatively and qualitatively) depending on the gestation, with early weight loss seen in more mature infants and absent in the most preterm.[25] More preterm infants are smaller and lighter at term-equivalent age (TEA) than more mature infants. Therefore, inclusion of gestation at birth is

important to adjust for these effects. Term infants who are born SGA after intrauterine growth restriction (IUGR) demonstrate catch-up growth after birth.[100] This means that they grow proportionately faster than infants who experienced appropriate in utero growth. There has been much less investigation into whether preterm infants born SGA demonstrate catch-up growth during their initial hospital admission, but inclusion of birthweight z-score in a database is important to investigate and, potentially, to adjust for catch-up growth when nutritional influences are investigated.

Other demographic factors have been found to influence neonatal outcomes. For example, maternal smoking has been associated with a modest increase in chronic lung disease rates.[101] Deprivation in the area of the mother's residence increases the risk of preterm birth.[102, 103] The effect of deprivation on outcomes for preterm infants has not frequently been investigated and it may be expected that universal access to advanced healthcare without direct cost at the point of delivery would mitigate effects of deprivation on neonatal outcomes. One study of infants in France and the UK did not identify an influence of deprivation on neonatal outcomes such as mortality or bronchopulmonary dysplasia (BPD) after adjustment for other determinants.[103] However, other factors were unevenly distributed, especially premature rupture of membranes and breastfeeding at discharge. Therefore, both maternal smoking status and markers of deprivation may influence growth and should be included on a database examining early preterm growth. Maternal age has also been positively associated with weight gain velocity.[104]

When considering the growth status of an infant, it is important to pay careful attention to the methods used to estimate the date of conception (by convention, using the "postmenstrual age" corresponding to the time of the mother's last menstrual period before conception). In the UK, the overwhelming majority of pregnancies are dated using an ultrasound scan at 11-14 weeks postmenstrual age. In studies of pregnancies initiated by in vitro fertilisation (in which the precise day of conception is known), such assessments are strikingly accurate, with a 95% confidence interval of −0.419 to 0.337 days.[105] Viable fetuses during this time window grow at a very predictable rate and so dating by this method is considered a sound assessment of postmenstrual age (and therefore the growth status of the infant at birth).

Perinatal factors, medications and disease states

Several disease states have been associated with slower growth in preterm infants, including respiratory distress syndrome with invasive ventilation, patent ductus arteriosus, necrotising enterocolitis (NEC) and sepsis.[106] Observational studies seek to adjust for demographic features to draw out an independent effect of these disease states on growth, but it is difficult to exclude residual confounding. Nevertheless, it is important to include these key variables in a database examining growth. Some drugs which are commonly used during the neonatal period are also

likely to influence growth. Corticosteroids are known to impair growth in certain patient groups, although some studies have not identified an influence on preterm infants.[107] Diuretics exert a complex effect on growth, with a short-term impact on weight likely due to fluid loss, but with some evidence that weight gain is generally faster over a longer period of time.[108]

Nutritional Intake

As has been explored in the introduction to this thesis, nutrient intakes are critical to the early growth of preterm infants. In practice, preterm infants are exposed to a wide range of nutritional products and accurately capturing all daily nutrient intakes is a complex task. It is also important to capture nutritional management approach, especially as it relates to the use of intravenous nutrition and the use of infant formula and breastmilk fortifier. The practicalities of capturing these data are further explored in the description of data sources below.

Blood test results

Preterm infants undergo regular monitoring of blood biochemistry values, both through point of care blood gas testing and formal laboratory blood analysis. Some biochemical indices have been implicated in poor growth, especially hyponatraemia and anaemia. Furthermore, blood culture results may be used to assess for episodes of sepsis. The identification of episodes of sepsis in the preterm infant is complicated by a number of factors. Preterm infants typically receive multiple courses of intravenous antibiotics during their hospital stay. Antibiotics may be initiated for a host of reasons, and often a problem is retrospectively identified to be due to a non-infective cause. Positive blood cultures may also be the result of contamination of the blood sample in the absence of true infection. Therefore, inclusion of blood tests indicative of inflammation may, in combination with blood culture results, be used to identify episodes of true bloodstream infection.

Metabolomic and genomic data

The possible influences of genomic and metabolomic factors on the growth of preterm infants are explored in the introduction to this thesis These factors produce a very large amount of data and present specific data management problems as explored later in this chapter.

Growth data

Growth is described by multiple metrics. Weight is the most frequently monitored measurement in preterm infants, although there may be significant difficulties in accurately measuring weight and errors in its entry onto databases. Head circumference and body length are measured less

frequently and may also be subject to errors. Limb circumference measurements are a useful marker of body composition during childhood but have not previously been well described in preterm infants.[51]

3.2.1 Database Standards for the Southampton Preterm Nutrition Database

The following section sets out the theories and standards which were used to guide the implementation of the Southampton Preterm Nutrition Database (SPND).

Relational databases

One aim of this doctoral project was to create a relational database for preterm growth and nutrition, which was compliant with Codd's theories and with the tidy data restrictions detailed below.

A relational database is a way of storing data adhering to certain fixed rules. The concept of relational data was introduced in the late 1960s by E.F. Codd.[109] It envisaged a future in which large volumes of data would be stored in databases and would be subject to requests to retrieve and analyse specific records. It recommends the organisation of data into tables (termed "relations") in which each row ("tuple") contains a distinct individual record (with their order being arbitrary) and in which columns are in a fixed order, each containing information concerning one uniquely-named domain. Codd recommended that one column in each table should contain a primary key which uniquely identifies the corresponding row. Different relations could be linked by a key. A relational database consists of a series of relations (i.e. tables) which can be linked to each other by the use of keys. One aim of this doctoral project was to create a relational database for preterm growth and nutrition, which was compliant with Codd's theories and with the tidy data restrictions detailed below.

Tidy data

The tidy data schema is a way of managing relational data which is further constrained in the meanings of the rows and columns.[110] It refines the relational model to insist that:

1. Each variable forms a column.
2. Each observation forms a row.
3. Each type of observational unit forms a table.

This approach to data management was developed by Hadley Wickham and can be implemented using the *dplyr* and *tidyr* packages for the R statistical programming language.[111] Commands from these packages were used to reshape data from the data sources into tidy format for inclusion in the SPND as detailed later in this chapter.

Database hosting

Data were hosted on a BC Platforms database (BC Platforms AG, Zurich). This is a secure online data repository which supports upload, storage, analysis and download of complex data relating to human health. It supports the storage of relational database tables alongside genomic data arising from next generation sequencing.

The data for this project were gathered, organised and stored by following a fixed protocol. The data processing pipeline is summarised in Figure 3-1. The next two sections will explore elements of this pipeline in further detail.

Figure 3-1. Summary of the processes required to curate data into a relational database.

3.2.2 Data Sources

This project relied on data which was routinely collected during clinical care or which was collated into research databases which were created prior to the start of the project. Clinical data were gathered from the BadgerNet data management system and from interrogation of UHS's database of blood results (Charts). Growth measurements (including limb circumference measurements) were extracted from the GAP database as well as from BadgerNet. Nutritional intake was recorded prospectively by the research nursing team using a bespoke Excel spreadsheet incorporating macros to reformat data (the SENNAT database).

The following sections explore each of these data sources in turn.

BadgerNet

BadgerNet is a nationally implemented programme designed to gather data on all neonatal unit admissions in the UK.[112] Data are entered locally by clinical staff (or by communication protocols with other clinical information systems) and are held on central servers managed by Clevermed Ltd. This database gathers baseline clinical information (including gestation, details of care and clinical conditions during pregnancy and initial stabilisation), daily data (including medications, respiratory support and measurements of weight, length and head circumference), discharge information (including diagnoses) and patient movements between neonatal units and their ultimate disposition from the database(i.e. whether the infant died, was discharged home or was discharged to a non-neonatal healthcare setting). Anthropometric measurements are carried out by trained clinical staff and are verified by an observer. Variability and accuracy have not been formally determined. Data concerning care delivered by a neonatal unit can be downloaded by

data managers from the unit in question. These data were provided as a collection of MS Access data tables. Subjects were identified by their UHS hospital number and national NHS number. Daily data were classified by a field which calculated the minutes between the first midnight after the infant was born and the start of day in question.

BadgerNet data are entered by clinical staff and the form of the data is designed to facilitate clinical care. Missingness is generally more common in these data than in the other data sources used, and only a handful of key fields mandate data entry. Where there are mandatory fields, individual clinicians may make unpredictable entries in an attempt to bypass these, when the relevant data are unavailable or inconvenient to gather. Latterly, data for BadgerNet has mostly been gathered by automated transfer from the Metavision clinical information system (iMDSoft, Tel Aviv) used on the neonatal unit in Southampton. This has produced a change in the pattern of data quality issues, with Metavision providing dummy data to BadgerNet when it is missing from Metavision's own SQL (structured query language) database. For example, when a head circumference value is not entered at the time of admission to the neonatal unit, a value of 10cm is entered into BadgerNet. Additionally, certain information needs to be manually transferred from Metavision to BadgerNet, a process which is commonly omitted, increasing missingness.

As BadgerNet data are used to build the NNRD (National Neonatal Research Database), it is likely that local Southampton BadgerNet data share some of the same issues seen in the NNRD, as discussed in Section 5.3 below. For example, in a validation study, the sensitivity for detecting whether an infant required supplemental oxygen at 36 weeks PMA was 86.7%, meaning that more than 10% of infants requiring supplemental oxygen were misclassified.[113] Sensitivity for abdominal surgery for NEC was 67.1%.

CHARTS

CHARTS is the bespoke clinical information system in use at UHS for recording medical test results. It holds values for laboratory blood tests. For a subset of infants in the study (recruited after around 2015), point of care blood test ('blood gas') values were also held. UHS staff downloaded data for specific blood tests during specific time periods in accordance with valid consent for research and these data were provided as a CSV file.

Laboratory test data were subject to rigorous quality control by laboratory staff. As a result, these data were generally reliable. Data were provided by UHS in a format which was not compliant with tidy data principles (see section 3.2.2.2) and required careful reformatting using the *dplyr* and *tidyr* packages during data processing. Where there were problems with sample processing, the numeric result was replaced by a text comment. In some cases, these comments were

provided in a standardised format, which were easily converted to error codes, but at other times they were given as free text which were more difficult to encode.

GAP Database

This MS Access database was formed during a previous research project. Its aim is to record weekly limb circumference measurements gathered during the Growth Assessment of Preterm Infants (GAP) Study, alongside head circumference and length measurements recorded during clinical care. It consists of a data entry form which receives limb circumferences, weight, length and head circumference measurements, and the presence of any indwelling lines in the subject's limbs. Limb circumference measurements are taken by specially trained research personnel. These data are then stored as an MS Access table. Subjects were identified by their GAP study number, assigned sequentially to each recruited infant. Data were held on UHS servers.

Data in the GAP database were generally of high quality, with errors limited to easily filtered inconsistencies in the formatting of the alphanumeric study number or unit of measurement errors.

SENNAT database

This database was created in 2011 to hold growth and nutrition data. It consists of an MS Excel database with bespoke macros to process data. Research nurses entered values for weight, length and head circumference when these values were gathered as part of routine clinical care. The file also contains a database of the nutrient content of all nutritional products used on the neonatal unit. Research nurses recorded daily intakes of nutritional products which were stored along with daily nutrient intakes calculated by macros within the spreadsheet. Subjects were identified by their UHS hospital number and each day of data was identified by a 'day number' formed by appending the day of life to the hospital number. Data were stored on UHS servers.

Data in the SENNAT database were entered by trained research staff and implausible values were uncommon. The data entry form and macros have minimal fidelity-checking integrated into their programming. This means that typing errors can become integrated into the stored data. These most commonly take the form of unit errors for measurements (e.g. weights being entered in grams rather than kilograms), missing decimal points, use of commas as a decimal separator or accidental duplication of numbers or decimal separators.

Table 3-1 gives an overview of the data sources used in this project.

Table 3-1. Data sources overview.

Source	Format	Research or Clinical	Subject Identifier	Day of Life Identifier	Data overview
SENNAT Database	MS Excel	Research	UHSFT Hospital Number	Day of life concatenated to patient number	Feed product intake, bespoke PN composition, weight, length, head circumference, derived values for nutrient intakes.
GAP Database	MS Access	Research	GAP Number	Date	Weight, length, head circumference, limb circumference measurements
BadgerNet Database	Downloaded to MS Access	Clinical	UHSFT Hospital Number	Minutes since first midnight after birth	Baseline clinical data concerning pregnancy and birth, details of clinical care (respiratory support, indwelling lines, medications), time and status of discharge.
Charts	Provided by UHS as CSV	Clinical	UHSFT Hospital Number	Date	Laboratory and point of care blood tests

UHS - University Hospital Southampton NHS Foundation Trust, CSV - comma-separated values table, PN - parenteral nutrition.

3.2.3 Data Processing Pipeline

Information used for this project were provided in different formats (table 3-1). Furthermore, different codes were used to identify individual infants and to signify the day of life of that infant for longitudinal data from the different sources.

I designed a pipeline to significantly automate the extraction of these data. The pipeline had seven aims:

1. To extract data for relevant infants (i.e. those for whom consent had been given for inclusion) from each data source.
2. To pseudonymise the data to the requirements set by ethical approvals.
3. To unify the identification code for each infant and each day of life.
4. To encode data held as free text.
5. To organise data in concordance with the principals of relational databases and tidy data.
6. To flag errors and correct them automatically where possible.
7. To derive secondary information from primary data, especially to calculate daily nutrient intake from volumes of nutritional products

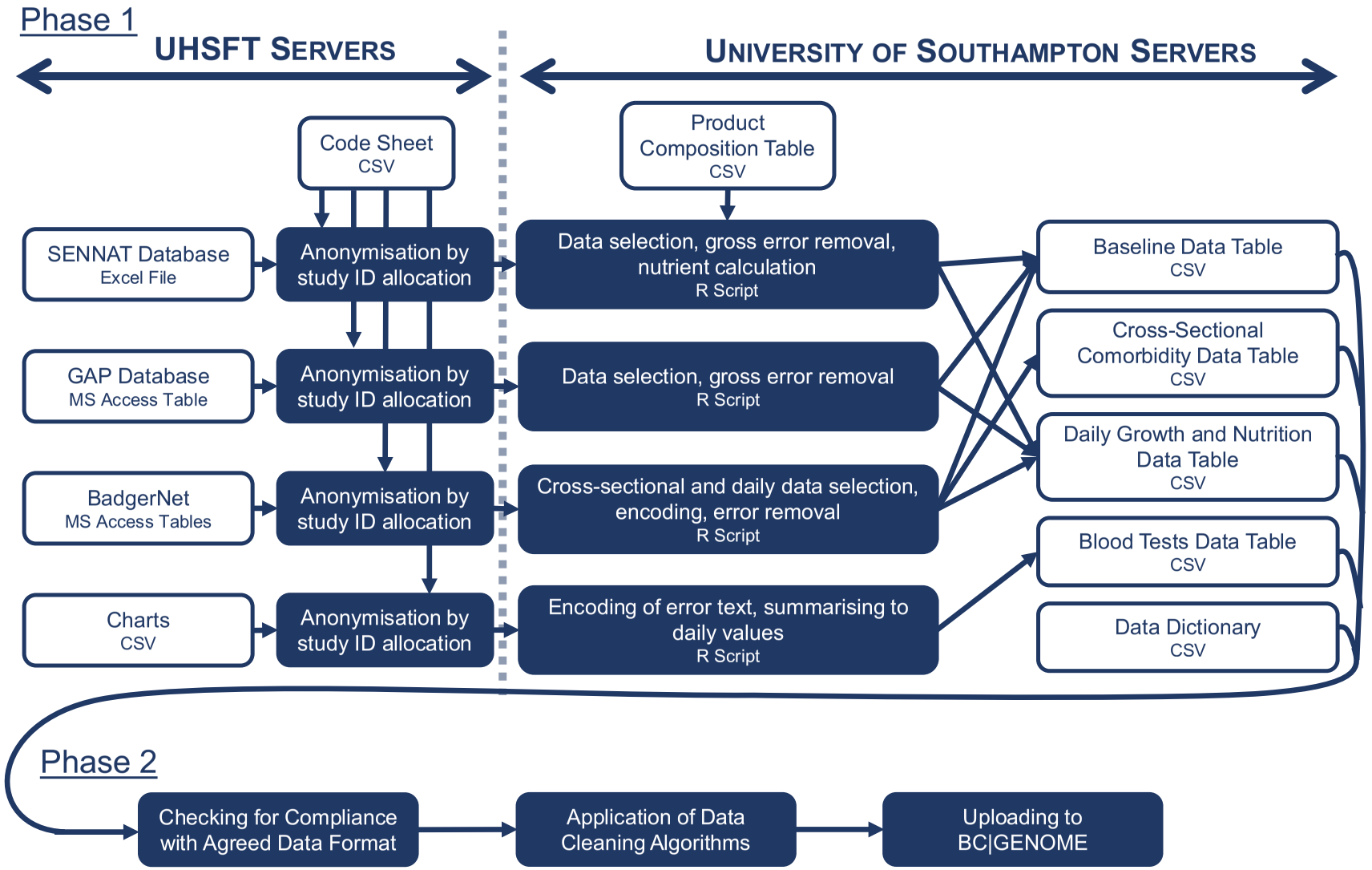


Figure 3-2. Overview of data processing pipeline used to extract data from data sources and prepare tables for entry onto the database.

Figure 3-2 illustrates in detail the data processing pipeline used to integrate, process and prepare data for entry onto the database. Phase 1 was conceptualised and written exclusively by me. Phase 2 was developed by members of the NIHR Southampton BRC Data Science Cross-Cutting Theme with me. During the initial phases of the research project, phase 2 processes were carried out by data science team members but I took over these tasks during the later stages of the project.

Practical application of database updates

Data processing for the pipeline involved running three R scripts. The scripts are held on an online GitHub repository at https://github.com/aneurinyoung/SPND (GitHub Inc, San Francisco). A guide to updating the database is included in the repository and is given here as Appendix 6. The three sequential R scripts are described in the following subsections.

Pseudonymisation

The first phase of data processing focused on replacing patient-identifiable data with codes. The R script was named *1_anonymisation.R*, although a key was kept so that the data could be matched back to patient-identifiable sources, meaning that the process was more accurately described as pseudonymisation rather than full anonymisation.

Minimal preparation of the original data sources was required as set out in Appendix 6. In summary, the following must be performed prior to execution of the R script:

1. Downloading BadgerNet data from the BadgerNet program into an MS Access file.
2. Making a copy of the current SENNAT database along with older versions of the database and removing all password protection.
3. Making a copy of the GAP database.
4. Placing all these files along with blood gas and pathology information into a folder on UHS servers.
5. Adding reference files to the folder.
6. Running *1_anonymisation.R*

A CSV file named *MASTER.csv* cross-referenced UHS numbers, NHS numbers, GAP numbers (a study ID from the GAP database) and a new unique study number (USN) created for the SPND. The R script checked whether an infant already had a USN allocated and allocated a new USN if one was not found. This prevented duplicate records for individuals when new updates to the database were made. The code also calculated day of life from the various day identifiers, allowing daily data to be combined between sources, whilst removing any data which could be used to derive the date of birth of the infant. The SENNAT database had undergone several

incremental version increments prior to the initiation of this doctoral project, with each version being incompatible with the others. Furthermore, there were multiple downloads of BadgerNet data and Charts files. Therefore, the script was designed to flexibly extract data from these different versions before combining them and deduplicating any repeat records.

Area-level deprivation data provided a specific challenge to the pseudonymisation step. BadgerNet stores postcodes of mothers and attempts to derive the lower super output area (LSOA) from these postcodes. Postcodes are patient-identifiable but storage of LSOA is considered not to contravene pseudonymisation rules. The postcode is often entered incorrectly into the BadgerNet database and BadgerNet has a very inflexible approach to postcode interpretation. Many LSOA values are not included in BadgerNet (even in situations where the postcode seems valid). Therefore, this R script disregarded the BadgerNet LSOA, attempted to standardise and repair the entered postcode and looked up the LSOA from a reference file.[114] The index of multiple deprivation (IMD) rank and decile were then derived from further reference files. Where a valid postcode could not be derived, the error was reported back to the user and a reference file was manually updated to correct the error. IMD rank is only available (in a comparable form) for postcodes in England, so infants whose mothers resided in the Channel Islands were excluded from IMD analyses.

Having removed all patient-identifiable data, this script produced a compressed data file containing nested data tables with data from different sources and different versions separated (along with some reference files including lookup tables from SENNAT which described the nutrient content of the nutritional products included).

Consolidation

This script ingested the data file produced by the pseudonymisation step. The first stage of this process was the conversion of feed and fluid intake data to daily nutrient intake data. This step was performed before combining the SENNAT versions, as lookup tables in the SENNAT files were only valid with data from that version. This calculation is a complex step, looking up 33 nutrients for over 100 nutritional products for nearly 600 infants and summarising these into daily values. In the most recent database update, this required over 10,000,000 individual calculations. This calculation can take a significant period of time and therefore the R script was written to perform these calculations in parallel. This allowed the individual calculations to be spread across multiple computer central processing unit (CPU) cores. During the early phases of the project, this process could take several hours, but optimisation of the code and the use of a modern laptop computer with eight available cores reduced the time required to 6 minutes. Further parallelisation using a high power computing cluster could reduce this processing time by around 75%.

After nutrient intakes were calculated, data from the different versions of SENNAT were combined and deduplicated. Data from the different sources were then combined to create three tables: one with a record for each infant, one with data for every individual care day and one with data from Charts which had been reformatted to comply with tidy data rules.

Encoding

In the final stage of data processing, the data were prepared for entry onto the database. This script processed those data to identify new weight measurements. The main body of the script searched the long text strings included in BadgerNet (listing, for example, all the drugs used in one day or all the diagnoses made for an infant) and encoded them as TRUE/FALSE values, rendering them compliant with relational database and tidy data principles. Code was also included to identify episodes of likely bloodstream infection. An episode of infection was identified if there was a positive blood culture result associated with a c-reactive protein (CRP) value of 20 or higher during the period from one day before the blood culture sample to five days after the blood culture sample. Daily data were summarised to provide cross-sectional counts (e.g. the number of days with a central venous line in situ for an individual) and to identify whether the infant was considered to have chronic lung disease on day 28 of life or at 36 weeks PMA.

Finally, all variable names were set to those required by the database and an Excel file was created with each tab containing a table in the format expected by BC Platforms. These data were then ready to undergo the specific cleaning processes described below.

3.2.4 Data Cleaning

Several elements of the above R scripts functioned to identify simple errors and to remove or mark them. The second phase of data processing was led by Hang Phan from the NIHR Southampton BRC Data Science Cross-Cutting theme, working with me. This phase used the Python programming language to perform more detailed analysis to identify errors and to correct or flag them.

Measurement data in this dataset were complex. Weight, body length and head circumference were recorded in BadgerNet, the GAP database and the SENNAT database. These sources were not always in agreement with each other, partly due to simple errors and partly because measurements were recorded at different intervals. BadgerNet stores weight data in two different fields, with one field automatically repeating the previous day's value if a new value is not entered. Therefore, the Python data first discarded these repeat measurements. It then created a linear regression model for weight measurements over time for that infant. Any

measurements outside two standard deviations of that regression line were discarded. Where the data sources had different values for the same day, the closest to the regression line was retained and the others were discarded. Figure 3-3 illustrates a weight chart for an example subject, with erroneous values removed and inconsistencies between databases resolved in favour of the more plausible weight value.

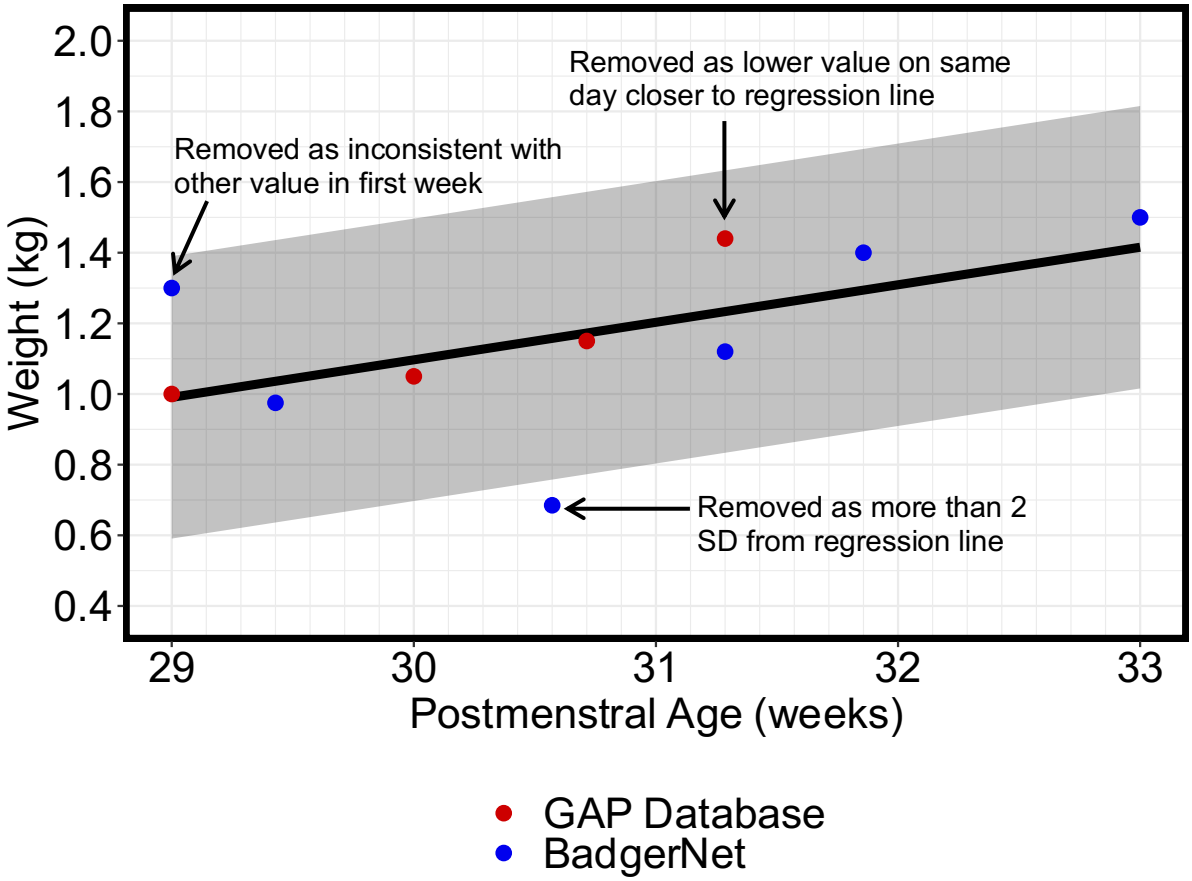


Figure 3-3. Example of actions of data cleaning Python code to remove erroneous weight values and to resolve inconsistencies between data sources. Black line indicates linear regression line. Shaded grey area indicates zone within two standard deviations from the linear regression line. SD – standard deviation.

Measurements at birth were also sometimes discrepant between the sources. The Python code used other records of the same measurement within the first week to resolve any differences in favour of the more plausible value. Finally, z-score values were calculated using the R package *childsds* with reference to the reanalysed UK1990 dataset.[115] This process produced tables in the format expected by the BC Platforms relational database. They were then uploaded to the database using the BC Platforms software.

3.2.5 Evaluation of the Southampton Preterm Nutrition Database

The following subsections detail the variables included in the database followed by a detailed exploration of the dataset, focusing on completeness.

Variables

The BC Platforms database comprised eight related tables: infant-level demographic data (named "baseline"), infant-level comorbidity data ("comorbidities_episode"), daily comorbidity data ("comorbidities_daily"), daily nutritional intakes ("fluid_nutrition"), daily bespoke PN data limited to days when bespoke PN was given ("bespoke_pn"), specimen-level blood sample data allowing for multiple rows during a single day when multiple blood sampling episodes occurred ("pathology"), specimen-level blood gas data ("gas") and daily measurement data limited to days when measurements were taken ("anthropometric"). In total, 310 variables were stored in the database, as detailed in Table 3-2 below.

Table 3-2. Data fields included in the Southampton Preterm Nutrition Database (separated by database table).

NAME	DESCRIPTION	DATA TYPE
Table: baseline		
STUDY_ID	Unique identifier of infant	Text
BIRTHMONTH	Birth month	Integer
BIRTHYEAR	Birth year	Integer
BIRTH_GA	Birth gestational age (in weeks)	Float
GENDER_PHENO	Gender by phenotyping	Choice
GENDER_GENO	Gender by genotyping	Choice
BWEIGHT	Birth weight (kg)	Float
BLENGTH	Birth length (cm)	Float
BOFC	Birth head circumference (cm)	Float
BWEIGHT_SDS	SDS (z-score) of birth weight	Float
BLENGTH_SDS	SDS (z-score) of birth length	Float
BOFC_SDS	SDS (z-score) of birth head circumference	Float
DISCHARGE_GA	Discharge gestational age	Float
ADMIT_FROM	Hospital admitted from	Text
DISCH_DEST	Discharge destination - 3 = baby died	Text
DISCH_HOSP	Discharge hospital	Text
BIRTHPLACE	Birthplace	Text
DESTIN_HOSP	Destination hospital	Text
AGECODE	Age code (for SDS calculation) - in years	Float
PROVIDER_NHS_CODE	Provider NHS code	Text
TIME_DIFF_BIRTH _DISCHARGE	Time difference between birth and discharge in minutes	Float
Table: comorbidities_episode		
STUDY_ID	Unique identifier of infant	Text
ANTEMAG	Whether antenatal magnesium sulphate given	Choice
ANTESTER	Whether antenatal steroids given	Choice
ANTESTER_COURSE	Whether complete course of antenatal steroids given	Choice
APGAR1	Apgar score at 1 minute of life	Number
APGAR5	Apgar score at 5 minutes of life	Number
APGAR10	Apgar score at 10 minutes of life	Number
BIRTHPLACE_CODE	Type of place born	Choice
IMD_DECILE	Decile of index of multiple deprivation (IMD) by maternal postcode	Number
IMD_RANK	Rank (out of 32,844 areas) of IMD by maternal postcode	Number
MATAGE	Maternal age (in years) at time of delivery	Number
MULTIPLICTY	Number of fetuses in pregnancy	Number
SMOKING	Whether mother smoked during pregnancy	Choice
TEMP	Temperate on admission (in °C)	Float
AREDF	Absent or reduced end-diastolic flow	Choice

NAME	DESCRIPTION	DATA TYPE
NEC_BA	Whether necrotising enterocolitis included in discharge diagnoses on BadgerNet	Choice
NEC_SE	Whether necrotising enterocolitis included in discharge diagnoses on SENNAT	Choice
NECSURG	Whether necrotising enterocolitis surgery recorded in BadgerNet	Choice
INOTROPE_DAYS	Number of days inotropic medications given	Number
PDA_BA	Whether patent ductus arteriosus (PDA) included in discharge diagnoses on SENNAT	Choice
PDA_SE	Whether patent ductus arteriosus (PDA) included in discharge diagnoses on BadgerNet	Choice
PDASURG	Whether PDA surgery recorded in BadgerNet	Choice
CANNULA_DAYS	Number of days with peripheral venous catheter in situ	Number
CVC_DAYS	Number of days with surgical central line in situ	Number
PICC_DAYS	Number of days with peripherally inserted central line in situ	Number
UAC_DAYS	Number of days with umbilical arterial catheter in situ	Number
UVC_DAYS	Number of days with umbilical venous catheter in situ	Number
ICUDAYS_2011	Number of days receiving intensive care (by BAPM 2011 standards)	Number
ANTIBIOTIC_DAYS	Number of days receiving antibiotics	Number
SEPSIS_FLAG	Whether at least one episode of sepsis during stay - see mapping for details	Choice
G_POS	Whether at least one episode of Gram positive sepsis during stay	Choice
G_NEG	Whether at least one episode of Gram negative sepsis during stay	Choice
YEAST	Whether at least one episode of yeast sepsis during stay	Choice
IVH1	Whether grade 1 intraventricular haemorrhage included in BadgerNet discharge	Choice
IVH2	Whether grade 2 intraventricular haemorrhage included in BadgerNet discharge	Choice
IVH3	Whether grade 3 intraventricular haemorrhage included in BadgerNet discharge	Choice
IVH4	Whether grade 4 intraventricular haemorrhage included in BadgerNet discharge	Choice
IVHU	Whether unspecified grade intraventricular haemorrhage included in BadgerNet discharge	Choice
CLD_D28	Whether chronic lung disease on day 28 of life	Choice
CLD_W36	Whether chronic lung disease at 36 weeks corrected gestational age	Choice
OXY_DAYS	Number of days receiving supplemental oxygen	Number
HF_DAYS	Number of days receiving high flow therapy	Number
CPAP_DAYS	Number of days receiving continuous positive airway pressure	Number
BIPAP_DAYS	Number of days receiving biphasic airway pressure	Number
VENT_DAYS	Number of days receiving ventilation	Number
HFOV_DAYS	Number of days receiving high frequency ventilation	Number
NITRIC_DAYS	Number of days receiving inhaled nitric oxide	Number

NAME	DESCRIPTION	DATA TYPE
DEXAMETHASONE_DAYS	Number of days receiving dexamethasone (a steroid)	Number
HYDROCORTISONE_DAYS	Number of days receiving hydrocortisone (a steroid)	Number
DIURETIC_DAYS	Number of days receiving a diuretic medication	Number
INSULIN_DAYS	Number of days receiving insulin	Number
Table: comorbidities_daily		
STUDY_ID	Unique identifier of infant	Text
UDN	Unique identifier of neonatal patients in the neonatal unit with day of care	Text
DAY_OF_LIFE	Day of life	Integer
OXY	Whether received supplemental oxygen this day	Choice
HF	Whether received high flow therapy this day	Choice
CPAP	Whether received continuous positive airway pressure this day	Choice
BIPAP	Whether received biphasic airway pressure this day	Choice
VENT	Whether received invasive ventilation this day	Choice
HFOV	Whether received high frequency ventilation this day	Choice
CANNULA	Whether peripheral venous catheter in situ this day	Choice
CVC	Whether surgical central line in situ this day	Choice
PICC	Whether peripherally inserted central line in situ this day	Choice
UAC	Whether umbilical arterial catheter in situ this day	Choice
UVC	Whether umbilical venous catheter in situ this day	Choice
DEXAMETHASONE	Whether received dexamethasone (a steroid) this day	Choice
HYDROCORTISONE	Whether received hydrocortisone (a steroid) this day	Choice
BUDESONIDE	Whether received nebulised budesonide this day	Choice
ANTIBIOTIC	Whether received an antibiotic this day	Choice
INOTROPE	Whether received an inotropic medication this day	Choice
INSULIN	Whether received insulin this day	Choice
DIURETIC	Whether received a diuretic medication this day	Choice
NITRIC	Whether received inhaled nitric oxide this day	Choice
ANTIBIOTIC_NAMES	Names of antibiotics given this day	Text
INOTROPE_NAMES	Names of inotropic medications given this day	Text
DIURETIC_NAMES	Names of diuretic medications given this day	Text
Table: fluid_nutrition		
STUDY_ID	Unique identifier of infant	Text
UDN	Unique identifier of neonatal patients in the neonatal unit with day of care	Text
MBM	Volume of Maternal breastmilk (MBM) given this day(ml)	Float
FORT_BOMB_5	Volume of 5ml Fortifier 'bomb' (2.2g in 5ml) given this day(ml)	Float
ABIDEC	Volume of Abidec supplement given this day(ml)	Float

NAME	DESCRIPTION	DATA TYPE
ALFAMINO	Volume of Alfamino formula given this day(ml)	Float
APT_CG_FIRST	Volume of Aptamil or C&G First formula given this day(ml)	Float
APT_PRETERM	Volume of Aptamil Preterm formula given this day(ml)	Float
BABIVEN	Volume of Babiven PN given this day(ml)	Float
BLOOD	Volume of Packed red blood cells given this day(ml)	Float
CHOLECAL	Volume of Cholecalciferol supplement given this day(ml)	Float
DALAVIT	Volume of Dalivit given this day(ml)	Float
DBM	Volume of Donor breastmilk given this day(ml)	Float
DEX10	Volume of Dextrose 10% fluid given this day(ml)	Float
DEX125	Volume of Dextrose 12.5% fluid given this day(ml)	Float
DEX15	Volume of Dextrose 15% fluid given this day(ml)	Float
DEX20	Volume of Dextrose 20% fluid given this day(ml)	Float
DEX5	Volume of Dextrose 5% fluid given this day(ml)	Float
NP_HYDROL	Volume of Hydrolysed Nutriprem formula given this day(ml)	Float
INFATRINI	Volume of Infatrini formula given this day(ml)	Float
INFATRINI_PEPTI	Volume of Infatrini Peptisorb formula given this day(ml)	Float
KETO_LIQ	Volume of Ketovite liquid supplement given this day(ml)	Float
NP2	Volume of Nutriprem 2 formula given this day(ml)	Float
PEPTIJ	Volume of Peptijunior formula given this day(ml)	Float
KCL20	Volume of Potassium chloride 20% supplement given this day(ml)	Float
PRE_CONC	Volume of Preterm concentrated PN given this day(ml)	Float
PRE_SOD	Volume of Preterm plus sodium PN given this day(ml)	Float
PRE_PN	Volume of Preterm PN given this day(ml)	Float
SALINE045	Volume of 0.45% saline fluid given this day(ml)	Float
SALINE09	Volume of 0.9% saline fluid given this day(ml)	Float
SMOF	Volume of SMOF lipid emulsion given this day(ml)	Float
NACL30	Volume of 30% sodium chloride solution given this day(ml)	Float
NAGLYPHOS	Volume of Sodium glycerophosphate supplement given this day(ml)	Float
SYTRON	Volume of Sytron iron supplement given this day(ml)	Float
ZINCSULPH	Volume of Zinc sulphate supplement given this day(ml)	Float
FULLTERM_PN	Volume of Term PN given this day(ml)	Float
FORT_BOMB_HALF	Volume of Half strength Fort Bomb (1g in 10ml) given this day(ml)	Float
FORT_BOMB_10	Volume of Fortifier 'bomb' (2.2g in 10ml) given this day(ml)	Float
KETO_TAB	Volume of Ketovite tablet supplement given this day(ml)	Float
MBM_HALF_FORT	Volume of MBM with half strength fortifier given this day(ml)	Float
MBM_QUART_FORT	Volume of MBM with one quarter strength fortifier given this day(ml)	Float

NAME	DESCRIPTION	DATA TYPE
MBM_THREEQUART_FORT	Volume of MBM with three quarters strength fortifier given this day(ml)	Float
MBM_FULL_FORT	Volume of MBM with full strength fortifier given this day(ml)	Float
NEOCATE	Volume of Neocate formula given this day(ml)	Float
PROT_HALF	Volume of Half strength protein supplement given this day(ml)	Float
PROT_QUART	Volume of Quarter strength protein supplement given this day(ml)	Float
PROT_THREEQUART	Volume of Three quarters strength protein supplement given this day(ml)	Float
PROT_BOMB	Volume of Protein supplement bomb given this day(ml)	Float
PROT_FULL	Volume of Full strength protein supplement given this day(ml)	Float
SANDOPHOS	Volume of Phosphate Sandoz supplement given this day(ml)	Float
KPHOS	Volume of Potassium phosphate supplement given this day(ml)	Float
CLINOLEIC	Volume of Clinoleic lipid solution given this day(ml)	Float
ALFACAL	Volume of Alfacalcidol supplement given this day(ml)	Float
ENFAMIL	Volume of Enfamil formula given this day(ml)	Float
FOLIC	Volume of Folic acid supplement given this day(ml)	Float
OLD_APT_PRETERM	Volume of Pre-September 2011 Aptamil Preterm Formula given this day(ml)	Float
OLD_D1_PRETERM_PN	Volume of Old day 1 preterm PN given this day(ml)	Float
OLD_D5_PRETERM_PN	Volume of Old day 5 preterm PN given this day(ml)	Float
OLD_MBM_FULL_FORT	Volume of Pre-September 2011 MBM with full fortifier given this day(ml)	Float
OLD_MBM_HALF_FORT	Volume of Pre-September 2011 MBM with half fortifier given this day(ml)	Float
OLD_MBM_QUART _FORT	Volume of Pre-September 2011 MBM with quarter fortifier given this day(ml)	Float
OLD_MBM _THREEQUART_FORT	Volume of Pre-September 2011 MBM with three quarters fortifier given this day(ml)	Float
OLD_NP2	Volume of Pre-September 2011 Nutriprem 2 formula given this day(ml)	Float
OLD_FULLTERM_PN	Volume of Old term PN given this day(ml)	Float
PREGESTIMIL	Volume of Pregestimil formula given this day(ml)	Float
SMA_FIRST	Volume of SMA First formula given this day(ml)	Float
VITE	Volume of Vitamin E supplement given this day(ml)	Float
VITK	Volume of Vitamin K supplement given this day(ml)	Float
ENERGY	Energy delivered this day(kcal)	Float
PROTEIN	Protein delivered this day(g)	Float
CARBOHYDRATE	Carbohydrate delivered this day(g)	Float
FAT	Fat delivered this day(g)	Float
SODIUM	Sodium delivered this day(mmol)	Float
CHLORIDE	Chloride delivered this day(mmol)	Float
POTASSIUM	Potassium delivered this day(mmol)	Float

NAME	DESCRIPTION	DATA TYPE
CALCIUM	Calcium delivered this day(mmol)	Float
PHOSPHORUS	Phosphorus delivered this day(mmol)	Float
MAGNESIUM	Magnesium delivered this day(mmol)	Float
IRON	Iron delivered this day(μmol)	Float
ZINC	Zinc delivered this day(μmol)	Float
COPPER	Copper delivered this day(μmol)	Float
SELENIUM	Selenium delivered this day(nmol)	Float
IODINE	Iodine delivered this day(nmol)	Float
MANGANESE	Manganese delivered this day(nmol)	Float
VITAMIN_A	Vitamin A delivered this day (international units)	Float
VITAMIN_D	Vitamin D delivered this day (international units)	Float
VITAMIN_E	Vitamin E delivered this day (international units)	Float
VITAMIN_K	Vitamin K delivered this day (international units)	Float
THIAMIN	Thiamine delivered this day(μg)	Float
RIBOFLAVIN	Riboflavin delivered this day(μg)	Float
VITAMIN_B6	Vitamin B12 delivered this day(μg)	Float
FOLATE	Folate delivered this day(μg)	Float
VITAMIN_B12	Vitamin B18 delivered this day(μg)	Float
BIOTIN	Biotin delivered this day(μg)	Float
PANTOTHENIC_ACID	Pantothenic acid delivered this day(mg)	Float
NIACIN	Niacin delivered this day(mg)	Float
VITAMIN_C	Vitamin C delivered this day(mg)	Float
TAURINE	Taurine delivered this day(mg)	Float
CHOLINE	Choline delivered this day(mg)	Float
CARNITINE	Carnitine delivered this day(mg)	Float
INOSITOL	Inositol delivered this day(mg)	Float
WWEIGHT	Working weight (kg)	Float
DIFFDAY	Difference between day of working weight and day of life(days)	Number
STUDY_ID	Unique identifier of infant	Text
Table: bespoke_pn		
UDN	Unique identifier of neonatal patients in the neonatal unit with day of care	Text
WWEIGHT	Working weight (kg)	Float
DIFFDAY	Difference between day of working weight and day of life(days)	Number
AQ_VOLUME	Volume of bespoke aqueous parenteral nutrition (PN) given this day(ml)	Float
BS_PRIM10	Volume of Primene 10% given this day(ml)	Float

NAME	DESCRIPTION	DATA TYPE
BS_GLUC50	Volume of Glucose 50% given this day(ml)	Float
BS_KCL20	Volume of Potassium Chloride 20% given this day(ml)	Float
BS_CAGLUC10	Volume of Calcium gluconate 10% given this day(ml)	Float
BS_PEDITRACE	Volume of Peditrace given this day(ml)	Float
BS_GLUC5	Volume of Glucose 5% given this day(ml)	Float
BS_NACL30	Volume of Sodium Chloride 30% given this day(ml)	Float
BS_MGSULPH50	Volume of Magnesium Sulphate 50% given this day(ml)	Float
BS_NAGLYPOS216	Volume of Sodium Glycerophosphate 21.6% given this day(ml)	Float
BS_ADDIPHOS	Volume of Addiphos given this day(ml)	Float
BS_SELENASE	Volume of Sodium Selenite (Selenase) given this day(ml)	Float
BS_ZINCSULPH	Volume of Zinc Sulphate given this day(ml)	Float
BS_KPHOS1742	Volume of Potassium Phosphate 17.42% given this day(ml)	Float
BS_ACETATE30	Volume of Sodium Acetate 30% given this day(ml)	Float
BS_ACETATE49	Volume of Potassium Acetate 49% given this day(ml)	Float
BS_IRONCHLOR	Volume of Iron III Chloride 0.1mg/ml given this day(ml)	Float
BS_VAMINOLACT	Volume of Vaminolact given this day(ml)	Float
BS_CACL147	Volume of Calcium Chloride 14.7% given this day(ml)	Float
BS_ASCORB	Volume of Ascorbic Acid 100mg/ml given this day(ml)	Float
BS_VITK	Volume of Vitamin K 10000mcg/ml given this day(ml)	Float
BS_WFI	Volume of Water for Injection given this day(ml)	Float
BS_SALINE045	Volume of Saline 0.45% given this day(ml)	Float
BS_GLUC10	Volume of Glucose 10% given this day(ml)	Float
BS_COPPERSULPHATE	Volume of CuSO4 given this day(ml)	Float
BS_NACL09	Volume of Sodium Chloride 0.9% given this day(ml)	Float
BS_GLUC20	Volume of Glucose 20% given this day(ml)	Float
BS_NAKREP	Volume of Sodium potassium replacement given this day(ml)	Float
BS_CUCL	Volume of Copper Chloride given this day(ml)	Float
LIPID_VOLUME	Volume of bespoke lipid solution given this day(ml)	Float
BS_CLINOLEIC	Volume of Clinoleic 20% given this day(ml)	Float
BS_SMOF	Volume of SMOF 20% given this day(ml)	Float
BS_SOLVITO	Volume of Solvito N (in IL 20%) given this day(ml)	Float
BS_VITLIPID	Volume of Vitlipid (in IL 10%) given this day(ml)	Float
BS_INTRALIPID	Volume of Intralipid 20% given this day(ml)	Float
Table: pathology		
STUDY_ID	Unique identifier of infant	Text
UDN	Unique identifier of neonatal patients in the neonatal unit with day of care	Text

NAME	DESCRIPTION	DATA TYPE
SPECI_TIME	Time of specimen (as a decimal fraction of a day from midnight to midnight)	Float
DAY_OF_LIFE	Day of life	Number
ABSOLUTE_RETICULOCYTES	Absolute reticulocyte count (count x 10^9/l)	Float
ALBUMIN	Serum albumin (g/l)	Float
ALKALINE_PHOSPHATASE	Serum alkaline phosphatase (U/l)	Float
ALT	Serum ALT (U/l)	Float
BLOOD_CULTURE	Blood culture result	Choice
CALCIUM	Serum calcium (mmol/l)	Float
CONJUGATED_BILIRUBIN	Serum conjugated bilirubin (μmol/l)	Float
CREATININE	Serum creatinine (mmol/l)	Float
CRP	Serum c reactive protein (mg/l)	Float
GRAM_STAIN	Blood culture Gram stain result	Choice
HAEMAGLOBIN	Haemoglobin (g/l)	Float
MAGNESIUM	Serum magnesium (mmol/l)	Float
NEUTROPHIL_COUNT	Neutrophil count (count x 10^9/l)	Float
PHOSPHATE	Serum phosphate (mmol/l)	Float
PLATELET_COUNT	Platelet count (count x 10^9/l)	Float
POTASSIUM	Serum potassium (mmol/l)	Float
SERUM_COPPER	Serum copper (μmol/l)	Float
SERUM_MANGANESE	Serum manganese (μmol/l)	Float
SERUM_SELENIUM	Serum selenium (μmol/l)	Float
SERUM_ZINC	Serum zinc (μmol/l)	Float
SODIUM	Serum sodium (mmol/l)	Float
TOTAL_BILIRUBIN	Serum total bilirubin (μmol/l)	Float
TOTAL_PROTEIN	Serum protein (g/l)	Float
TRIGLYCERIDES	Serum triglycerides (mmol/l)	Float
UREA	Serum urea (mmol/l)	Float
VITAMIN_A	Serum vitamin A (μmol/l)	Float
VITAMIN_D	Serum vitamin D (nmol/l)	Float
VITAMIN_E	Serum vitamin E (μmol/l)	Float
WHITE_CELL_COUNT	White cell count (count x 10^9/l)	Float

Table: gas		
UDN	Unique identifier of neonatal patients in the neonatal unit with day of care	Text
GAS_TIME	Time of blood gas (as a decimal fraction of a day from midnight to midnight)	Float

NAME	DESCRIPTION	DATA TYPE
DAY_OF_LIFE	Day of life	Number
GAS_BASE_EXCESS	Base excess	Float
GAS_CALCIUM	Ionized calcium (mmol/l)	Float
GAS_CHLORIDE	Chloride (mmol/l)	Float
GAS_GLUCOSE	Glucose (mmol/l)	Float
GAS_HAEMOGLOBIN	Haemoglobin (g/l)	Float
GAS_HCO3	Bicarbonate (mmol/l)	Float
GAS_LACTATE	Lactate (mmol/l)	Float
GAS_PCO2	Carbon dioxide partial pressure (kPa)	Float
GAS_PH	pH (kPa)	Float
GAS_PO2	Oxygen partial pressure	Float
GAS_POTASSIUM	Potassium (mmol/l)	Float
GAS_SODIUM	Sodium (mmol/l)	Float
Table: anthropometric		
STUDY_ID	Unique identifier of infant	Text
UDN	Unique identifier of neonatal patients in the neonatal unit with day of care	Text
PROVIDER_NHS_CODE	Provider NHS Code	Text
CORRECTED_GA	Corrected gestational age (weeks)	Float
DAY_OF_LIFE	Day of life	Float
WEIGHT	Weight (kg)	Float
LENGTH	Length (cm)	Float
OFC	Head circumference (cm)	Float
LMUAC	left mid-upper arm circumference measurements (mm)	Float
RMUAC	right mid-upper arm circumference measurements (mm)	Float
LMTC	left mid-thigh circumference measurements (mm)	Float
RMTC	right mid-thigh circumference measurements (mm)	Float
BODYWATER	Total body water as a proportion of bodyweight on this day	Float
AGECODE	Corrected gestational age (years)	Float
WEIGHT_SDS	SDS value of weight	Float
LENGTH_SDS	SDS value of length	Float
OFC_SDS	SDS value of head circumference	Float
WEIGHT_FFLAG	Flag for filtering weight measurement	Choice
LENGTH_FFLAG	Flag for filtering length measurement	Choice
OFC_FFLAG	Flag for filtering OFC measurement	Choice
WEIGHT_SOURCE	Source of weight data	Text
LENGTH_SOURCE	Source of length data	Text

NAME	DESCRIPTION	DATA TYPE
OFC_SOURCE	Source of OFC data	Text

3.3 Introduction to Mathematical Models of Growth

Mathematical approaches may be used to describe the pattern of growth in a population or for an individual.

JM Tanner wrote:

> *Many attempts have been made to find mathematical curves which fit, and thus summarize, human and animal growth data. Most have ended in disillusion or fantasy; disillusion because new data failed to conform to them, fantasy because the system eventually contained so many parameters (or 'constants') that it became impossible to interpret them biologically.*[116]

Despite this warning, serious attempts at describing growth mathematically had been made prior to Tanner's writing. Efforts in the in three decades since publication of his book have taken advantage of new mathematical approaches and then-unforeseeable improvements in computing power to describe growth using models which can analyse large datasets to form simple but mathematically sound growth curves.

This section explores the different methods used in these modelling approaches and provides justification for the methods used in later chapters when modelling the growth of infants in the different cohorts considered in this research programme.

3.3.1 Modelling Growth in Populations

Growth over time is a complex process, with considerable variation between individuals. Several techniques have been used to summarise the normal, typical or expected growth pattern in a population. The SuperImposition by Translation And Rotation (SITAR) method was developed to illustrate the pattern of growth within a population, using a single mean curve.[117] It also provides a means to assess the impact of fixed effects on that growth pattern. Growth charts are generally constructed from measurements of infants at birth and aim to classify infants by their weight centile at birth and to set a target growth trajectory. In practice, such charts are most commonly generated using the LMS (lambda, mu, skew) method. Each of these approaches is discussed in the following sections.

SITAR Modelling

The SITAR modelling method was developed by Prof Tim Cole and was based on earlier work on shape invariant models.[118] SITAR (and shape invariant models more generally), assumes that there is an underlying growth pattern in the population and that each subject's growth differs from this underlying pattern in a limited number of ways. Specifically, SITAR provides a summary mean

curve (a natural cubic spline curve, usually with three degrees of freedom) for a cohort and describes each subject's deviation from the mean curve in terms of the subject's:

- overall size (α), analogous to a shift along the y-axis of a chart of time against measurement value
- timing of the age at peak growth velocity (β)
- velocity of growth (γ), analogous to the gradient of a line on the chart described above.

These random effects are illustrated in Figure 3-4.

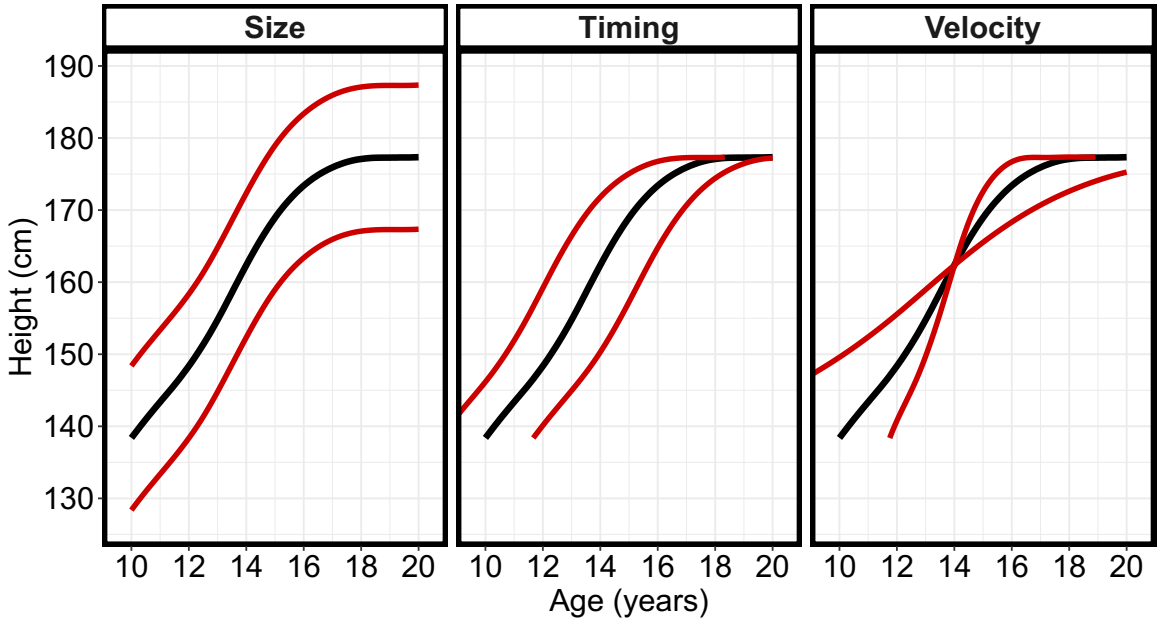


Figure 3-4. The effect of changes in each of the random effects of a SITAR model, with the mean line shown in black and the effects of increasing or decreasing the random effect in red. Adolescent height data for males from the WHO dataset are used as an example.[119]

A SITAR model can be described with the following formula:

$$y_{it} = \alpha_i + h\left(\frac{t - \beta_i}{\exp(-\gamma_i)}\right)$$

Where: y_{it} is the measurement value (e.g. weight) for subject i at age (or postmenstrual age) t.

$h(t)$ is the underlying cubic spline curve

α_i is the size metric described above for that subject

β_i is the timing value described above for that subject

γ_i is the velocity value for that subject

In practice, optimising the cubic spline curve and the random effects for each individual is an iterative process of non-linear mixed effects modelling which requires considerable computing resources.

The SITAR model provides two advantages over simpler ways to describe growth. Firstly, the mean line is smoothed in a way which more closely reflects actual growth patterns than simply taking the mean or median weight on a particular day of life. Figure 3-5 uses weight gain data for infants born in England in 2014-2018 after 31 completed weeks gestation to illustrate that using mean or median values for each day of life produces a line which is insufficiently smoothed to accurately reflect real weight gain patterns.

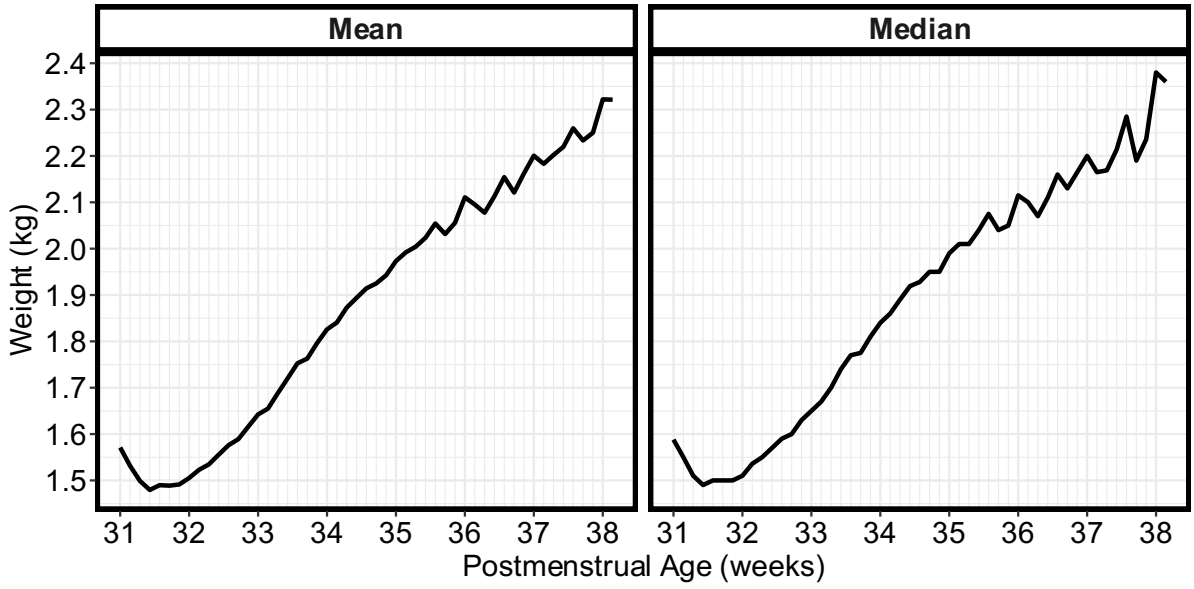


Figure 3-5. Weight gain of 9082 infants born at 31 weeks completed gestation in England during 2014-2018, adjusted by moving the birth gestation of each infant back to 31 weeks exactly and summarised by taking the A. mean and B. median weight value for infants weighed on that day.

Secondly, the addition of fixed effects to the model allows the researcher to understand the impact of a factor on the overall size of the individual, the time of peak growth velocity and the magnitude of growth velocity. The example given in the paper which introduces SITAR is the effect of oxandrolone treatment on the pubertal growth of people with Turner's syndrome.[117] This paper demonstrates that oxandrolone treatment increased growth velocity and illustrates the effect using the SITAR summary mean curve for the treatment and placebo groups. In the context of this doctoral research project, cohort identity (i.e. English infants born 2006-2011, English infants born 2014-2018 and the Southampton Preterm Nutrition Database cohort) was included as a fixed effect to assess the differences in growth patterns between the cohorts.

SITAR models require the input of carefully cleaned data, with removal of spurious weight values. This process is complicated and must be automated when dealing with very large numbers of subjects, such as those included in the English datasets. The practical application of SITAR modelling, including the data cleaning methods applied, is given in Chapter 6.

Advantages and Disadvantages of SITAR Models

SITAR modelling provides an effective way to summarise the growth in a defined population over a specific growth phase. The ability to include fixed effects in the modelling process allows granular assessment of the ways that treatments, conditions or cohort memberships affect the growth pattern seen. It also provides a mathematically robust way to describe the way the growth of one individual differs from the mean growth of the population.

Whilst SITAR models describe growth in a population, they cannot be used to monitor the growth of an individual over time as they do not produce growth standards or targets. In common with other forms of growth modelling, they can only describe the growth in a population but they do not define growth patterns which are optimal or desirable.

LMS Growth Charts

The LMS method uses a series of paired ages (or postmenstrual ages) along with weight or other measurement values to establish centile lines describing the distribution of measurement values within a population. Three values are defined for each age timepoint. M is median value of all the measurements at that age. L is a power value used to remove skewness from the data. Anthropometric measurements are usually right-skewed (with a longer tail to the right of the median than to the left on a standard histogram). Raising each value to a certain power will remove this skewness and result in normalised data. The LMS method uses a method described by Box and Cox in the 1960s to discover the optimal power by which to raise the measurements in order to normalise the distribution of values.[120] The result of this is that L represents a series of values required to generate normally distributed data at each time point. S is the coefficient of variation. This value is calculated by dividing the standard deviation of the measurements at that timepoint by the mean value. Once a value for L, M and S has been calculated at each timepoint, each of the values can be plotted against age. Each of these plots is then smoothed, with simple shapes adequately described by straight lines formed by linear regression or curved polynomials. More complex shapes (as is often the case for M) may need to be described as cubic splines.

Figure 3-6-A illustrates LMS curves using data from the WHO weight standard for male children aged from one month to four years of life (which were all smoothed using cubic splines).[121] Centile lines can then be derived from these L, M and S values using the following formula:

$$C = M(1 + LSZ)^{\frac{1}{L}}$$

Where C is the measurement value for standard deviation score Z. Lookup tables can be used to convert centile values to standard deviation scores for use in this formula. Connecting these values draws centile lines to form a standard growth chart (Figure 3-6-B). By convention, lines are

drawn at intervals of two thirds of a standard deviation (0.4th, 2nd, 9th, 25th, 50th, 75th, 91st, 98th and 99.6th centiles).

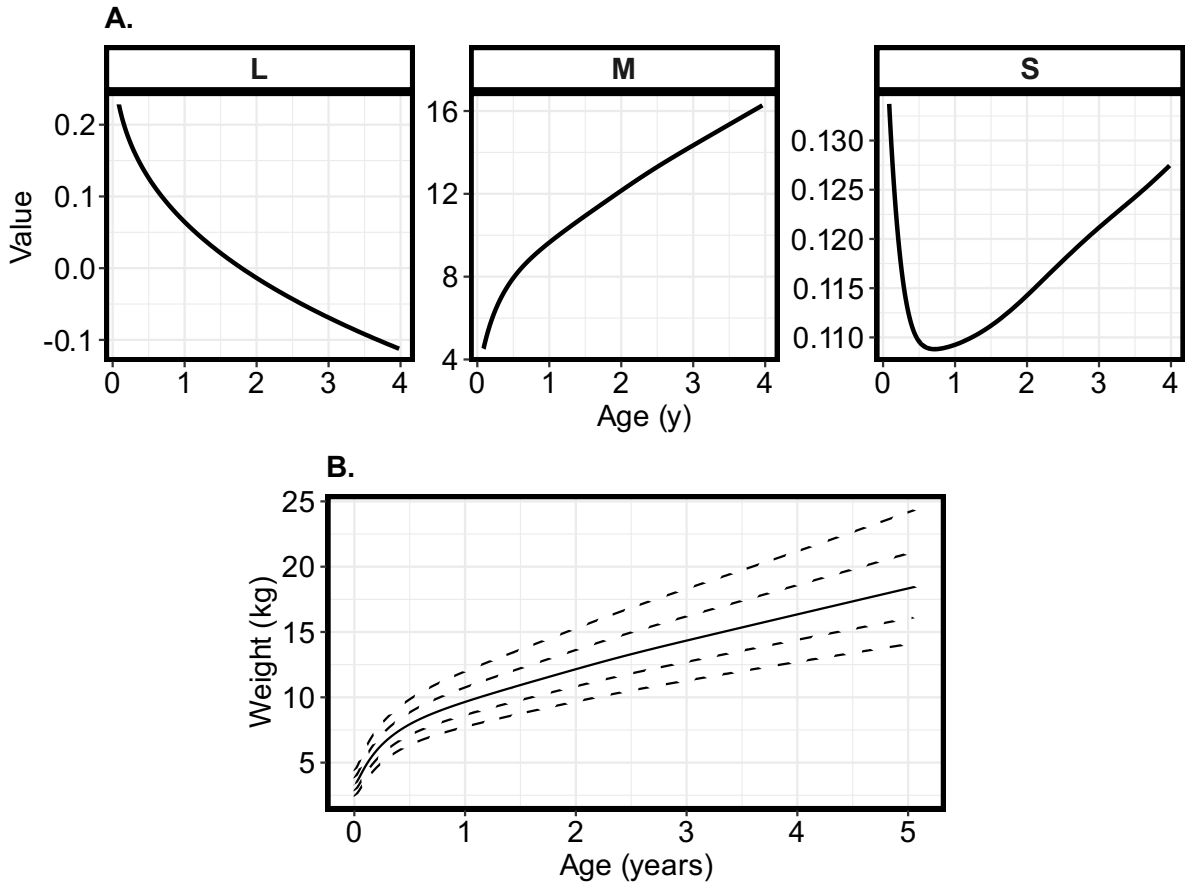


Figure 3-6. A. Example L, M and S curves using data from the WHO growth standards for male children aged from one month to four years of life.[121] B. Growth chart derived from LMS values with dashed lines at -2, -1, +1 and +2 standard deviation scores from the median (solid line).

Cross-sectional and Longitudinal Data

Most growth charts which are currently in widespread use for preterm infants were created using the LMS method with birthweight data. This means that all data used for creating growth centiles is derived from infants at the point of birth. In other words, growth charts created in this reflect the growth of the fetus in utero. An alternative strategy is to use longitudinal growth data from preterm infants as they grow. Growth charts created in this way reflect a combination of size at birth and subsequent growth. Therefore, they may provide a more appropriate standard to measure postnatal growth of preterm infants.

However, such charts present problems of their own. When the independent variable is set as the postmenstrual age (PMA), the centile at each timepoint is made up of a combination of at-birth values for infants born at that time and values for other infants who are older (but were born at an earlier gestation). This makes the actual meaning of the centile lines difficult to interpret and limits their use as a way to track infant growth. An alternative strategy would be to limit each

chart to a distinct gestational window, although the complex nature of the changes in weight during the first two weeks of life mean that, even after restricting each model to cover one gestational week, charts will reflect birthweights and weights after postnatal weight loss at the end of the first week. One approach to remedy these issues would be to use day of life as the independent variable, having selected a gestational window in which similar growth patterns are expected.

As explored in the introduction chapter, recent growth charts developed by the WHO and the INTERGROWTH-21st project have sought to define growth standards (i.e. expected healthy growth) by limiting their cohorts to children who were free from disease and were fed according to international guidance.[31] This approach has clear advantages in the term-born population, where most infants will be free from significant disease and, depending on the population selected, many will be exclusively breastfed. The strictures of selecting "healthy" infants fed to current standards are much more difficult to apply to a very preterm population. It could be argued that prematurity is, in itself, a disease state. Furthermore, some degree of respiratory distress is essentially universal at the earlier gestations, with most infants also experiencing other distinct diseases or complications of prematurity. Even if a "healthy" cohort could be identified, it is not yet certain what constitutes an optimal nutritional approach for these infants.

The ideal growth chart for preterm infants would:

1. Reflect the actual growth pattern of preterm infants after birth, including early weight loss (or static weight).
2. Be based on a population of infants without significant comorbidities (including chronic lung disease and necrotising enterocolitis).
3. Be based on a population of infants who all received adequate nutrition (although this is difficult to define).
4. Contain enough infants (in each gestation band) to build stable and reliable LMS models, even at the extremes of small and large infants.

The last of these points is not trivial. For a growth chart to be useful, it should capture the growth patterns of infants born small for gestational age, but LMS centiles are most reliable when a significant number of the subjects lie around that centile line. Constructing growth charts with, for example, a 0.4th centile line would require a large number of infants born below the 0.4th centile, which would only be possible with a very large starting population.

There is currently no population of very preterm infants which is large enough and has sufficiently detailed information to create the ideal growth charts described above. Chapter 4 explores the use of data from the SPND to form LMS charts and discusses the challenges and limitations of using these data.

3.3.2 Modelling the Growth of Individuals

As well as population-level growth summaries, there are situations where it is useful to model the growth of an individual subject. The reasoning behind this approach and the methods used are described in this section.

Error

Anthropometric measurements are prone to error. Errors can arise for multiple reasons and are typically classified as random errors or systematic errors. Taking weight as an example, random error may occur when the instruments used are imprecise or there is variation in the method used to weigh infants. Systematic errors may occur when the method used to weigh an infant persistently over- or under-estimates weight. For example, the practice of leaving nappies or respiratory equipment on the infant during weighing may cause systematic overestimation of weight.

Further errors may be introduced during the recording of measurements. Units may be confused (for example use of a value in grams when a value in kilograms is called for) and typing errors may introduce error or movement of the decimal point. Different measurement values may be entered onto different clinical systems for the same day of life.

In addition to measurement errors, there may be variation in measurement values which do not reflect the true growth of the individual. For example, acute illness may cause oedema (fluid retention) in the infant. This will cause the weight of the infant to increase but does not reflect true growth of the subject's tissues. Therefore, the infant will have appeared to have thrived during a period when true growth was actually arrested, and then to have lost weight during recuperation.

Value Interpolation

Measurements of infants are undertaken intermittently. During this study, infants were weighed twice-weekly and had their length and head circumference measured weekly. In practice, the intervals between measurements were not entirely predictable as infants were weighed or measured only on certain days of the week (but always at birth) and measurements were sometimes missed due to the infant being too unwell to measure. Models used to assess influences on growth benefit from weight values present at every timepoint (i.e. on every day of life for this study). Therefore, a strategy is required to estimate the actual weight of an infant on any given day.

Error, random variation and the need for interpolation of values mean that models to smooth measurements for each individual are useful. The degree of smoothing can be controlled to avoid under- and over-fitting and the model equation can be used to produce interpolated values on each day.

Modelling Approaches

Attempts to model childhood growth have typically used polynomials, linear splines or cubic splines.[122, 123] Other biological systems have been described using the broken stick model, a specific application of linear splines.[124]

Before trialling these models, some initial cleaning of the weight data was required, to exclude extreme outliers and implausible values. Implausible values were defined as those over 60000g and below 60g (which are so low or so high that they could not be a factor-of-ten error) and were removed. Likely factor-of-ten errors were corrected, based on visual inspection of weight value distribution in the national dataset. The following algorithm describes the process:

```
If weight is greater 60000 or less than 60, discard it.
If weight is above a line defined by 2500/11*CGA-1700/11 where CGA
is the corrected gestational age in weeks, divide it by 10
(correcting for transcription errors).
If weight is below a line defined by 125/7*CGA-2500/7, multiply it
by 10 (correcting for transcription errors).
If weight has not been adjusted or discarded by the above steps and
it lies above the line
 650/3 CGA-11500/3 or below the line 400/13 CGA-6700/13, discard
it.
Otherwise, retain it.
```

A subgroup of infants was then selected based on the following inclusion criteria:

1. Latest weight value on or after day 42 of life *and*
2. At least one weight value within seven days (before or after) day 42 of life.

These criteria were designed to select infants for whom there was sufficient data to model the first six weeks of life and to exclude a handful of infants who were transferred out of Southampton early in life and then returned when they were more than six weeks old. This resulted in the selection of 375 infants.

Three sets of models were then made individually for each infant:

1. Polynomial models of order 1 to 10.
2. Natural cubic spline models with 3 to 10 knots.
3. Broken stick models with 3 to 10 knots.

R was used to generate models using the base features of R, the *splines* package and the *brokenstick* package. Each of these modelling approaches is described below, with example models for each.

Polynomial Models

Polynomial modelling is an extension of linear regression modelling, in which the dependent variable can change in response to the independent variable and to powers of the independent variable. The number of powers which can be used is described as the degree of the polynomial. In effect, increasing degrees of polynomials allow for increasingly complex curves to be applied. An example of increasing the degree of polynomial for the modelling of the weight gain for an individual infant is given as figure 3-7.

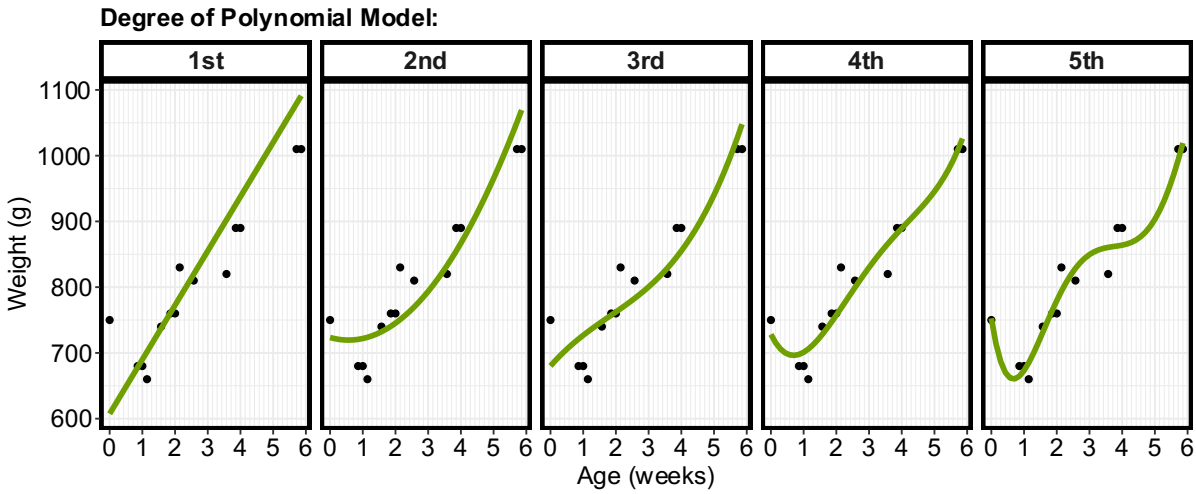


Figure 3-7. Polynomial models for smoothing weight values for an example infant. Dots are actual weight values and green line is the polynomial smoothing model using polynomials up to the degree specified above.

The first order polynomial is simply a linear regression line and is underfitted, not accurately describing the complex weight gain pattern seen in early life. The second and third polynomial models likewise fail to capture the actual growth of the infant. The fourth order polynomial model captures the complexity of early weight change. However, overfitting is becoming apparent in this model, as its complexity in the later weeks does not reflect real growth but is impacted by random error. The fifth order polynomial model is clearly overfitted.

Polynomial models have the advantage of being simple to fit. However, it is unusual for the underlying biological growth pattern to truly follow a polynomial pattern with the result that such models tend to be underfitted or overfitted.

Natural Cubic Spline Models

A spline is a piecewise polynomial, which is to say that it is a series of connected polynomial curves. A cubic spline restricts the polynomials used to a third order polynomial, meaning that each segment is a simple curve. The complexity of the overall curve can be controlled by defining the number of segments which are allowed in the model, with each boundary between one segment and the next (or one segment and the beginning or end of the curve) defined as "knots". Natural cubic splines exert restrictions on the relationships between different segments at the knots, essentially "stiffening" the curve and preventing rapid changes in velocity.

By allowing gradual transition between smooth growth periods, natural cubic splines have been shown to efficiently describe human growth in several different contexts, including the growth of young children.[122] Their use in SITAR and LMS models is described above. Figure 3-8 illustrates a natural cubic spline curve fitted to the data for the same infant as was used above. Due to the restrictions on natural cubic splines, the number of knots makes little difference to model fitting and so two knots (six degrees of freedom) were used, as is common in the field.[117]

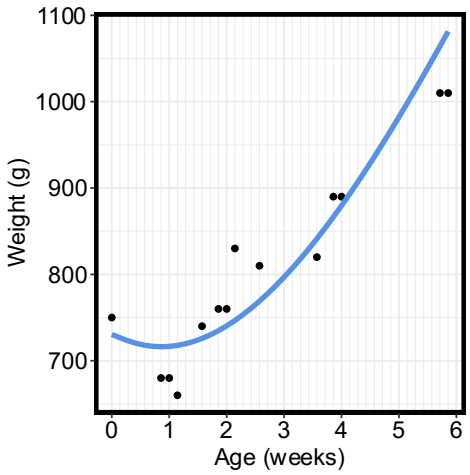


Figure 3-8. Natural cubic spline model (with two knots) for smoothing weight values of the infant also used in Figure 3-13. Dots are actual weight measurements and blue line is the smoothed cubic spline curve.

This curve can be seen to describe the growth curve more naturally than the polynomial models, although the strictures of natural cubic spline modelling mean that it does not fully capture the initial weight loss of this infant. The simplicity of the curve prevents overfitting. Cubic b-splines were also generated but frequently did not form plausible growth curves, with very large deviations from actual weight measurements, especially where data were more sparse than usual.

Broken Stick Model

The broken stick model seeks to smooth and interpolate data by connecting a number of straight lines. It has been applied to growth where there are abrupt changes in growth and where segments can be described linearly.[124] Figure 3-9 shows broken stick modelling of the same infant as used above, with varying numbers of knots.

In a similar phenomenon as seen with polynomial modelling, models with too few knots were grossly underfitted, but when sufficient knots were used to capture early weight loss (seven), the remaining points were significantly overfitted. The sharp inflection (for example as seen in the right-most plot) is not a plausible pattern for the weight gain of an infant.

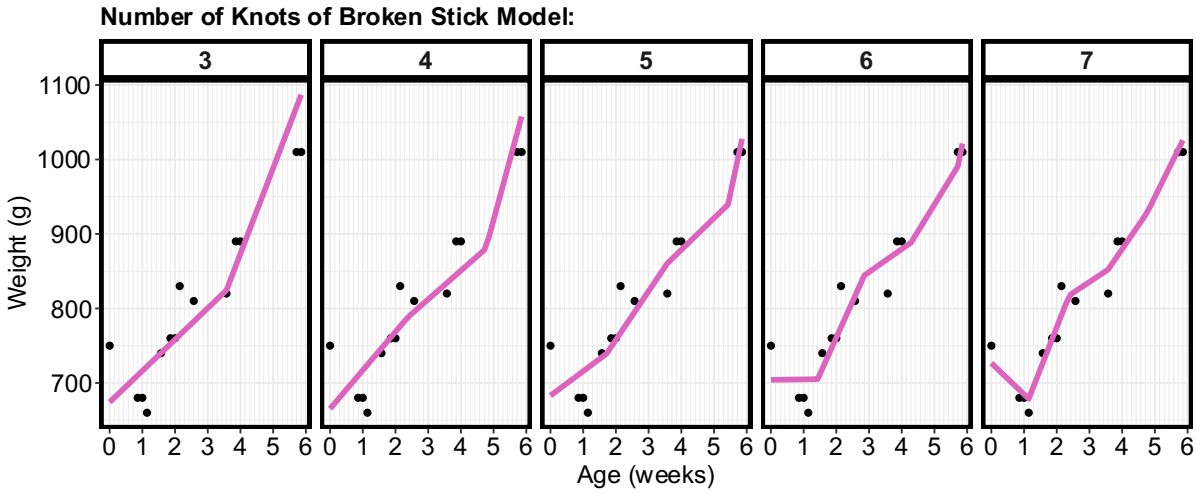


Figure 3-9. Broken stick model of the growth of the individual infant also used in Figure 3-13 and Figure 3-14, at varying number of knots. Black dots are actual weight measurements and pink line represents the smoothed model.

Model Selection

Each variation of each model was superimposed onto the actual growth measurements and was visually inspected for underfitting and overfitting (as illustrated in Figures 3-13 to 3-15). Polynomial models were underfitted below the fourth degree but often overfitted above the third degree. Broken stick models did not form plausible growth curves as described above. Natural cubic spline curves were resistant to overfitting at all knot values. Occasionally, they seemed to underestimate the initial weight loss of infants. Interpolations were also made by forming a SITAR model of the whole cohort and then interpolating values based on the random effects assigned to each infant. This method was too inflexible to accurately reflect unusual but plausible growth patterns (such as rapid early weight gain followed by weight plateau). After consideration of all tested approaches to modelling, natural cubic splines (with two knots) were selected for use in further analysis of the influences on early weight gain. Head growth and length growth generally follow simpler growth patterns, meaning that use of simpler models (especially simple

polynomials) may be adequate to model their growth patterns. However, the natural cubic spline method is flexible enough that it could be used for all measurements.

3.4 Methods for Assessing Associations

A key element of this doctoral project was to assess the association between clinical and nutritional variables and the growth of preterm infants. The most longstanding method of assessing for such associations is linear regression. Modern computing technology and mathematical techniques have facilitated the development of machine learning algorithms, including supervised machine learning algorithms which can be trained to predict outcomes (dependent variables) based on inputs (independent variables) and unsupervised models which seek to discover novel patterns. Each of these methods is considered in turn below.

Linear and Logistic Regression Modelling

Linear regression describes the relationship between a scalar response (i.e. a continuous dependent variable) and an explanatory variable (which may be continuous or categorical). It assumes that there is a linear relationship between the explanatory variable and the response variable, calculating the best linear fit line, usually employing the least squares method.

Multiple linear regression is an extension of this method, allowing multiple explanatory variables, each presumed to have a linear relationship with the response variable. This allows for the determination of the effect of each explanatory variable independent of the other explanatory variables (often described as the effect size when other variables are “held fixed”). Multiple linear regression models have advantages over more complex machine learning models. They require very modest computational resources, even when cohort sizes are large. The effect size is also easy to understand. For example, the result of a multiple linear regression model for the effect of weight gain velocity on protein intake might conclude “After adjustment for gestation, every increase in 1g/kg/day of protein intake was associated with a rise in growth velocity of 3g/kg/day”. Therefore, multiple linear regression has been used in this project when the priority has been to describe the influence of a particular factor on continuous outcomes of interest.

Logistic regression assesses the association between one or more explanatory variables (which may be continuous or categorical) and one binary response variable. It calculates the change in the probability of the response variable associated with a change in the explanatory variable. Such models are useful when trying to predict a binary outcome. They can be more difficult to understand than a linear regression model, as the output is the change in the log of the odds of the outcome in response to a change in the explanatory variable. Simple conclusions can be drawn, for example “The chance of weight dropping more than two centile spaces is reduced by

increasing protein intake, after adjustment for birthweight". In practice, such models may be useful when risk prediction is the desired output. In this project, logistic regression models have been formed when training models to predict a binary outcome, for example in efforts to risk-stratify infants based on their characteristics. More complex machine learning algorithms also often require a binary outcome and so logistic regression can often be used to provide a baseline for assessing the usefulness of more complex models.

Multiple linear regression and logistic regression were performed in R.

Machine Learning Approaches

Machine learning models seek to classify individual cases based on a set of explanatory variables. In supervised machine learning, models are provided with the outcome variable (typically a binary outcome) and seek to discover relationships between the explanatory variables and the outcome in question. In unsupervised machine learning, individuals are grouped according to features discovered by the model. Both approaches were used in this project. Supervised techniques were used in a similar way to logistic regression, predicting the risk of a particular growth outcome based on the explanatory variables (which may be clinical or nutritional). Unsupervised learning encompasses clustering techniques such as k-means clustering, which was explored as a means to classify the growth outcome of the infants in the study. The following sections describe the underlying principles of selected machine learning techniques and their application to this project.

Random Forests

Random forests are usually implemented as supervised methods for classification into categorical groups using decision trees. Decision trees are algorithms which predict the allocation of a subject to a category using a series of decision points based on predictor variables. The random forests approach generates a large number of decision trees. Predictions are then based on a "voting" system in which the returned prediction for group allocation is decided by the result of the majority of the generated decision trees. Random forests have been used in healthcare research, especially in risk prediction models.[125] They have been used in one study of the growth of preterm infants.[126] Random forests have the benefit of relatively simple set-up, with few requirements for data transformations or complex data preparations. Importantly, the "voting" system of many individual trees reduces the impact of covariance on the outcomes of the model. In this project, I used the *randomForest* package[127] for R to generate random forest models.

Interpreting the Results of Random Forest Models

In common with many machine learning approaches, random forests report their predictions without explaining the contribution of individual factors, as issue described as the "black box"

problem.[128] Observational research often seeks to understand the influence of specific factors on an outcome of interest. Traditional black box machine learning methods do not provide this information as a matter of course, limiting their usefulness. In response to this limitation, techniques have been developed to adapt machine learning techniques or to apply post-hoc analysis to explain the way that models reach their predictions, a field known as explainable artificial intelligence or interpretable machine learning.

One popular approach to interpreting random forest models is SHAP (Shapley additive explanations). The SHAP method assesses the impact of individual features (i.e. predictor factors) on the final prediction of the model. This provides a measure of the importance of each predictor and the size of the effect at different values of the predictor. In this project, I used the *treeshap* package[129] for R, which uses a modified form of the SHAP process which is optimised for tree-based models. It is important to remember that machine learning models are reliant on the underlying data supplied to them and testing using an independent dataset would be required to assess their generalisability to larger populations.

K-Means Clustering

K-means clustering is an unsupervised machine learning method seeking to identify groups in data. Individuals are plotted by the values of the chosen variables (which have been scaled). An iterative process is then employed to group the individuals in space to minimise the distances between individuals in the same cluster whilst maximising the distances between individuals in different clusters. The number of clusters must be decided by the user, but algorithms exist to determine the optimum number of clusters. The elbow method can be used to decide the number of clusters by visually assessing the number of clusters beyond which there is little improvement in the difference between within-cluster variation and between-cluster variation. This process can be used to classify subjects by predictor variables or by outcome. It is useful for reducing dimensionality in highly dimensional data. For example, the intake of different nutrients may be very highly correlated with each other, causing problems for regression techniques. K-means clustering can be used to divide infants into a small number of groups with differing patterns of nutrient intake.

Chapter 4 sets out the results of applying these methods to the SPND.

Principal Component Analysis (PCA)

PCA is a machine learning approach which aims to capture the variance of data using fewer dimensions than are included in the source data. It is especially useful when data have a very large number of variables (i.e. highly dimensional data) and when variables are collinear.

In theory, each observation is plotted in a multidimensional space which each dimension consisting of a variable in the original dataset. A line can then be drawn through this hyper-dimensional space which maximises the variance between sample values (i.e. the distance between them in this space). The position of an observation along this line is the value for the first principal component. A second line can be drawn (orthogonally to the first, to minimise covariance between the principal components) and the position of an observation along that line is the value for the second principal component. In practice, it is very difficult to visualise or conceptualise this process when more than three or four dimensions are used (i.e. more than three or four source variables). In practice, there are well-developed statistical packages (for example, the *prcomp* command in R) which can quickly calculate principal component values for very highly dimensional data. Data must first be scaled and centred. Analysis of principal component models can also provide "loadings", the influence of each of the original variables on the value for each principal component. These loadings indicate which of the original variables contributed most to the variance between observations.

In summary, the first component is a single number which represents the most possible variance within the underlying data (a linear regression value of all source variables). The second component is the next-best discriminator, with the restriction that it must be orthogonal to (and therefore not covariant with the first). Each following principal component follows this pattern. Thus, a series of non-collinear values are generated for each subject which explain the maximum proportion of the variation in the original data in the fewest possible components.[130]

In this thesis, PCA was used when the intended analysis technique (e.g. linear regression) did not cope well with data which had many variables, with significant covariance between the variables. These factors were present especially in metabolomic and genomic data, which often contained hundreds or thousands of variables. PCA provided a means to summarise these data into fewer variables whilst retaining as much variance as possible and removing covariance. The loadings of the PCA models also provided insights into the variables which contained the most variance.

Chapter 4 Modelling Influences on Growth

This section deals with assessments of clinical and nutritional factors on the growth of very preterm infants. The specific methods used in this chapter are described. The results section opens with a detailed description of the cohort of infants in the Southampton Preterm Nutrition Database, followed by the preparation and assessment of data, and concludes with general linear regression and machine learning-based techniques. The results and limitations are then discussed.

The work in this chapter was carried out by Aneurin Young with advice from his supervisors and from Dr Guo Cheng.

4.1 Background

The growth of preterm infants is influenced by myriad interacting factors. Nutrient intakes are bound to be influential, but differences in the demographic features of the infant, the severity and type of their comorbidities and the care which is provided to them are also likely to exert powerful influences on their growth.

Taking gestation as an example, several studies have identified differences in the patterns of weight gain seen in infants born at different gestations.[25, 131] Major morbidities, such as chronic lung disease and necrotising enterocolitis, have also been shown to have profound effects on the early growth of preterm infants.[132] Commonly used medications may also have an impact on growth. Particular attention has been paid to dexamethasone, a corticosteroid medication which is commonly used to treat chronic lung disease in preterm infants. The effect of dexamethasone on growth is likely to be dose-dependent and recent studies using different protocols have found differing effects, from no impact on growth[133] to transient inhibition of growth followed by catch-up growth[134] to persisting growth deficit.[135] Factors present during pregnancy may also influence growth. For example, maternal smoking during pregnancy is an established risk factor for intrauterine growth restriction. However, the influence of smoking on postnatal growth is less certain, with one study suggesting faster postnatal growth in term-born infants exposed to maternal cigarette smoking *in utero*.[136]

The influence of nutritional intake on the growth of preterm infants has been investigated using observational and experimental techniques and the key studies in this field are summarised in the first chapter of this thesis. In summary, manipulation of parenteral nutrition to provide higher

macronutrient intakes has generally resulted in faster growth of preterm infants. Likewise, supplementation of breastmilk with a multicomponent fortifier promotes faster growth.

Shortly after birth, infants exhibit a diuresis, leading to reduction in total body water and a resultant early weight loss. The magnitude of weight loss to be expected is uncertain and different patterns of weight loss have been identified in different cohorts of preterm infants.[25, 131] However, there is a growing consensus that assessment of weight gain should exclude this early period of weight loss.[137]

This chapter aims to use detailed observational data to define which factors are influential on the weight gain and head circumference growth of preterm infants. Quantifying these effects will identify modifiable care approaches which could improve the growth of preterm infants.

These observational data contain complex covariance and collinearity patterns, presenting challenges to traditional linear regression approaches. Therefore, random forest modelling and SHAP approaches were used, as described in the shared methods chapter above.

4.2 Aims of Modelling Influences on Growth

The aims of this chapter are:

- To describe the specific cohort of infants included in the Southampton Preterm Nutrition Database
- To describe covariance and collinearity between clinical and nutritional factors
- To identify demographic, clinical and nutritional factors which influence the weight gain and head circumference growth of preterm infants.

4.3 Methods

All infants in the Southampton Preterm Nutrition Database were included in analysis. Age at weight nadir was calculated for each infant. Change in weight z-score and change in head circumference z-score were calculated from day seven of life (when almost all infants had reached nadir weight) to day 42 of life for each infant.

Weight gain and head circumference growth were modelled and smoothed for each infant using cubic splines. The z-score for each measurement at day seven of life was subtracted from the z-score at day 42 of life to derive measures of growth over the first six weeks of life.

Where a dichotomous outcome of overall growth was required, infants exhibiting a gain or loss in z-score of less than one third for both weight gain and head circumference were defined as having experience adequate growth. This measure split infants into two groups which were close in size.

As these measurements were made during the first six weeks of life, this classification roughly reflects infants who are on pace to drop one whole centile space (a drop of two thirds of a z-score) prior to reaching term corrected age.

Weight gain and head growth were also summarised by week, with the change in weight z-score and the change in head circumference z- score calculated during each week-long period and the mean nutrient intake calculated during the week in question. For a dichotomous outcome, a drop of 0.1 or less was used to define adequate growth (being roughly equal to a drop of two thirds of a z-score across the whole time period). Neurodevelopmental outcomes at two years corrected age were acquired from BadgerNet.

Covariance between nutrients was explored using correlation coefficients. Daily nutrient intake values (per kg of body weight of the infant at the time) were extracted from the SPND and Pearson correlation coefficients were calculated between each pair of nutrients. Coefficients were calculated using the *rstatix* package for R. Mean nutrient intake values during the first six weeks of life were calculated for each infant and these mean values were subjected to the same analysis. Variance inflation factors (VIFs) are a measure of how correlated a factor is with the other factors used in a model. VIF values were calculated for generalised linear regression models using the *vif* function of the *car* package for R. VIF values greater than 5 were taken to indicate significant covariance.

Table 4-1 shows the predictor variables included in regression modelling.

Table 4-1. Predictor variables used in regression models.

Gestational age at birth (weeks)
Sex
Birthweight Z-score
Whether absent or reversed end-diastolic flow in pregnancy
Admission temperature (°C)
Maternal age (years)
Whether mother smoked in pregnancy
Whether antenatal steroids given
Apgar score at 1 minute of life
Apgar score at 5 minutes of life
Whether antenatal magnesium sulphate given
Index of multiple deprivation rank
Days receiving supplemental oxygen
Days of invasive ventilation
Days of non-invasive ventilation
Days of dexamethasone treatment
Days of hydrocortisone treatment
Days of treatment with diuretic drugs
Days of treatment with inotropic drugs
Days of treatment with insulin
Days of treatment with inhaled nitric oxide
33 individual macro- and micro-nutrients as described in the methods chapter above
Days receiving breastmilk
Days receiving parenteral nutrition
Minimum blood haemoglobin value
Minimum serum sodium
Year and month of birth
Whether diagnosed with necrotising enterocolitis
Whether diagnosed with patent ductus arteriosus
Whether diagnosed with intraventricular haemorrhage (grade III or IV)
Whether singleton or multiple pregnancy

Generalised Linear Regression

Multiple linear regression of change in weight and head circumference z-score was performed, including all the above predictor variables, using the *lm* function of R. Logistic regression of the binary outcomes used the *glm* function. 10-fold cross-validation with five repeats was used to reduce the effect of overfitting of linear regression models to the underlying data (with no

significance change in results seen when more folds were used). Nutrient intakes were summarised by taking the mean average intake of each nutrient during the first six weeks of life.

Random Forests

Random forest models were created to assess relationships between nutrient intakes, clinical variables, weight gain and head growth. The *randomForest* (v4.7) package for R, with the number of trees set to 500. In most cases, fewer than 500 trees are required, but this number was chosen as it was very likely to optimise accuracy but nevertheless required less than one minute of computing time to form the models.

SHAP values were calculated and visualised using the *treeshap* (v0.1.1) and *shapviz* (v0.2.0) packages for R.

4.3.1 Clustering by Feeding Approach

An alternative strategy to reduce the impact of nutrient intake collinearity was to instead search for patterns of feeding approach associated with weight gain and head circumference growth. Daily feeding data were summarised to provide the volume of each of the following product types for each day of life for each infant: parenteral nutrition, breastmilk, fortified breastmilk and infant formula.

K-means clustering was performed using the *kmeans* command in R. The optimum number of clusters was identified as four using the elbow method. Associations between weight gain, head growth and feeding approach cluster were explored using multiple linear regression.

4.4 Results

4.4.1 The SPND Cohort

At the time of submission 621 infants are included in the Southampton Preterm Nutrition Database (SPND) arising from the SPIN and GAP studies (as described in Chapter 3). Table 4-2 sets out basic demographic information for those infants.

Table 4-2. Demographic details of infants in the Southampton Preterm Nutrition Database.

Gestation (weeks + days), median (range)		27^{+5} (23^{+0} to 31^{+6})
Birthweight (g), mean ±SD		981 ±283
Male, number (%)		340 (55%)
SGA*, number (%)		149 (24%)
Born at UHS, number (%)		554 (89%)
Multiplicity of Pregnancy, number (%)	Singleton	411 (66%)
	Twin	178 (29%)
	Triplet	25 (4%)

* Small for gestational age (birthweight below 10th centile according to reanalysed UK1990 data[115]).

Figure 4-1-A illustrates the number of infants born at each completed week of gestation.

The distribution of raw birthweights was slightly right-skewed (Figure 4-1-B) and would not be expected to show a normal distribution due to the distribution of gestations. The reanalysed UK 1990 birthweight dataset[115] was used to calculate birthweight z-scores (i.e. the deviation from the median birthweight for the gestation, expressed as number of standard deviations). The distribution of these z-score was left-skewed (Figure 4-1-C) and the rate of small for gestational age (SGA; birthweight z-score less than 10th centile) was 24%, indicating an excess of infants with intrauterine growth restriction (IUGR). Head circumference distributions were less skewed (Figure 4-1-D-E).

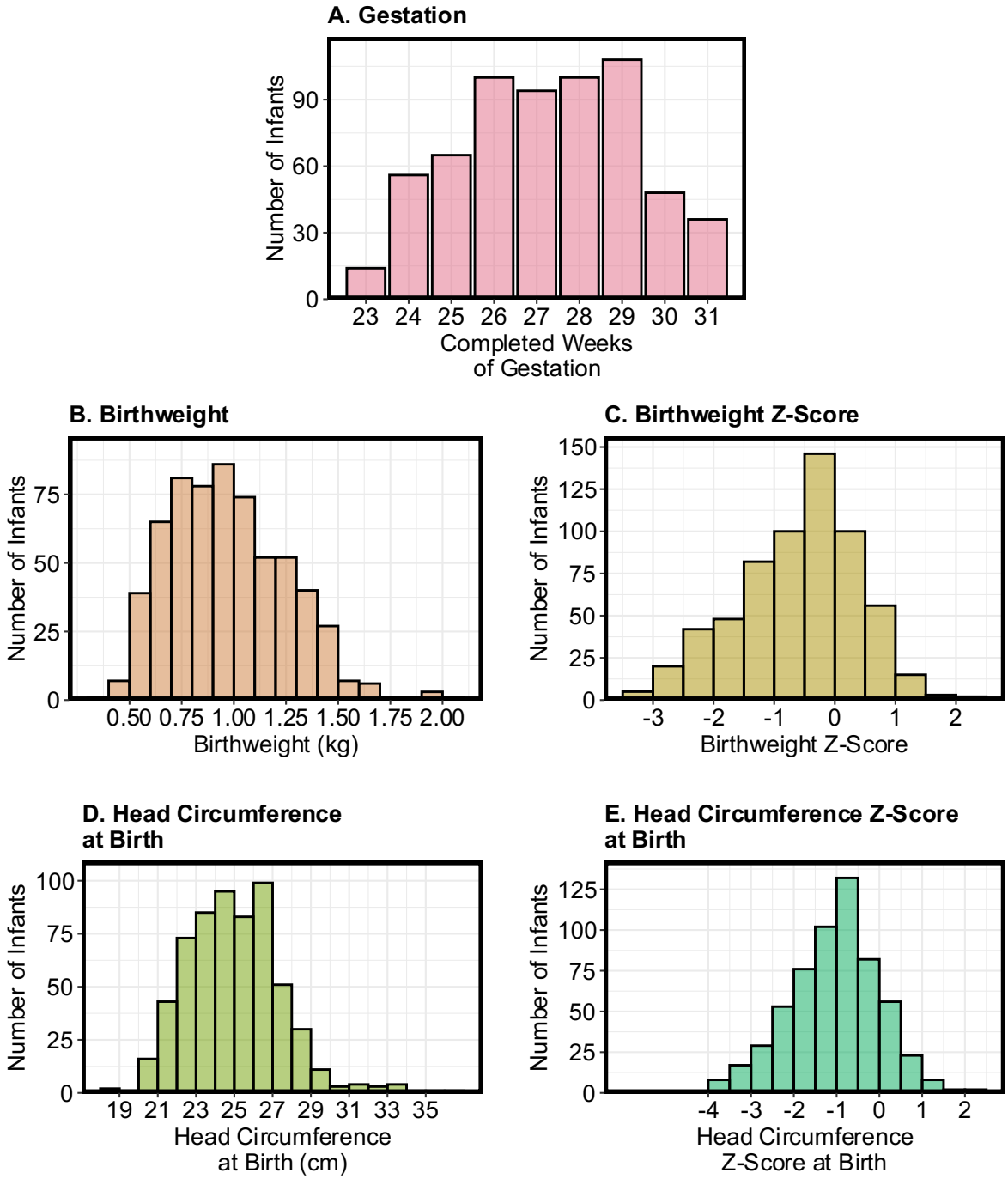


Figure 4-1. Demographic features of infants included in the Southampton Preterm Nutrition Database. A. Number of infants born at each gestation; B. Birthweight distribution; C. Birthweight z-score distribution; D. Head circumference at birth distribution; E. Head circumference z-score distribution.

Table 4-3 describes the discharge destination of the 593 infants included in the SPND. Overall mortality for the cohort was 6%, with around half of surviving infants being discharged home from UHS and half being transferred to other healthcare settings.

Table 4-3. Discharge Destination of infants included in the SPND.

Discharge Destination, number (%)	Home	302 (49%)
	Other hospital	282 (45%)
	Died	37 (6%)
	Unknown	0
Mortality by Gestation, number (%)	23 weeks	5 (36%)
	24 weeks	6 (11%)
	25 weeks	6 (9%)
	26 weeks	12 (12%)
	27 weeks	0 (0%)
	28 weeks	2 (2%)
	29 weeks	3 (3%)
	30 weeks	2 (4%)
	31 weeks	1 (3%)

The median length of stay in UHS was 54 days (IQR 30-80 days).

Figure 4-2-A illustrates that the most preterm infants generally had longer periods at UHS than more mature infants. Length of stay for infants born before 24 weeks gestation was lower than those born at 24 or 25 weeks of gestation. However, this was likely due to the small number of infants in that group and much higher mortality. When only surviving infants were included, infants born at 23 weeks of gestation had a similar length of stay to those born at 24 weeks (Figure 4-2-B). Length of stay was shorter for infants who were transferred to other hospitals, illustrating the fact that the database collected complete neonatal stay information for infants of parents local to Southampton, but often captured only the first phase of neonatal care of infants who were “repatriated” closer to home during their neonatal treatment period.

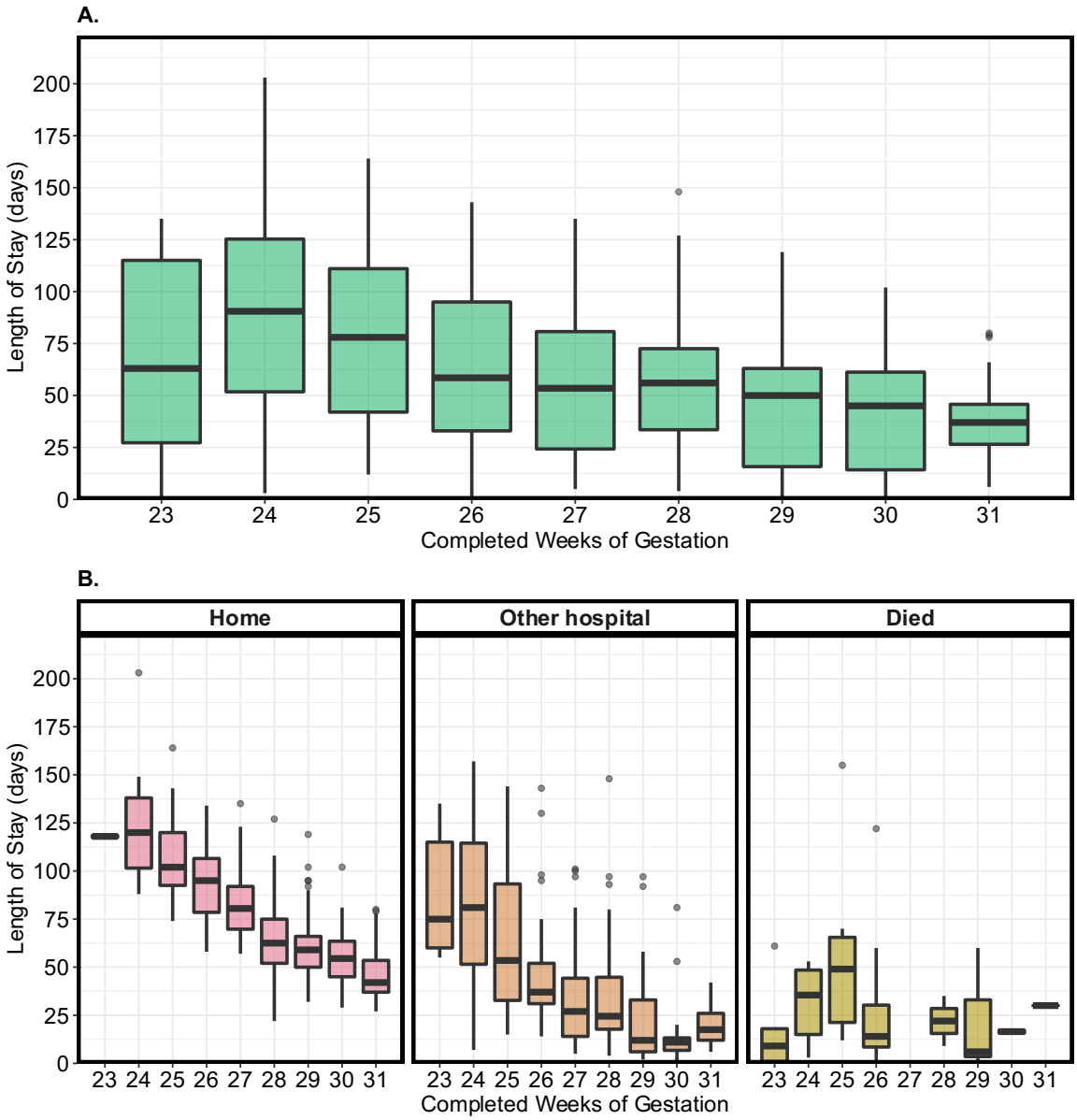


Figure 4-2. Boxplot of length of stay at UHS for A. all infants in the SPND, grouped by gestation; B. Infants separated by discharge destination from UHS. Line indicates median, box encompasses interquartile range, whiskers indicate limits of values within 1.5x IQR and dots are outliers.

Pregnancy and Perinatal Comorbidity Data

Maternal factors included smoking, age and the index of multiple deprivation (IMD) decile of the mother's address. Smoking status was missing for 128 infants (21%), with smoking confirmed in 94 cases (19% of those for whom smoking status was available, similar to the national average). IMD was missing for 29 infants (5%), mostly due to their mother residing outside England.

Mean maternal age at delivery was 30.3 years (very close to the national average for all births which rose during the study period to 30.7 in 2020). Index of multiple deprivation decile (a measure of the level of deprivation in the postcode of the mother's address, available for 95% of cohort) did not show an obvious pattern (Figure 4-3).

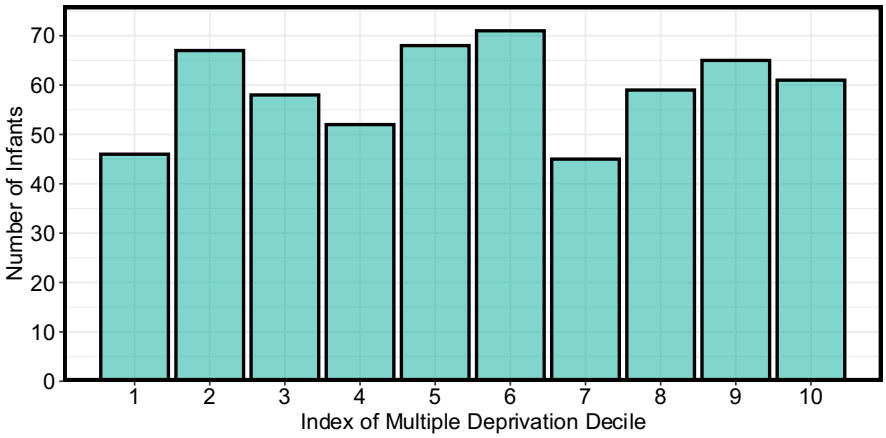


Figure 4-3. Distribution of index of multiple deprivation decile for infants in the SPND (compared to national classifications). Decile 1 is the most deprived and decile 10 is the least deprived.

Apgar scores were recorded as a marker of severity of illness at birth. Apgar scores at one minute and five minutes of life were available for 94% of infants, with the mean score rising from 4.7 to 6.9 between the two timepoints. Apgar score at 10 minutes was recorded less frequently (missing for 144 infants, 24%). Temperature on admission was recorded in all cases. Absent or reduced end-diastolic flow in the umbilical artery was confirmed for 58 infants (9%).

Nutritional Intake

Nutritional intake was recorded for 34,610 care days. Mean intakes of macro- and micro-nutrients are recorded in Table 4-4.

Table 4-4-4. Mean daily nutrient intake for 33 selected macro- and micro-nutrients, expressed per kg of infant body weight per day, with standard deviation from the mean.

Nutrient	Units	Mean Intake, /kg/day (±SD)
Energy	kcal	111 (±34)
Protein	grams	3.04 (±1.1)
Carbohydrate	grams	13.2 (±4)
Fat	grams	5.08 (±2.1)
Sodium	mmol	4.01 (±2.3)
Chloride	mmol	3.2 (±1.8)
Potassium	mmol	2.88 (±1.2)
Calcium	mmol	2.37 (±1.4)
Phosphorus	mmol	2.31 (±1.1)
Magnesium	mmol	0.325 (±0.17)
Iron	μmol	24.3 (±29)
Zinc	μmol	12.5 (±8.2)
Copper	μmol	1.12 (±0.69)
Selenium	nmol	34.4 (±20)
Iodine	nmol	137 (±100)
Manganese	nmol	690 (±940)
Vitamin A	IU	1790 (±1300)
Vitamin D	IU	341 (±270)
Vitamin E	IU	4.0 (±2.6)
Vitamin K	IU	23 (±25)
Thiamine	μg	427 (±270)
Riboflavin	μg	626 (±410)
Vitamin B6	μg	567 (±380)
Folate	μg	35.6 (±19)
Vitamin B12	μg	0.291 (±0.16)
Biotin	μg	4.12 (±1.9)
Pantothenic acid	mg	1.16 (±0.55)
Niacin	mg	3.57 (±1.9)
Vitamin C	mg	35.3 (±26)
Taurine	mg	11.6 (±6.1)
Choline	mg	17.2 (±6.3)
Carnitine	mg	1.18 (±0.85)
Inositol	mg	60.3 (±46)

Nutrient intakes change in relation to the age of the infant and Figure 4-4 demonstrates the rising nutrient intake during the first six weeks of life. These calculations pool intakes from all enteral and parenteral sources. Reference values from ESPGHAN[23] and NICE[54] set out intake recommendations separately for enteral and parenteral nutrition, making it difficult to apply a blanket recommended range of intakes to these data.

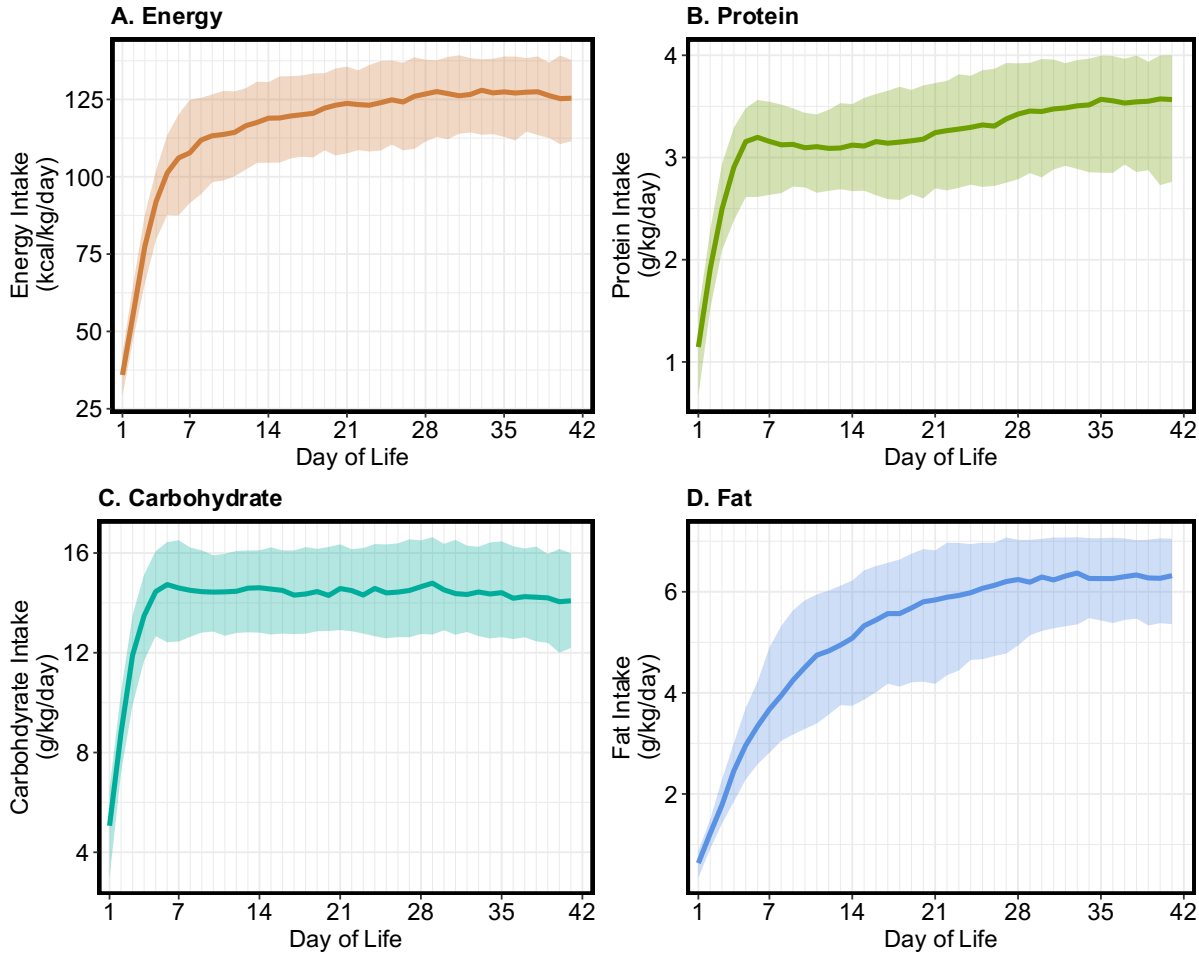


Figure 4-4. Macronutrient intake by age of the infant (starting with first full day of life), solid lines showing the median intake on that day of life and shaded area illustrating interquartile range.

Inclusion of specific feed and fluid volumes in the database provided an opportunity to describe nutritional approach as well as total nutrient intake. Figure 4-5 shows the mean intake of different types of nutritional product as infants became older. It demonstrates the normal practice of starting parenteral nutrition shortly after birth followed by initiation of feeds with maternal or donated breastmilk with subsequent addition of breastmilk fortifier (a multicomponent product designed to provide additional protein, carbohydrate, vitamins and electrolytes to preterm infants fed with breastmilk). Infant formula is avoided during the early weeks of life.

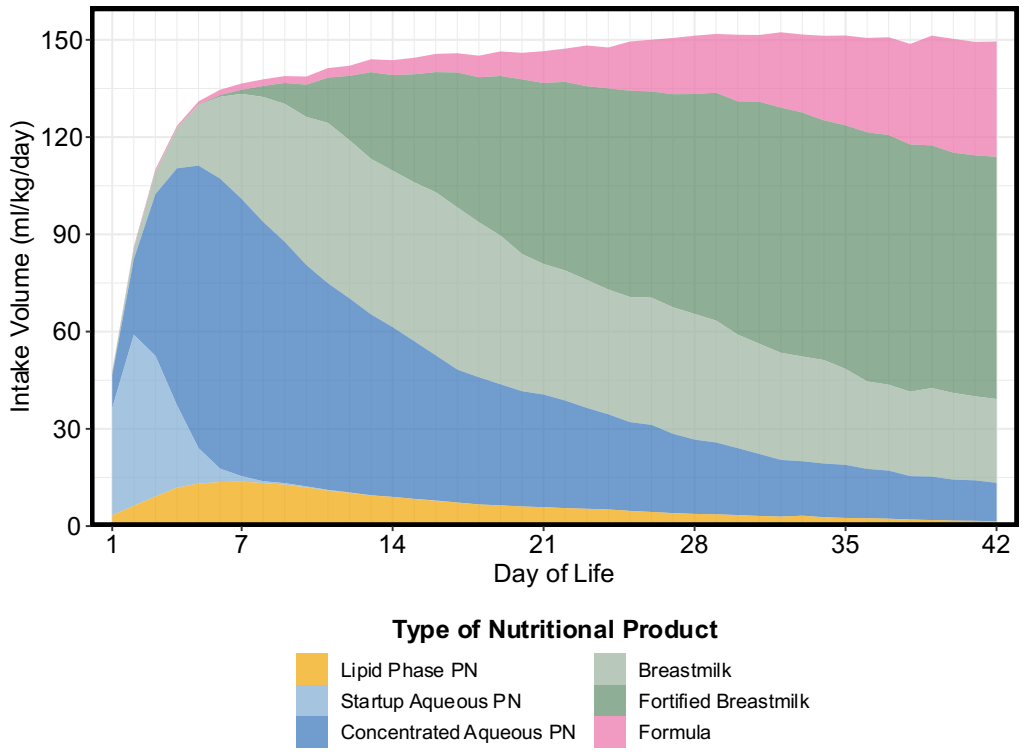


Figure 4-5. Mean intake of types of nutritional products during the first seven weeks of life (starting from the first full day of life). PN – parenteral nutrition.

These data also provided a way to examined relationships between demographic factors and nutritional management. Table 4-5 shows the relationship between gestation at birth and the number of days of parenteral nutrition. Data were not normally distributed due to a minority of infants who received a very long period of parenteral nutrition (62 received more than six weeks of PN and 11 received more than 10 weeks of PN). The median number of days of parenteral nutrition (PN) for the whole cohort was 19 (IQR: 12-29.5).

Table 4-5. Median days of parenteral nutrition received, stratified by gestation at birth (limited to infants who survived to discharge from Southampton).

Gestational Age (completed weeks)	Days of PN, median (IQR)
23	40 (39 to 64)
24	37 (28 to 45)
25	33 (22 to 44.5)
26	24.5 (20 to 32.25)
27	19 (12 to 25)

Gestational completed weeks)	Days of PN, median (IQR)
28	17 (12 to 25.25)
29	13 (9 to 19)
30	11 (7.25 to 14)
31	11 (7 to 16.75)

Clinical Course, Clinical Outcomes and Complications

A range of short-term clinical outcomes were gathered into the database including important complications of prematurity such as necrotising enterocolitis, chronic lung disease, sepsis and intraventricular haemorrhage. Exposure to indwelling vascular access lines and to certain drugs were also recorded. Markers of severity of illness included the number of days of intensive care

and the duration of invasive and non-invasive respiratory support required. 133 infants (21%) had at least one episode of sepsis (as defined in 3.2.1.4).

Growth Outcomes

20,011 weight values were recorded on the database. There was wide variation in the number of weight measurements taken for each infant (median 30, IQR 14-47). There were fewer records for the other measurements (7001 for body length and 7809 for head circumference).

Covariance of Nutrients

Nutrient intakes were analysed on a daily and per-subject level for covariance. Figure 4-7 demonstrates the covariance seen in these data. These figures demonstrate that intakes of most nutrients were strongly correlated with intakes of most other nutrients thus limiting estimates of their individual impact. Some micronutrients were almost perfectly correlated, being delivered in a fixed ratio in nutritional products (e.g. thiamine and Vitamin A had a Pearson correlation coefficient of 0.96). Important macronutrients were also strongly correlated (carbohydrate-protein: 0.87, carbohydrate-fat: 0.62, fat-protein: 0.67). A handful of nutrients were negatively correlated with others (e.g. sodium and chloride, which were strongly correlated with each other but negatively correlated with most other nutrients).

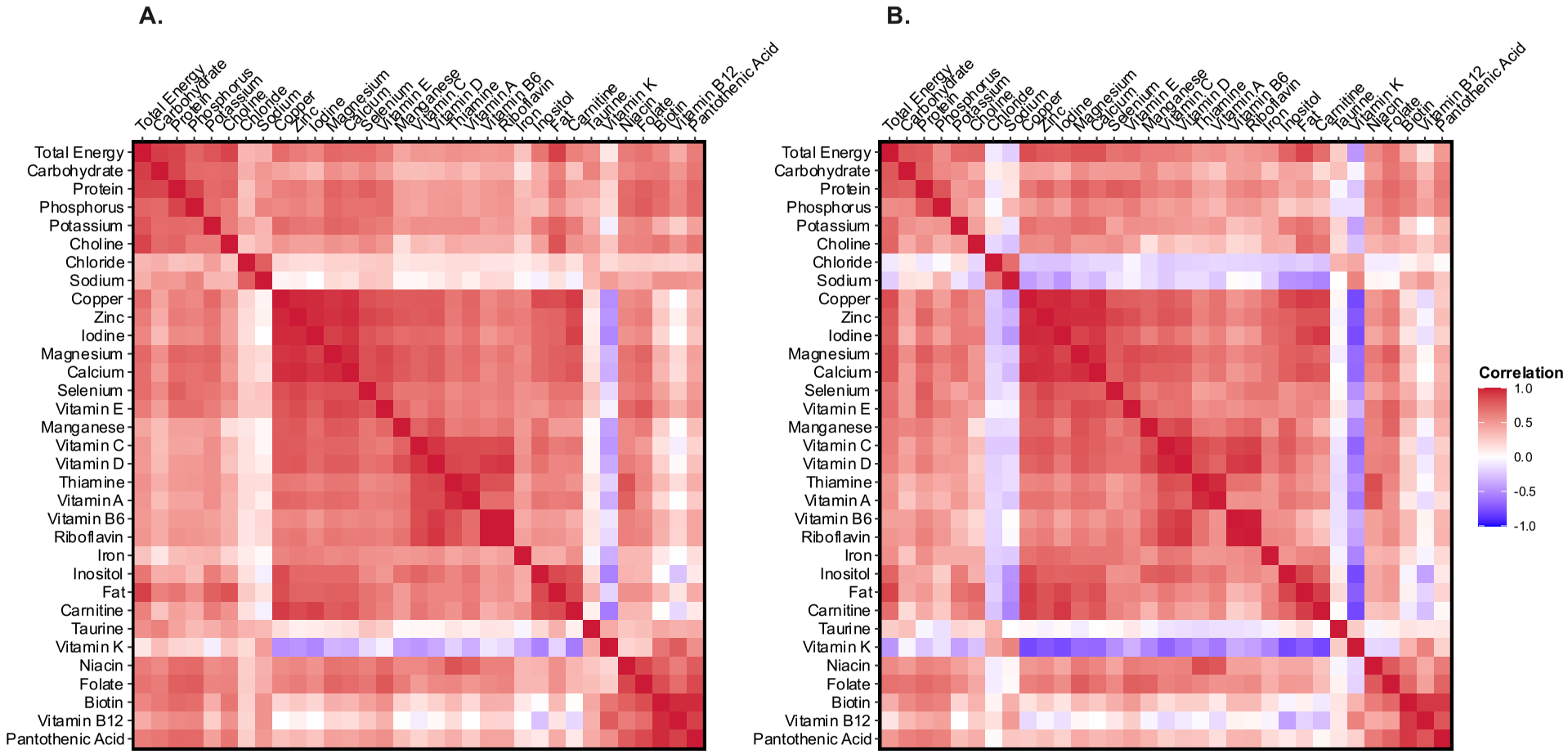


Figure 4-6. Covariance matrices to show Pearson correlation coefficients between nutrient intakes (/kg/day), ordered to highlight clusters A. for every day of nutrition data for infants up to six weeks old (n=20,380 days) and B. mean intakes across the first six weeks of life for infants spending all of their first six weeks in Southampton (n=366 infants).

Generalised Linear Regression Modelling of Nutrient Intakes on Weight Gain and Head Growth

The degree of influence of all nutrient intakes and other factors on change in weight z-score from the seventh to the 42nd day of life was assessed using linear regression and independent variables listed in Table 4-5. Across 333 infants with complete weight data, 26% of the variability in weight z-score change was explained by nutrient intake and other measured covariates (R^2 value of 0.26). while 37% of variance was explained for 219 babies with data for change in head circumference z-score). As expected, VIF values were very high for many of the nutrients and were above 5 for every nutrient except iron.

Random Forest Models

Random forest models were created to assess relationship between weight gain and head circumference growth from day seven of life to day 42 of life and nutritional and clinical variables. Inclusion of all nutrients led to unstable models and so nutrient intakes were limited to three macronutrients (carbohydrate, protein and fat) along with total energy intake.

When weight gain was considered, the model explained 23% of the variance (compared to 26% explained by multiple linear regression). When head circumference growth was considered, the model explained 48% of the variance (compared to 37% explained by multiple linear regression). When the dichotomous variable indicating good weight gain and head growth was used, the out-of-box accuracy (analogous to cross-validated accuracy for logistic regression) was 75%.

Feature Importance and Effects

Figure 4-7 compares the most important features identified by SHAP for each of the three random forest models. For weight gain, protein intake was the most important factor, followed by fat intake and total energy intake. For head circumference growth, the most important factors were markers of clinical severity (with the number of days of invasive ventilation being the most important, associated with slower head growth) and nutritional factors were less important. For the composite dichotomous outcome, the most important factors were a mixture of clinical factors (days of ventilation being the most important) and nutritional factors (protein, energy and fat intake being the second to fourth most important factors).

Figure 4-8 shows the effect sizes of the four most important factors for each random forest model. Figure 4-9 compares the associations for nutrient intakes across the different models. This figure illustrates a plateau effect for protein intake. Increasing protein intake was associated with faster weight gain and with faster head circumference growth, but in both cases this effect plateaued at 3.5g/kg/day, with higher protein intakes not being associated with faster growth.

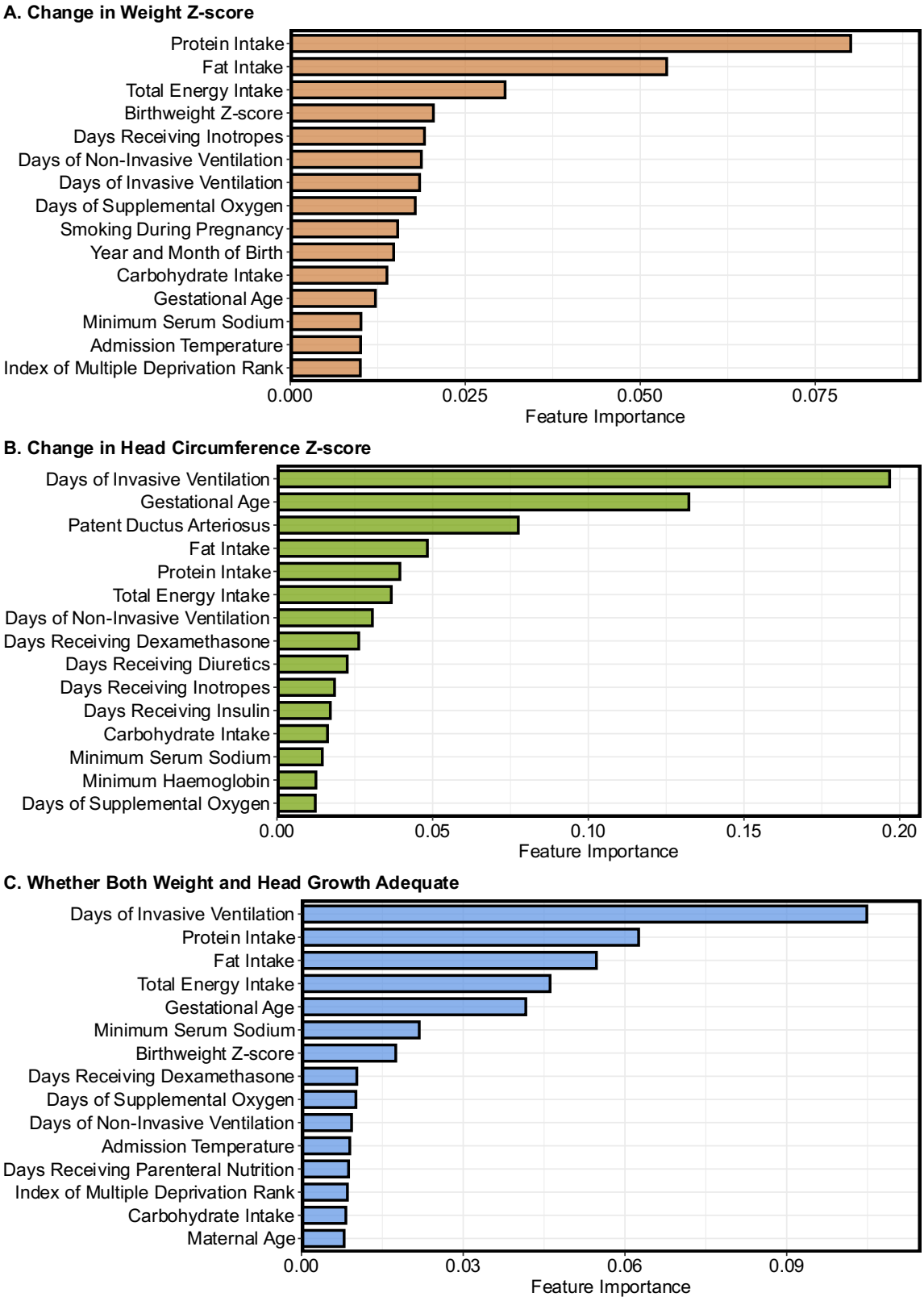


Figure 4-7. The most important 15 features, ranked by importance, for random forest models investigating influences on A. weight gain, B. head circumference growth, and C. a dichotomous variable indicating adequate growth of both, from day 7 to day 42 of life.

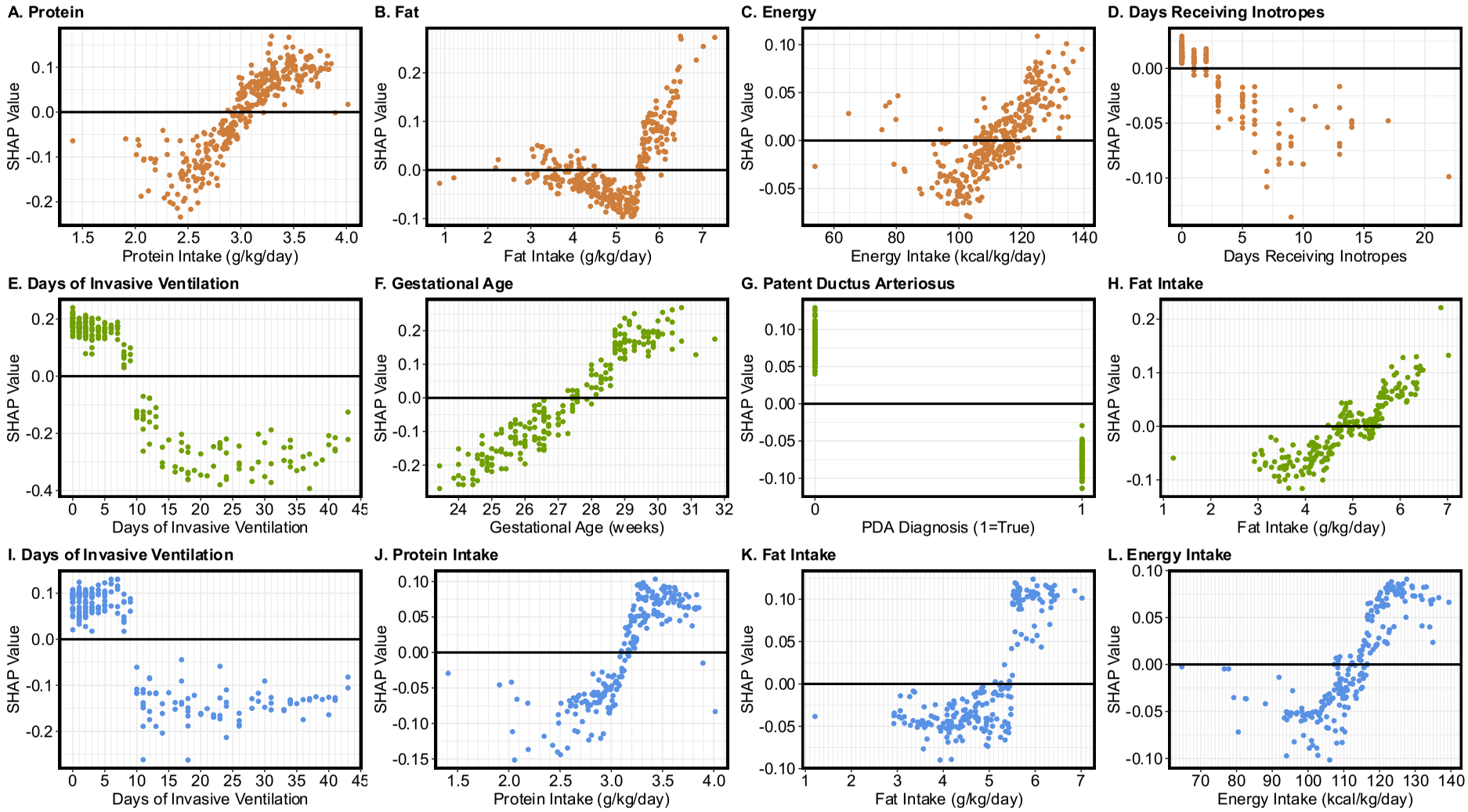


Figure 4-8. SHAP values for the most important factors in random forest models of A-D. weight gain, E-H. head circumference growth, and I-L. A composite dichotomous outcome of adequate growth of both, from day 7 to day 42 of life.

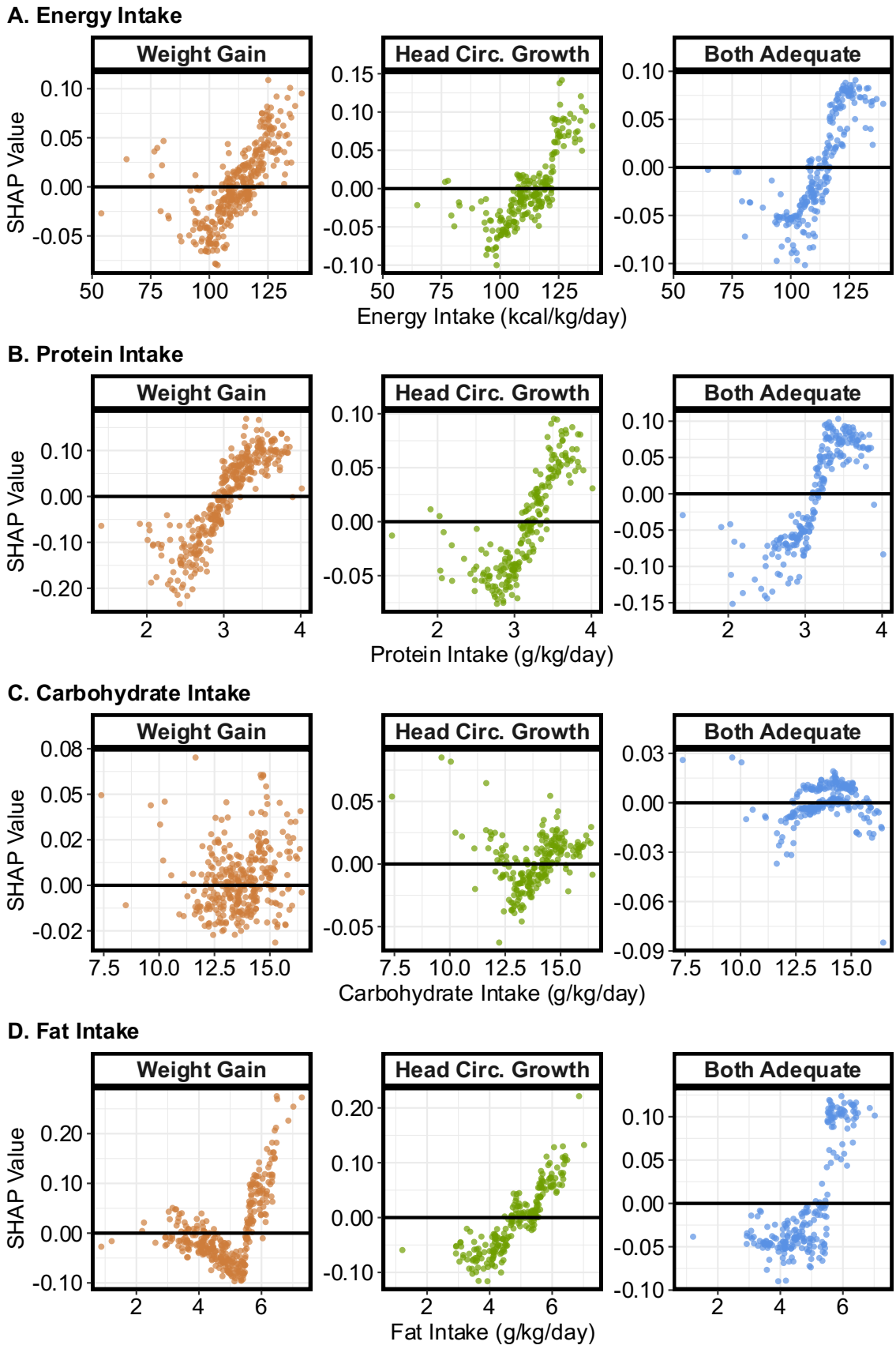


Figure 4-9. Effect of A. total energy intake, B. protein intake, C. carbohydrate intake, and D. fat intake on weight gain, head growth and a dichotomous marker of both growing adequately, from day 7 to day 42 of life.

Random Forest Models for Each Week of Growth

Data were then summarised for each week of life for each infant. This generated 1682 individual weeks of weight gain data and 1211 weeks of head circumference data. A mixed effects linear regression model was created to assess the impact of repeated measures. The random effect was not statistically significant, indicating that taking repeated measures on each individual could be safely disregarded. When weight gain was considered, the model explained 21% of the variance. When head circumference growth was considered, the model explained 52% of the variance. When the dichotomous variable indicating good weight gain and head growth was used, the out-of-box accuracy was 77%.

Feature Importance and Effects

The 15 most important features for each model are illustrated in Figure 4-10. Protein intake was by far the most significant feature for weight gain, whereas clinical factors were more important for head growth. When the categorical adequate growth metric was used, the most important factors were clinical variables, although protein intake was the fourth most significant variable.

Figure 4-11 shows the effects of the four most important factors for each random forest model. Values for weight gain suggest a similar pattern for protein intake as seen in models considering the whole first six weeks combined. Infants with lower birthweight z-scores grew faster (in z-score change terms) than those with higher birthweight z-scores. Advancing postnatal age was associated with faster weight and head circumference growth.

Figure 4-12 compares associations for macronutrient intakes across the different models. Higher energy intake was associated with faster weight gain and head growth. Improvements in weight gain and head growth were not seen when protein intake exceeded 3.5g/kg/day, in a pattern very similar to that seen when the first six weeks of life were combined above. Weight gain was faster with the highest fat intakes (above 6.5g/kg/day), whereas increases in head growth velocity plateaued when fat intakes exceeded 4g/kg/day.

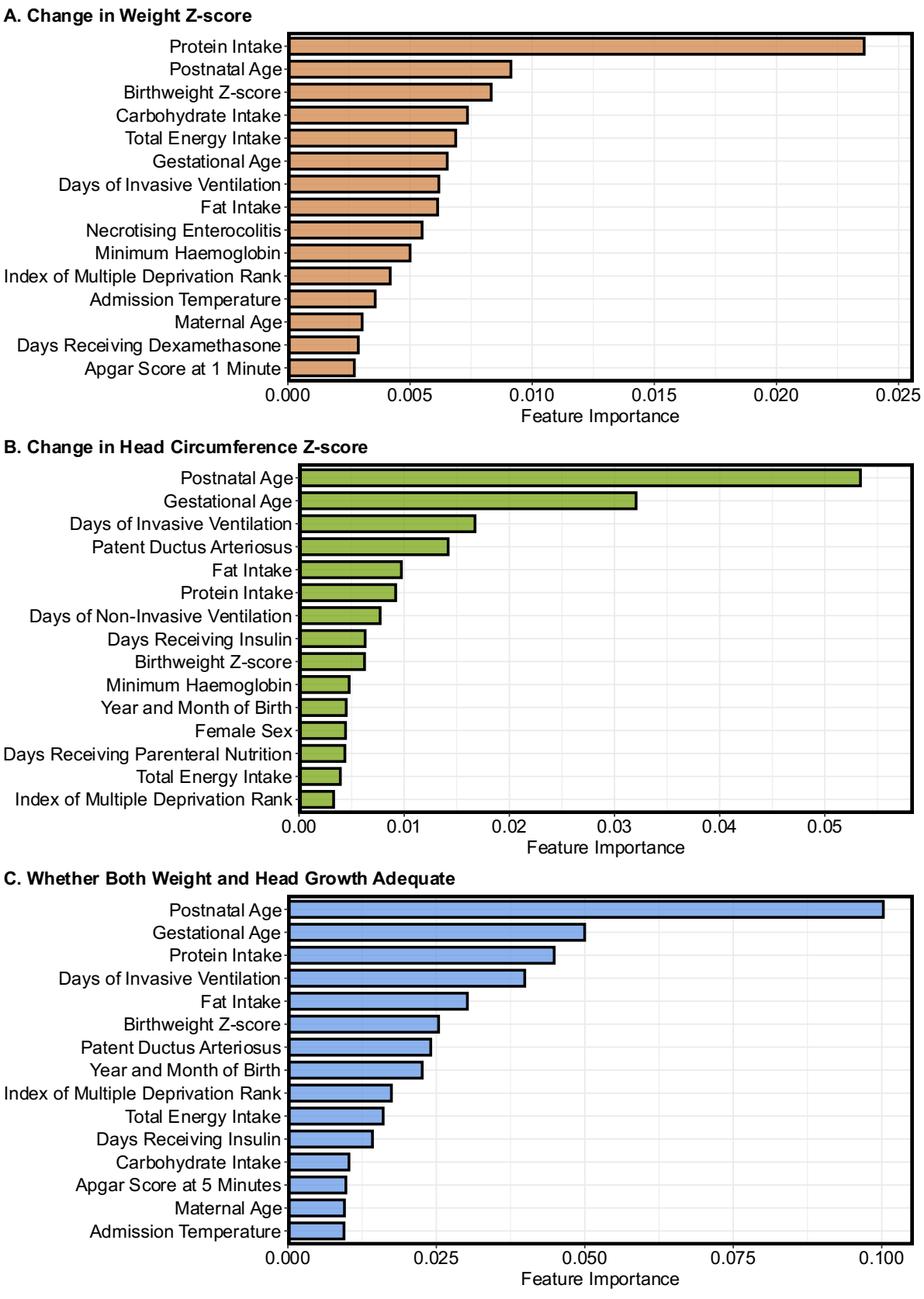


Figure 4-10. The most important 15 features, ranked by importance, for random forest models investigating influences on A. weight gain, B. head circumference growth, and C. a dichotomous variable indicating adequate growth of both, with each week of life summarised for each infant.

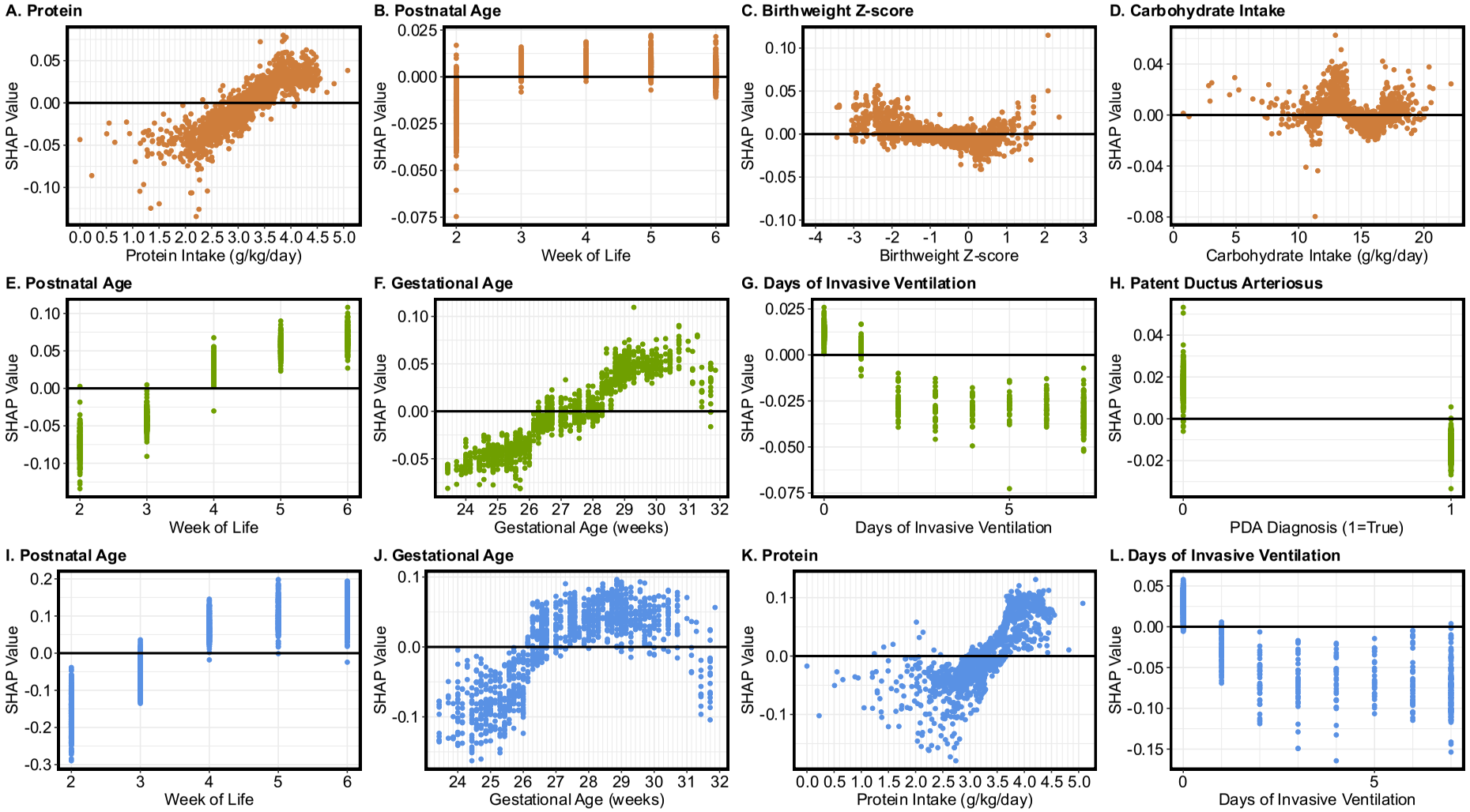


Figure 4-11. SHAP values for the most important factors in random forest models of A-D. weight gain, E-H. head circumference growth, and I-L. A composite dichotomous outcome of adequate growth of both, with each week of life for each infant summarised.

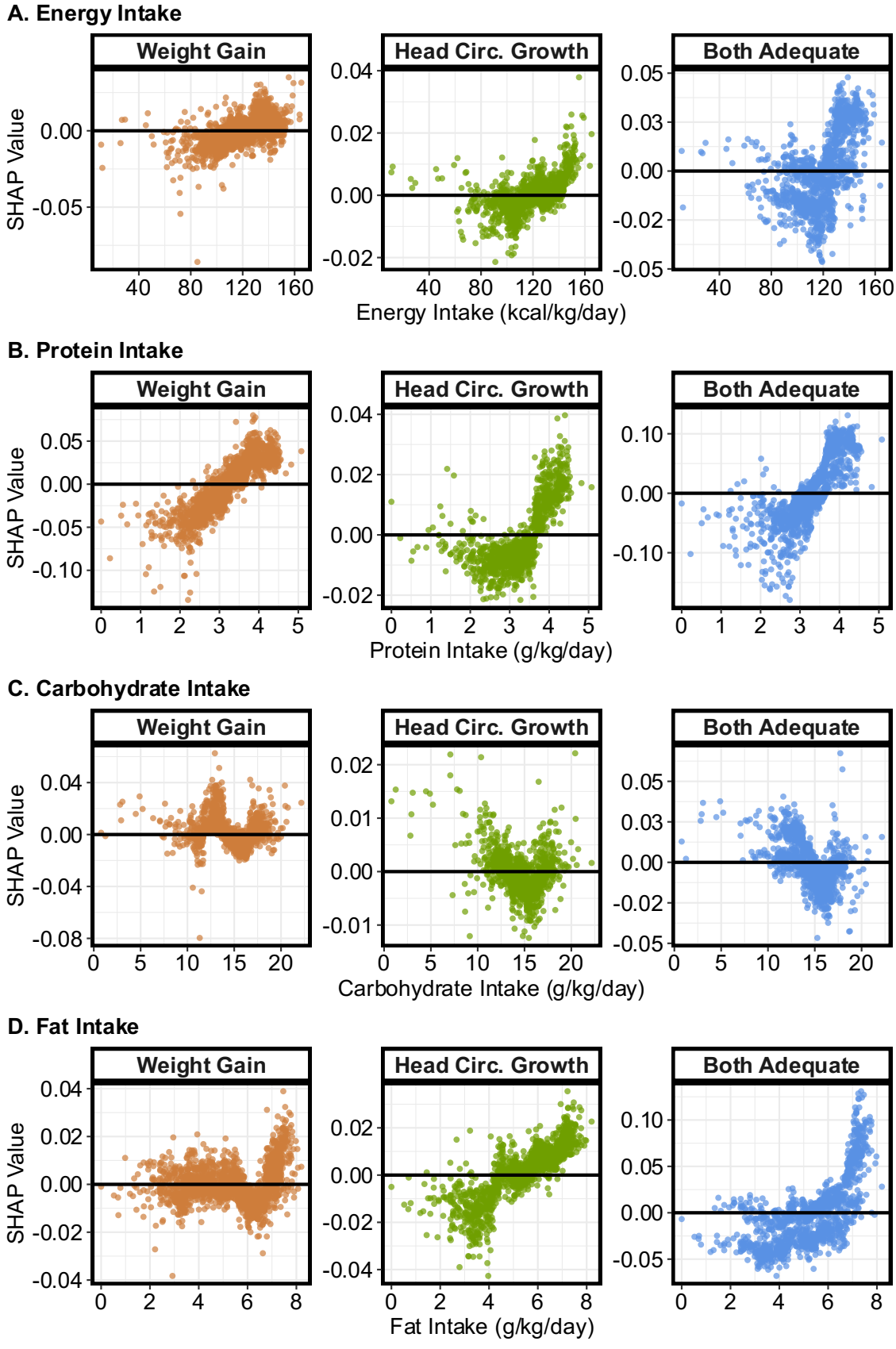


Figure 4-12. Effect of A. total energy intake, B. protein intake, C. carbohydrate intake, and D. fat intake on weight gain, head growth and a dichotomous marker of both growing adequately, with each week of life summarised for each infant.

4.4.2 Clustering by Feeding Approach

390 infants had sufficient information available for the first 42 days of life to be included in feed type clustering. K-means clustering by feeding approach identified four clusters. The mean volume of each type of nutritional product given on each day in each group is illustrated in Figure 4-13.

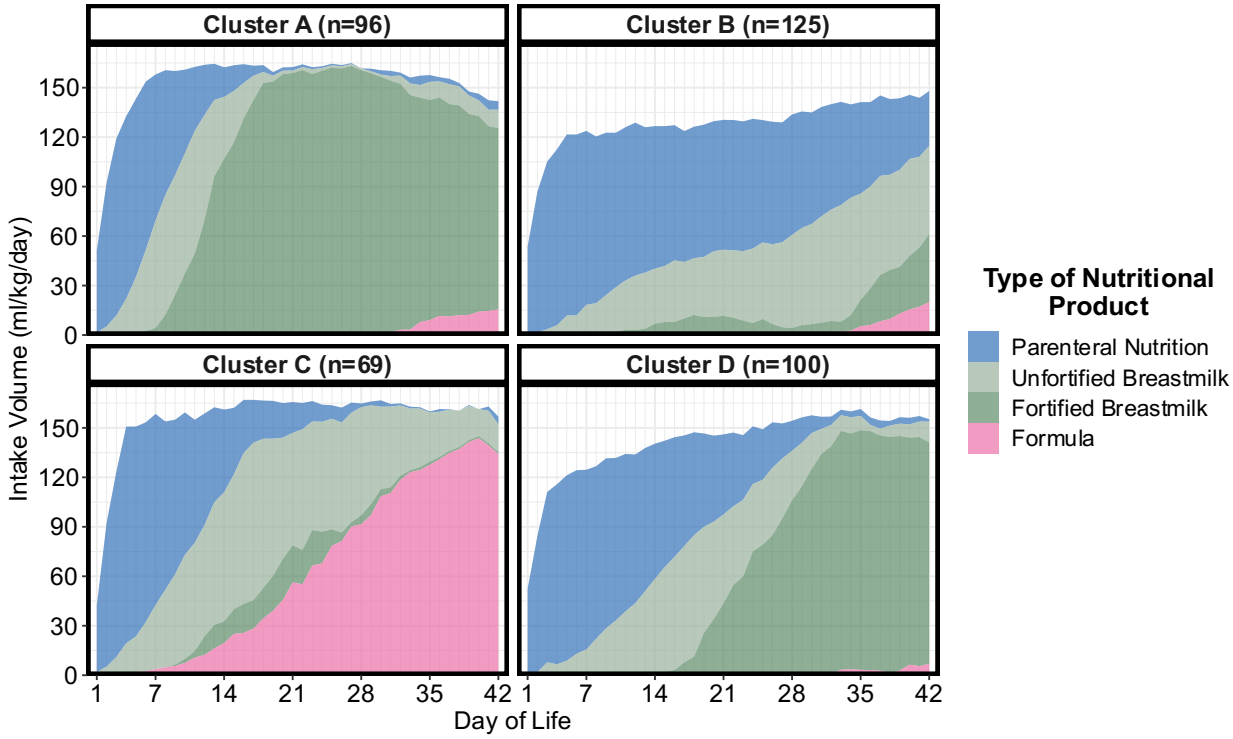


Figure 4-13. Mean intake of four different types of nutritional product on each day of life for each of four clusters identified by k means analysis.

Cluster A represented infants receiving a relatively short period of parenteral nutrition followed by rapid fortification of breastmilk, broadly in keeping with the nutrition guideline in place on the neonatal unit at the time of writing. Cluster B reflected infants having a prolonged period of parenteral nutrition, with some provision of breastmilk but with little fortified breastmilk or infant formula. Cluster C represented a cohort of infants receiving little fortified breastmilk but establishing formula feeding. Cluster D was intermediate between cluster A and cluster B, representing infants with delayed transition from parenteral nutrition to breastmilk and then slow introduction of breastmilk fortifier.

There was a significant association between date of birth and feeding approach cluster, with clusters A and B much more likely in recent years and clusters C and D being common in infants born and recruited earlier in time (see Figure 4-14 for analysis of variance (ANOVA) results). Therefore, year and month of birth (expressed as a decimal fraction of the year of birth) was

included as a covariate to adjust for other changes in care which may have occurred during the time period in question.

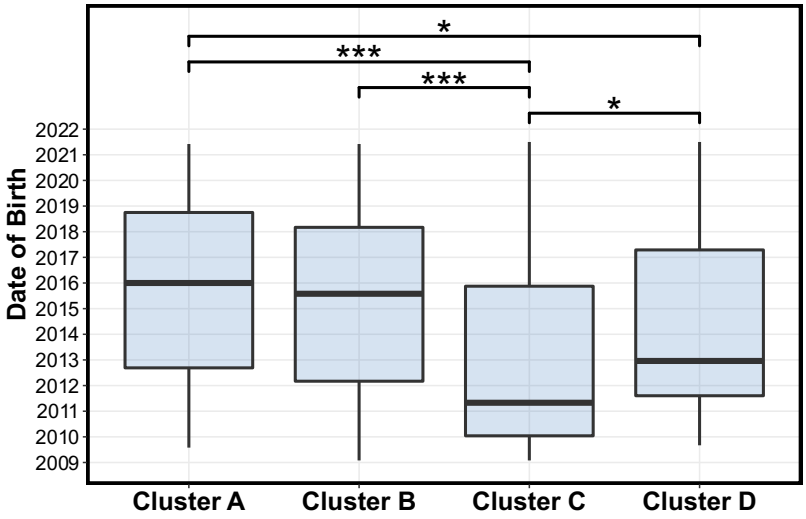


Figure 4-14. Box plot of date of birth by feeding approach cluster. Horizontal lines are the median, boxes encompass the interquartile range and vertical lines indicate range. Comparison by one-way ANOVA. Unmarked comparisons not significant. * p<0.05; *** p<0.001.

Regression of Feeding Approach Cluster and Growth Metrics

Multiple linear regression was used to assess associations between cluster identity and growth metrics for day 7 to day 42 of life, using the same covariates as those shown in table 4-5 but with all nutrient intake values removed, along with the number of days of PN and breastmilk. VIF values were borderline for cluster membership (4.4-5.5 depending on the cohort used for analysis). A covariance matrix indicated several collinearities, with the number of days of central access being very different between clusters (due to the requirement for central access during parenteral nutrition) (Figure 4-15). Generally speaking, infants in Clusters B and D were more premature infants who required more intensive care. Days of central line access was removed as a covariate in further analysis as it was almost completely captured by cluster membership. This reduced VIF values for cluster membership to below 3.5.

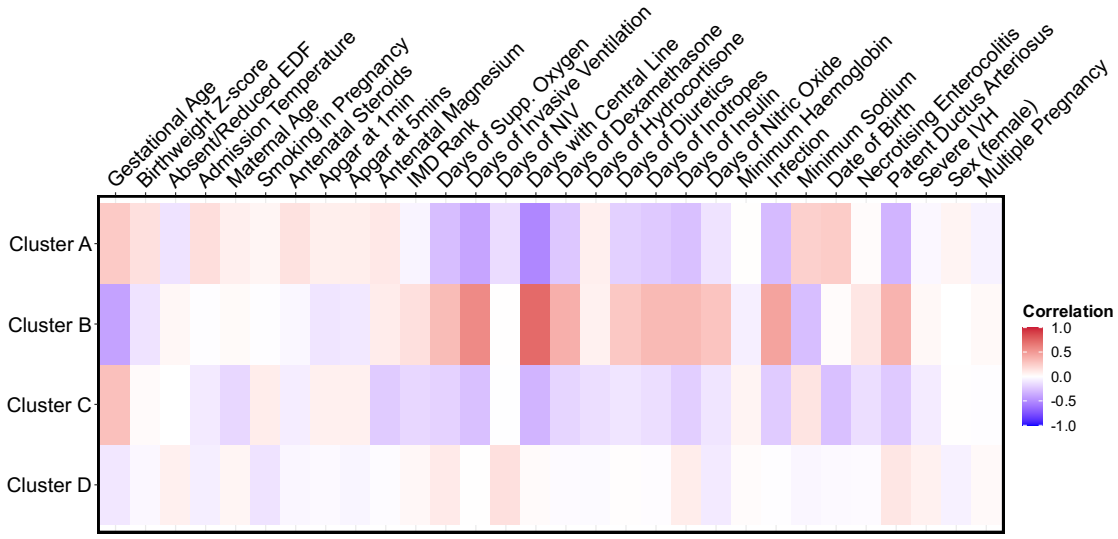


Figure 4-15. Correlation matrix showing systematic differences in clinical variables between nutritional approach clusters.

Multiple linear regression was used to assess the adjusted association between nutritional approach cluster and change in weight z-score from day seven to day 42 of life, with Cluster A taken as the reference group. Covariates were the same as those shown in table 4-5 but with all nutrient intake values removed, along with the number of days of PN and breastmilk. After adjustment, there was no difference in weight gain between Cluster A and Cluster B (p=0.46). Infants in Cluster C had faster weight gain than Cluster A (estimated difference 0.22 SD scores, p=0.01) and Cluster D had slower weight gain than Cluster A (estimated difference -0.20 SD scores, p=0.01).

When head circumference was used the dependent variable, no statistically significant differences were found between Cluster A and Cluster C (p=0.92) or Cluster D (p=0.17). However, infants in Cluster B had slower head circumference growth (estimated difference -0.29 SD scores, p=0.04).

Regression of Feeding Approach Cluster and Neurodevelopmental Outcome

Only 143 infants had basic neurodevelopmental data stored on weeks. When models were created to assess relationships between nutritional approach cluster and neurodevelopmental outcome in these infants, VIF values for cluster membership were very high and could only be reduced by removing important covariates. Therefore modelling was not attempted using these data.

4.5 Discussion

In this chapter, I have identified key nutritional, demographic and clinical care factors which influence early weight gain and head growth in preterm infants.

The SPND cohort provides a highly detailed data set, combining demographic, clinical and nutritional data on over 500 preterm infants cared for in Southampton. It provides a platform from which to explore the influences on their growth. The depth, detail and reliability of these data render the SPND a powerful tool for identifying the factors which influence the growth of these infants.

There are two major limitations of the SPND dataset. Firstly, infants were only recruited after informed written consent from his or her parents, and consent was often deferred until the second week of life. Therefore, infants who died during the first week of life, or those who were very unstable during this period and did not become stable before their death, were under-represented in the dataset. This means that the cohort was not an accurate cross-section of infants admitted to the neonatal unit. The second important limitation was that detailed data collection was not possible after infants had been discharged from Southampton. With half of infants "repatriated" to their local neonatal unit prior to discharge home, it was not possible to track all infants up to the point of discharge home. Most infants remained in Southampton for their first six weeks of life and so the analysis in this chapter focussed on this period of nutrition and growth. Whilst this is likely to encompass an important period of growth, the lack of later data limits the ability of this doctoral project to identify influences on growth beyond that early postnatal period.

It is currently impractical to gather detailed feeding and nutrition data from other neonatal units as those data are held locally and funding is not available to extract data from the local paper-based and computerised clinical systems. With the increasing adoption of computerised clinical information systems, digital transfer, extraction and organisation of these data may become easier. Current ethical approvals are limited to data held by University Hospital Southampton and extraction of data from elsewhere would require changes to this ethical approval and gaining local research and development support at each of several local neonatal units.

Some of the demographic features identified in the cohort were unexpected. At 19%, the maternal smoking rate in the cohort was substantially higher than the current rate in the Hampshire, Southampton and Isle of Wight CCG which is 9.5%.[138] However, rates of smoking have been falling during the period of data collection (it was as high as 17.2% in a different geographical area limited to the City of Southampton in 2011). The high rate of smoking in the mothers for whom records are available may have been due to clinicians being more likely to report smoking status when the mother smoked. However, there is an association between smoking and preterm birth[139, 140] meaning that that the true rate of smoking may be higher in this cohort than in the general population of pregnant women. The significant level of missing data for smoking status limited the power of this project to identify any effect of maternal smoking on postnatal growth.

The preterm infants it this study were distributed evenly across the deprivation deciles, despite the fact that preterm birth is known to be more common in more deprived populations in England (Figure 4-4).[141] Although the City of Southampton is more deprived than the national average, many surrounding areas are significantly less deprived and so the underlying population distribution of IMD is difficult to determine.[142] The even distribution of IMD decile in the database may reflect general low levels of deprivation in the region counterbalancing the expected excess of preterm births amongst the most deprived women.

The data for this project exhibited many of the problems commonly encountered in observational studies. Firstly, there is significant correlation between many of the possible predictor variables. Specifically, intakes of different nutrients were mostly significantly correlated. This made it impossible for normal generalised linear regression methods to identify the influence of each individual nutrient on growth metrics. Random forest models are less prone to collinearity problems than regression models, but they do become unstable when factors are very collinear. For each decision tree generated within the random forest, if one variable is chosen as a branching point then other variables collinear with the first are less likely to be chosen. This means that feature importance becomes unstable and it becomes difficult to assess whether a feature (e.g. intake of a specific nutrient) is an important determinant of the outcome (e.g. weight gain) or whether it is simply correlated with another feature which determines the outcome. The significant correlation between micronutrients and macronutrients meant that it was not possible to examine any influence of micronutrients, and only macronutrients were included in machine learning models. It is difficult to conceive of any statistical method which could identify the influence of highly collinear micronutrients on growth in these observational data.

Notwithstanding these problems, I was able to identify some nutrient intake features associated with faster weight gain and head circumference growth. Specifically, increasing protein intake was associated with faster weight gain and head growth, but this effect plateaued at around 3.5g/kg/day. This finding is in good agreement with guidance derived from measuring fetal accretion rates of protein alongside the obligate urinary nitrogen loss of the preterm infant.[143] Current European guidelines recommend a protein intake of 3.5-4.5g/kg/day[23] and it has previously been noted that this wide range limits the applicability of these guidelines in practice.[144] The data presented here provide evidence that aiming for protein intakes in the upper limit of this range may not increase growth to the same degree as intakes at the lower end of the range. Findings for total energy intake were less clear, although it seems that intakes above 125kcal/kg/day may not have been associated with increased weight gain or head growth. Providing 125kcal/kg/day alongside 3.5g/kg/day of protein would lead to a protein-energy ratio of 36kcal per gram of protein. This is consistent with a recommended range of 30-40kcal per gram of protein.[144] It should be noted that the models used in this doctoral project take nutrient intakes as

a mean average over a period of time. It is unclear from these data whether increased nutrient intakes after a period of deficit would support growth.

It is possible that this association between meeting nutritional requirements and good growth is confounded by the general health of the infant in question. It may be the case that infants who are generally healthy are destined to grow well. Those infants will also receive and tolerate nutritional intakes which are in line with current guidelines. Therefore, the apparent response of growth to this nutritional intake may be confounded by the general health of the infant. Where possible markers of ill-health have been included as covariables in regression models and machine learning approaches, but it is not possible to exclude the risk of residual confounding.

Taking a wider view, the feature importance values from random forest models indicate that weight gain was critically dependent on nutrient intakes but that head growth was more influenced by clinical features and was significantly impacted by common complications of prematurity, such as prolonged invasive ventilation or persistently patent ductus arteriosus. It is not possible to define the reason for this effect from the available data. It may be the case that head growth is preserved in the presence of malnutrition (in preference to promoting fat lay-down and the growth of other organs). Insofar as nutrient intakes were important, fat intake was more predictive of head growth than other macronutrients, with intakes of fat of at least 4g/kg/day associated with faster head growth.

When feeding approach clusters were used as predictors of growth, the outcome for weight gain provided some support for the current nutritional approach implemented in Southampton. Infants who rapidly established feeding with fortified breastmilk (Cluster A) gained weight at a similar pace to those who received prolonged parenteral nutrition delivered at the full intended rate (around 100ml/kg/day, Cluster B). Infants who spent a prolonged period of time receiving unfortified breastmilk alongside lower volumes of parenteral nutrition (Cluster D) gained weight more slowly. Infants transitioning rapidly onto formula feeding (Cluster C) gained weight more quickly than any other group. However, it is recognised that early formula feeding is associated with an elevated risk of necrotising enterocolitis.[145] The precise impact of formula feeding on overall mortality and morbidity remains uncertain, but it is likely that the use of formula or donated breast milk as a feed for preterm infants will be determined by these risks rather than by their impact on weight gain in clinical practice. Infants fed with a prolonged period of high-volume parenteral nutrition (Cluster B) had a slower rate of head circumference growth than those in other clusters receiving significant early enteral feeding. This may be due to a real deleterious effect of parenteral feeding, by more physiological absorption of fat during enteral feeding (considering that fat intake was an influence on head circumference growth) or due to unmeasured confounding. Infants in Cluster B were more preterm, suffered more complications and required more days of intensive care than those in other clusters. The regression models

adjusted for many markers of disease severity but it remains possible that the association between membership of this cluster and slower head growth was confounded by the severity of illness seen in these infants.

Neurodevelopmental data were sparse for this cohort and that scarcity of data made it impractical to assess the influence of nutrient intakes or feeding patterns on neurodevelopmental outcome with appropriate control of covariates. It is likely that some further neurodevelopmental data has been gathered for many of these infants, but it has not been stored on a version of BadgerNet which can be accessed from University Hospital Southampton. If further ethical approval could be obtained (and further researcher time secured), it may be possible to gather these data from other Trusts and from narrative records of infants in Southampton, but that was outside the scope of this doctoral project.

The findings from this chapter lead to number of conclusions which can be used to guide the personalised nutritional care of preterm infants. Firstly, weight gain and head growth are supported by average intakes of 3.5g/kg/day of protein and 125kcal/kg/day of total energy. Increases above these rates may not lead to further increases in growth. Secondly, weight gain is likely to be improved by an approach of maximising parenteral nutrition before briskly transitioning to fully fortified breastmilk feeds rather than prolonged periods with a mixture of parenteral nutrition and unfortified breastmilk. Thirdly, head growth is more strongly determined by clinical factors and comorbidities than nutritional factors but establishing early enteral feeding and providing a mean fat intake of at least 4g/kg/day may encourage more rapid head growth.

Chapter 5 Changes in Growth Over Time

This chapter uses weight data for very preterm infants born in England to examine changes in weight gain from 2006 to 2018. It also considers differences between the weight gain of infants in England as a whole and those born in Southampton. Finally, the influences on weight gain in the national cohort are briefly considered.

The work in this chapter was carried out by Aneurin Young, supported by his supervisors and with the expert advice of Professor Tim Cole (Professor of Medical Statistics, Institute of Child Health, University College London). The work in this chapter (excluding consideration of growth in Southampton) was published in a peer reviewed article in Archives of Disease in Childhood: Fetal and Neonatal Edition in 2022 (Appendix 4).[4] SITAR models were formed using the IRIDIS High Performance Computing Facility of the University of Southampton.

5.1 Background

As discussed in the introduction to this thesis, current international guidelines recommend that the growth of the preterm infant should match that of the fetus in utero. Professor Cole published in 2014 national data that showed that very preterm infants in England rarely achieve this (during the period 2006-2011).[25] Mean weight gain curves of the infants fell by more than two centile spaces between birth and term corrected age (corresponding to a fall in standard deviation score of around 1.5). This work used the SuperImposition by Translation And Rotation (SITAR) growth curve model to summarise the growth of groups of infants into a single summary curve.[117] This is discussed in Chapter 3 of this thesis.

Since 2011 there have been changes in neonatal practice, prioritising early nutrition including the earlier introduction of parenteral nutrition and ongoing improved nutritional provision to preterm infants. Use of infant formula has also reduced, with growing availability of donated human breastmilk. Data published from Southampton prior to my research fellowship suggested that early postnatal growth of infants born very preterm was better and brought closer to in-utero growth in a local cohort.[26] In that study, Andrews and co-workers plotted the median weight of infants born at varying gestations on the centile curves of the reanalysed UK1990 growth data, producing lines which were close to weight gain charts reflecting fetal growth. Head

circumference and length growth were less similar to fetal growth. In this chapter, SITAR was used with an updated data set to re-assess the differences between the growth of infants in Southampton and those born elsewhere in England.

Until recently, there was little data available to assess whether changes in the nutritional approach and clinical care had led to changes in the weight gain of preterm infants outside this local setting. In a recently published dataset Greenbury and co-workers examined changes in growth during the period 2008 to 2019 in England, in a process with some similarities to the methods used in this chapter.[132] The group used mean values rather than SITAR curves and paid particular attention to weight change in the first two weeks of life. These findings concerning a significantly overlapping group of infants provide a useful way to cross-check findings across different methods. The similarities and differences in the findings between Greenbury's work and mine are considered in the discussion section of this chapter.

The national English dataset used for this chapter provided some important demographic information about each infant but only contained very limited feeding information. The influences of these demographic and feeding factors on weight gain were examined by multiple linear regression.

5.2 Aims

This chapter aims to:

- Assess changes in weight gain patterns of preterm infants born in England during 2006-2011 and during 2014-2018.
- Compare growth in England as a whole with the growth of infants born in Southampton.
- Assess the impact of demographic features at birth, complications of prematurity and level of neonatal care unit on the growth of preterm infants in England.

5.3 Methods

The National Neonatal Research Database

National data were received from the National Neonatal Research Database (NNRD). The NNRD is described as:

> […] a national resource holding real-world clinical data captured in the course of care on all admissions to NHS neonatal units in England, Wales, Scotland and the Isle of Man. […] Data in the NNRD comprise the Neonatal Data Set (ISB1595), an approved NHS Information Standard and include demographic details, daily records of interventions

and treatments throughout the neonatal inpatient stay, information on diagnoses and outcomes, and follow-up health status at age two years.

National Neonatal Research Database (NNRD), Health Data Research UK[146]

The database is held by the Neonatal Data Analysis Unit (NDAU), based at Imperial College London. Data are entered by individual neonatal units during the course of routine clinical care and are then transferred to NDAU for inclusion in the NNRD.[147] In practice, this process is normally carried out by clinicians entering details onto the BadgerNet data management system (Clevermed Ltd, see section 3.2 for further information). Infants admitted to neonatal units in the UK are automatically entered onto the NNRD, with parents able to opt out of inclusion if they choose. Data from the NNRD can be requested by researchers working on neonatal topics, with a fee levied to support database management and extraction, the cost being based on the number of records requested and the number of data items for each record.

A validation study of the NNRD, led by NDAU members, identified variable reliability of the different data items.[113] Certain demographic features were also found to have high missingness (16.5% missing maternal Lower Super Output Area, used to assess deprivation, and 10.2% missing maternal ethnicity).

Study Initiation and Acquisition of Data from the NNRD

National growth data were acquired to assess changes in growth patterns over time, specifically compared to a previous cohort of infants born during 2006-2011.[25] Data from this historic cohort were limited to England and so only English data were requested for the re-assessment of growth. This exercise was formulated as a research study named Re-Growth (Reassessing the growth of infants born below 32 weeks' gestation in the UK, 2014-2018). I conceptualised the study with my supervisors and wrote the protocol (along with the ethical approval application). It was sponsored by UHS and received NHS REC authorisation (Oxford A, 20/SC/0073). The protocol for the study is given as Appendix 7. Application to NDAU for data extraction was made by me using the proforma provided by NDAU (Appendix 8). Data for all infants born before 32 weeks of gestation and cared for during the period from 2014 to 2018 in England were obtained from the NNRD.

The application for the extraction and provision of data from the NNRD was put before the NDAU board and approved. Data were extracted by NDAU and were provided to me by secure download of three comma-separated values (CSV) files: one containing details for each infant, one providing details for each episode of care experienced by each infant (multiple episodes of care are generated for an infant when the baby is transferred from one neonatal unit to another) and daily information for each infant. An anonymised code was included to cross-reference individual subjects across the three tables.

Historical English Growth Data

Weight data for preterm infants born in England during 2006-2011 were supplied by Professor Cole from the work he had completed in 2014. These data consisted of only weights, gestations and postmenstrual age values.

Southampton Data

Weight data for infants cared for in Southampton were obtained from the Southampton Preterm Nutrition Database, collated as described in Chapter 3. A detailed description of this cohort of infants is given in Chapter 4 of this thesis.

Data Analysis

Weight data were filtered using the algorithm described in chapter 3. In summary, values above 60,000g or below 60g were discarded, likely transposition errors caused by incorrect decimal places were corrected, and values which were implausible for the postmenstrual age and were not likely transposition errors were discarded. Upper age limits were set so that at least 10% of infants remained in the dataset at that time, to maintain adequate data within each group for modelling. This approach is consistent with that used by Cole in his 2014 paper.[25]

SITAR

Weight gain was summarised using the SITAR growth curve model with the *sitar* package in R. The SITAR method is described in more detail in Chapter 3 of this thesis. Cleaned data from the 2006-2011 cohort were also re-analysed to compare the results for the two eras. The era of the study (i.e. 2006-2011 versus 2014-2018) was included as a fixed effect in combined models using all infants, to test its effect on size, timing and intensity.

SITAR curves were fitted for male and female infants combined as their growth patterns are known to be similar.[25, 148] Separate growth curves were formed for each week of gestational age at birth, with 22 and 23 weeks of gestation combined due to small numbers. This matched the approach of the earlier cohort. Weight measurements more than four residual standard deviations from these mean growth curves were then excluded. For each individual growth curve, weight measurements were examined as triplets, velocity of weight gain was calculated for each measurement and those with an implausible velocity were excluded. Final SITAR models were formed from the resultant data.

For comparison between English and Southampton data, infants were selected from the Southampton Preterm Nutrition Database (SPND) who were born from the beginning of 2014 to

the end of 2021, representing the current nutritional approach in Southampton. Infants born before 24 weeks were excluded as there were too few for analysis in the Southampton weeks. Infants born on or after 30 weeks of gestation were also excluded, as data gathering for these infants started during the period in question. SITAR models were formed separately for the two groups, split by gestational week, for visual comparison and a combined model using the place of birth as a fixed effect was used for statistical comparisons. Reanalysed UK1990 growth curves[27, 115] were illustrated using reference data within the *childsds* package for R, with the curves for males and females averaged.

Weight z-scores were calculated for measurements taken around two weeks postnatal age and for subsequent alternate weeks. Differences between English and Southampton changes in weight z-score were analysed using two-sided student's t-tests, taking p-values below 0.05 as significant.

Regression Analysis

Weight gain was calculated using weights at birth, around 36^{+0} weeks gestation (between 35^{+0} and 36^{+6} weeks). Weight SD scores were calculated using the reanalysed UK1990 data, which excluded infants born prior to 23 weeks of gestation.[149] Weight gain was defined as the change in SD score from birth to 36 weeks.

Multiple linear regression models were used to assess demographic and clinical factors which had been identified by clinical researchers as likely to influence growth. The Bayesian information criterion (BIC) was used to assess model efficiency and to select the most appropriate variables for inclusion. For each model variable, the t-statistic was calculated. The t-statistic is calculated by dividing the effect size estimate by the standard error of the estimator. Values higher than 2 or lower than -2 are taken to indicated statistically significant positive or negative associations respectively.[150]

5.4 Results

5.4.1 The English National Cohort

Data were provided for 37,700 infants. Demographic details are set out in Table 5-1.

Table 5-1 Demographic details of infants included in the national dataset received from the NNRD.

Gestation (weeks + days), median (range)		29^{+3} (21^{+3} to 31^{+6})
Birthweight (g), mean ±SD		1200 ±386
Male, number (%)		20535 (54%)
SGA*, number (%)		5629 (15%)
Born in NICU Centre**, number (%)		20786 (55%)
Multiplicity of Pregnancy, number (%)	Singleton	27763 (74%)
	Twin	9079 (24%)
	Triplet	790 (2%)

* Small for gestational age (birthweight below 10^{th} centile).

** Neonatal Intensive Care Unit as defined by the British Association of Perinatal Medicine (BAPM)[151]

Figure 5-1-A illustrates the number of infants born at each completed week of gestation, demonstrating the expected increase as gestation increases. The number of infants born at 23 weeks of gestation is lower than would be expected, likely due to an excess of deaths in the delivery room in this group, either due to planned non-intervention or due to unsuccessful resuscitation. The NNRD includes only infants admitted alive to a neonatal unit. Like the data from Southampton, the distribution of birthweight z-score is left-skewed (Figure 5-1-C) and the rate of small for gestational age (birthweight z-score less than 10^{th} centile) was 15%, indicating an excess of infants with intrauterine growth restriction (IUGR). At first glance, this seems counterintuitive, in that a cohort which includes all infants should have 10% of individuals classified as SGA and the peak of the histogram should be at zero. This finding has been previously recognised by a different research group in an overlapping dataset.[132] The reference data used to calculate these z-scores are derived from infants born in the UK from 1983 to 1993. It is hypothesised that increased survival of growth restricted fetuses to delivery and admission to the neonatal unit has led to a reduction in mean birthweight of preterm infants. The distribution of head circumferences is less skewed (Figure 5-1-D) but the median head circumference z-score at birth is similarly below zero (Figure 5-1-E).

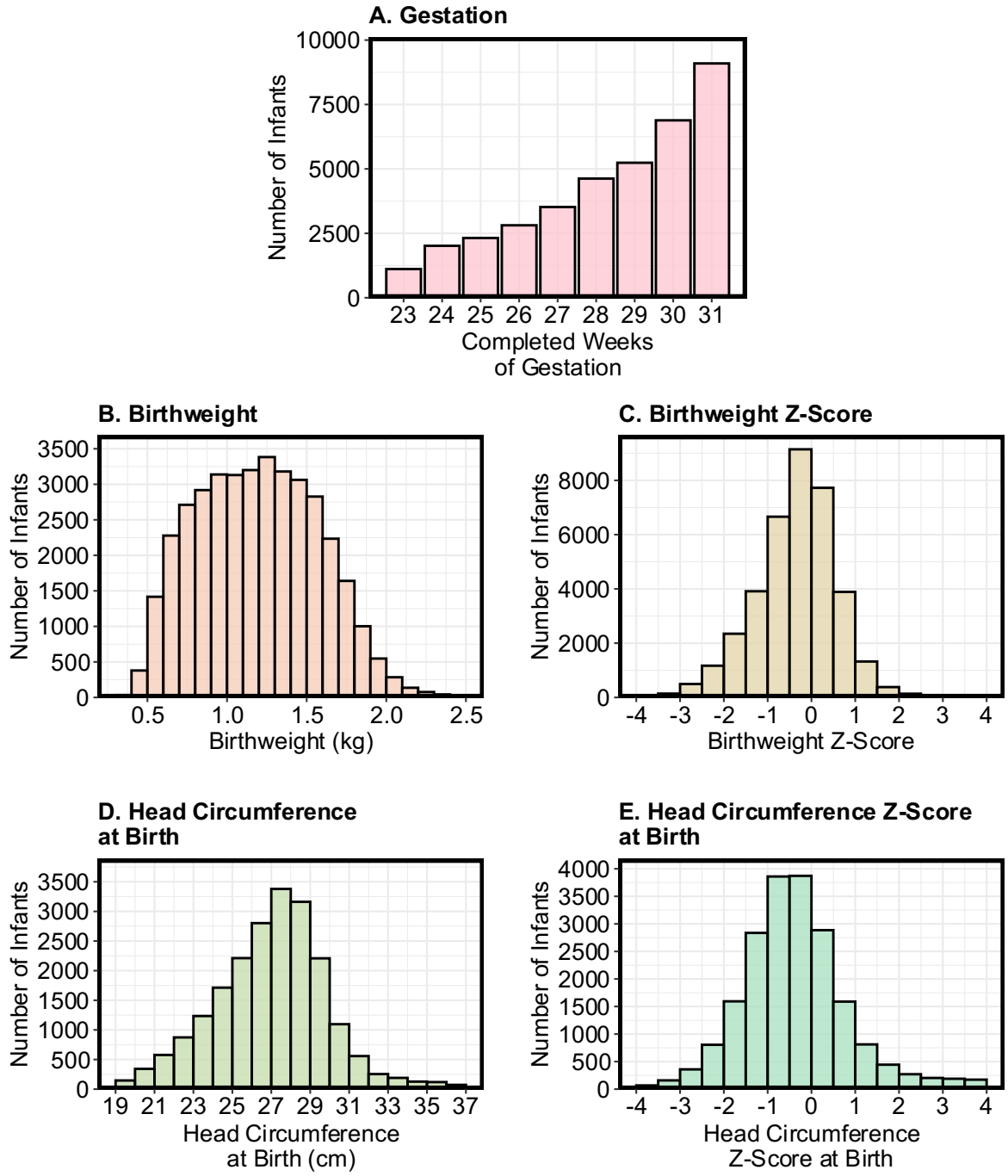


Figure 5-1. Demographic features of infants included in the national dataset received from the NNRD. A. Number of infants born at each gestation; B. Birthweight distribution; C. Birthweight z-score distribution; D. Head circumference at birth distribution; E. Head circumference z-score distribution.

Variables

Table 5-2 sets out the data items provided by NDAU, split into the three linked tables described above.

Table 5-2. Data fields in the national dataset provided from the National Neonatal Research Database

VARIABLE	DECRIPTION
table: baby level	
Anon_ID	Anonymised primary key for infant
BirthMonth	Birth month
BirthYear	Birth year
POBNDAUCode	Anonymised code of the place of birth
POBLevel	Level of neonatal unit at place where infant was born
Birthlength	Birth length (cm)
BirthHeweekscumference	Birth headweeksumference (cm)
GestationWeekweekspleted weeks gestation at birth	
GestationDays	Days in addition to completed weeks at birth
Gender	Sex
LSOA	Lower super output area of mother's address
FetusNumber	Number of fetuses in pregnancy
Apgar_1min	Apgar scores at 1, 5 and 10 minutes
Apgar_5min	
Apgar_10min	
RecTime	When two-year follow-up occurred (in relation to time of birth and actual month)
RecTimeMonth	
RecTimeYear	
growth_weight	Weight at two-year follow-up (kg)
growth_length	Length at two-year follow-up (cm)
growth_head_circ	Head circumference at two-year follow-up (cm)
dev_normal_less_3month_delay	Whether development was normal or mildly, moderately or severely delayed at two-year follow-up
dev_mild_delay	
dev_moderate_delay	
dev_severe_delay	
table: episode of care level	
Anon_ID	Anonymised primary key for infant
EpisodeNumberBaby	The number of the episode of care for the baby
ProviderNDAUCode	Pseudonymised code for place where care was provided
ProviderLevel	Level of neonatal unit
AdmitTimeAnon	Age of infant at time of this admission (mins)
DischTimeAnon	Age of infant at time of discharge from this episode (mins)
DischargeDestination	Place discharged to (e.g. home, died, another hospital)

VARIABLE	DECRIPTION
DateofDeath	Age of infant at time of death (mins)
table: daily data	
Anon_ID	Anonymised primary key for infant
DayDateAnon	Minutes from midnight on day of birth
DayProviderNDAUCode	Pseudonymised code for place where care was provided
DayWorkingWeight	Weight on this day (g)
DayHeadCirc	Head circumference this day (cm)
DayLength	Length this day (cm)
ParenteralNutrition	Whether had parenteral nutrition this day (binary)
GlucoseElectrolyte	Whether had IV fluids this day (binary)
DayEnteralFeeds	Types of milk given on this day
FormulaName	Types of formula given on this day
VolumeMilk	Volume of milk this day (ml)
MajorSurgeryToday	Whether had major surgery this day
SurgeryVPShunt	Whether ventriculoperitoneal shunt sited this day
ROPSurgery	Whether had surgery for retinopathy of prematurity this day
DiagnosesDay	List of diagnoses recorded on this day
RespiratorySupport	Whether invasive, non-invasive or no ventilation given
AddedO2	Whether supplemental oxygen delivered (and by which route) this day
NonInvasiveRespiratorySupport	Type of non-invasie ventilation this day
VentilationMode	Whether conventional or oscillatory ventilation given this day
PulmonaryVasodilator	Whether received pulmonary vasodilating drugs this day
Inotropesgiven	Whether received inotropic drugs this day
TreatmentforPDA	Whether and what treatment given for patent ductus arteriosus this day
NECTreatment	Whether necrotising enterocolitis treatment given this day (and whether surgical or medical)
LinesIn	Types of indwelling vascular lines in situ this day

Missingness and Data Accuracy

The NNRD is known to have variable coverage of different data types. Some data are required to enter and infant onto BadgerNet or can be inferred from the site and timing of recording. These data items have low missingness. For example, data are complete for birth year and birth month. Similarly, only 9 out of the 37,700 infants had a missing birthweight value. In contrast, head circumference at birth was missing for 43% of infants and length at birth was missing in 94% of cases. Neurodevelopmental follow-up data were available for only 37% of infants (although approximately 10% would not have reached two years corrected age at the time of data extraction). The missingness in follow-up data was non-randomly distributed: infants with follow-up had a mean gestational age at birth of 29.6 weeks and those without follow-up had a mean gestation at birth of 28.4 weeks, after exclusion of infants who died before discharge (t-test $p<0.001$).

Length of Stay

Total length of admission to any neonatal unit was significantly influenced by gestation at birth (Figure 5-2-A). Generally, the most preterm infants spent longer receiving neonatal care. When all infants were included, the total length of admission for infants born after 22 and 23 completed weeks of gestation were very low due to the much higher mortality in those groups (see Table 5-3).

Table 5-3. Mortality and discharge destination of infants in the national dataset

Discharge Destination, number (%)	Home	32,664 (86.6%)
	Ongoing inpatient care	1,889 (5.0%)
	Died	3139 (8.3%)
	Unknown	8 (<0.01%)
Mortality by Gestation, number (%)	22	41 (70.7%)
	23	601 (53.9%)
	24	677 (33.6%)
	25	439 (18.9%)
	26	374 (13.3%)
	27	264 (7.5%)
	28	280 (6.1%)
	29	158 (3%)
	30	149 (2.2%)
	31	157 (1.7%)
Age at discharge, days (IQR)		50 (34 to 77)

When analysis was restricted to infants who survived to discharge from the neonatal unit, the pattern was as expected (Figure 5-2-B). The NNRD aims to capture information for all infants who are admitted to a neonatal unit. This means that infants who die in the delivery room were not

included in these data. However, infants who die shortly after admission are included (in contrast to the Southampton data), so the median age at death was much lower in this dataset (Figure 5-2-B).

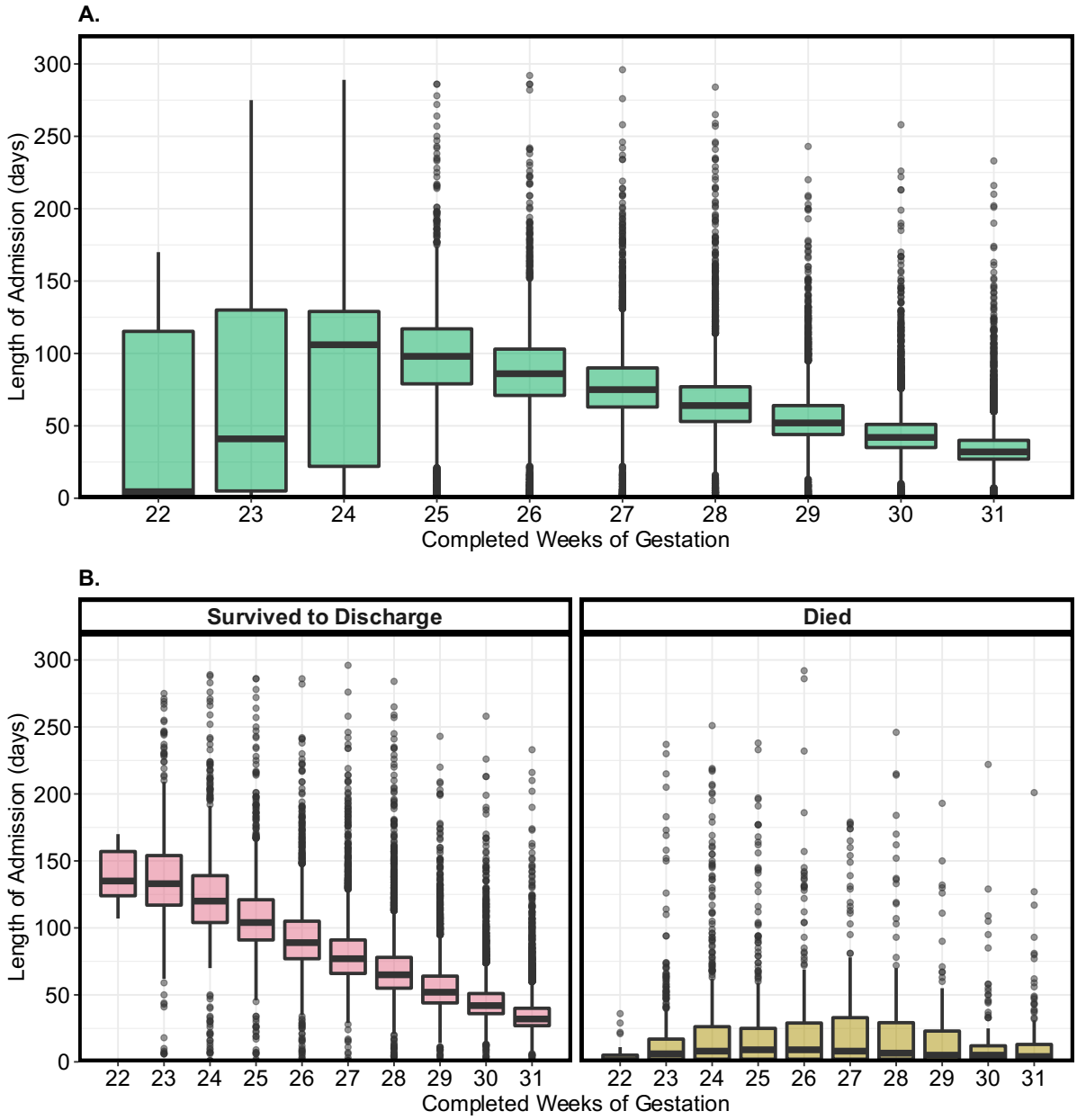


Figure 5-2. Boxplot of the total length of admission for infants in the national dataset. A. all infants grouped by gestation; B. Infants separated by survival status. Line indicates median, box encompasses interquartile range, whiskers indicate limits of values within 1.5x IQR and dots are outliers. A handful of infants with a length of admission over 300 days have been omitted from plots for clarity (but have been included in analysis).

Measurement Data

The NNRD stores weight measurements differently from how head circumference and length measurements are stored. Whereas length and head circumference values are recorded only on the days those measurements are taken, a weight value is recorded every day. This means that

values from actual weighing episodes are repeated every day until the next time an infant is weighed. The storage of weight data in this way makes it impossible to know whether an infant has been weighed on a specific day; if the weight value is different to the previous day's then it can be inferred that the infant has been weighed, but an infant whose weight has not changed may or may not have been weighed. Therefore, an algorithm can be used to detect whether there is a "new" weight value on that day. When this algorithm was applied, there were 747,241 new weight values in the database (33% of the total number of days). This value implies that infants were weighed every three days on average. Current practice for weighing frequency is not well-described in the UK, although publicly available guidelines from the East of England and the West of Scotland both recommend weighing infants two or three times each week.[152, 153] The actual number of weight values derived in this way may be artificially expanded by small deviations in the daily weight value recorded (for example, using different rounding), in the absence of actually weighing the infant on the day the weight was changed. The median number of weight measurements for each infant was 17 (IQR 11 to 27). Head circumference and length measurements were much sparser, with the median number of head circumference measurements of 1. 80% of infants had no recorded length measurement. The scarcity of these data rendered the NNRD data unsuitable for assessing linear growth of infants.

Some of the weight values present in the database were clearly spurious, with values as low as 1g and as high as 33.4 metric tonnes. Applying the filtering approach described in section 3.3.2.3, 3424 weight values were implausible. Of these, 2484 were likely factor-of-ten errors which could be corrected, leaving 940 weight values which were discarded (0.001% of the total weight values).

Feeding Data

The NNRD does not hold detailed nutrient intake data. As set out in Table 5-2 above, it does capture whether parenteral nutrition (PN) was given and the type of enteral feed given on each day. From this, some nutrition data could be derived. After excluding infants who died before the end of their first week of life, 65% of infants received PN prior to the end of their first full day of life and 81% received PN at some point in the first six weeks of life. PN use was more common in infants born at earlier gestations (Figure 5-3). Infants born at 22 weeks of gestation were less likely to receive early PN than other extremely preterm infants.

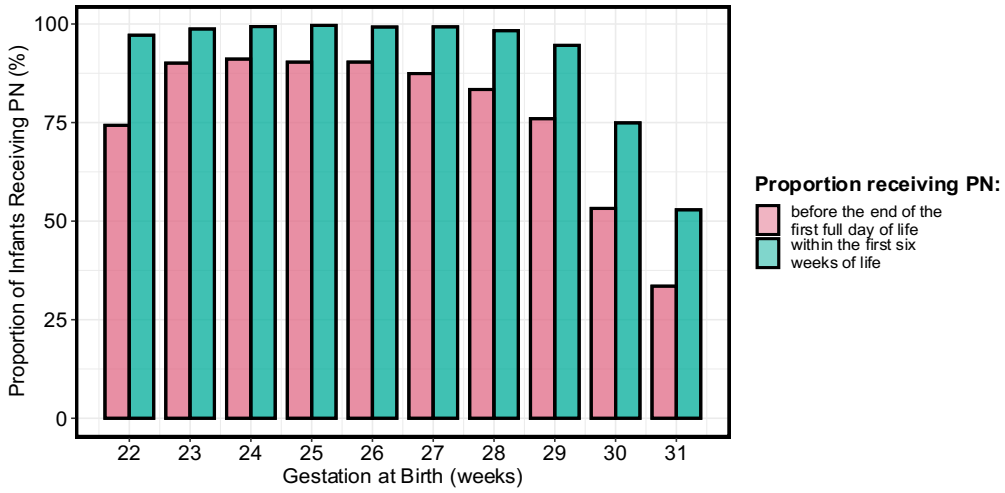


Figure 5-3. The proportion of infants receiving parenteral nutrition before the end of the first full day of life and within the first six weeks of life, split by gestation (after excluding infants who died before the end of their first full day of life).

Some inferences can also be made about enteral feeding. For example, 93% of infants received some maternal breastmilk during their neonatal stay, 77% received some infant formula, 47% received breastmilk fortifier and 31% received donated breastmilk (all after excluding infants who died before the end of their first full day of life).

Clinical Comorbidity Data

The clinical information provided in the NNRD dataset was less detailed than in the Southampton database. Some information on respiratory support was available (indicating that 67% of the cohort received invasive ventilation at some point during their stay). The data also indicated that 15% of the cohort received medical management for necrotising enterocolitis (NEC) and that 2.6% had surgical treatment for NEC (although it should be noted that validation studies on the NNRD have indicated that these data are not completely reliable.[154]

5.4.2 SITAR Growth Curve Analysis

Figure 5-4-A shows SITAR curves for infants born 2014-2018 grouped by completed weeks of gestation and superimposed on UK-WHO weight centiles with the sexes averaged. Figure 5-4-B shows the same data with weight on a logarithmic scale, so that the slope of each curve represents relative weight gain (*i.e*. g/kg body weight/day). Each curve crosses centiles downward, indicating that weight gain did not keep pace with expected weight gain for the equivalent fetus in utero

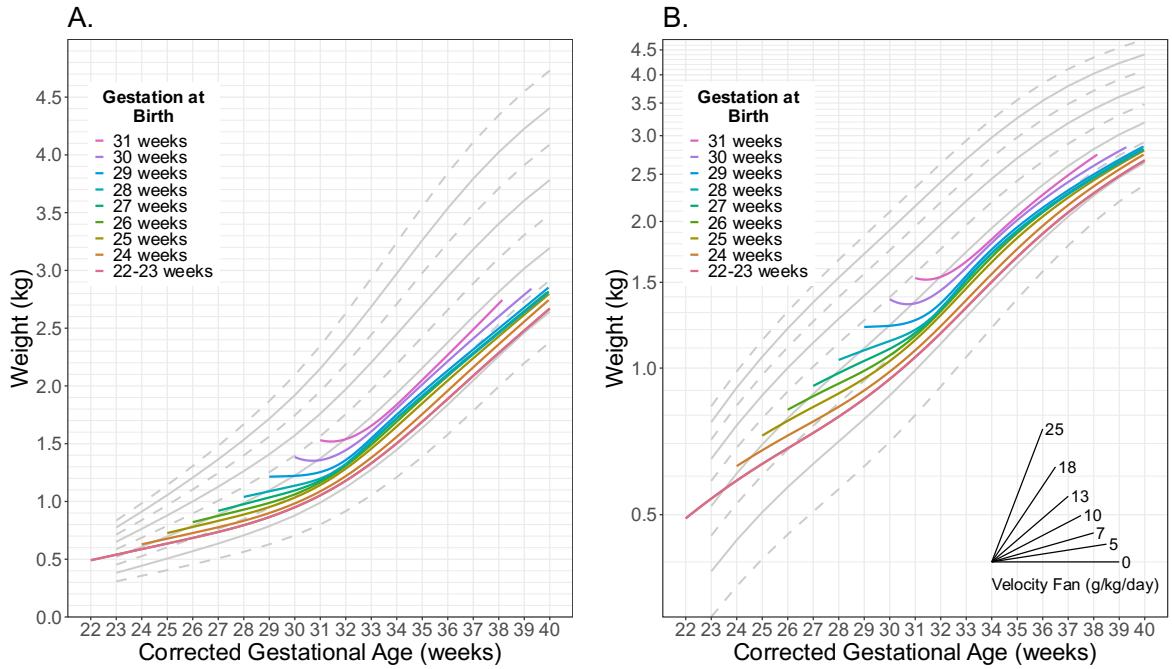


Figure 5-3**. SITAR curves for weight by completed weeks of gestational age for the English 2014-18 cohort superimposed on reanalysed *UK1990* centile lines (sexes averaged) plotted with A. an absolute weight scale; and B. a logarithmic scale of weight with a velocity fan demonstrating proportional weight gain in g/kg/day**

Comparison with Data from 2006-2011

Figure 5-5 compares SITAR curves from English infants during the two eras, 2006-2011 and 2014-2018. For infants born from 24 to 28 weeks of gestation, weight at birth was similar. Birthweight was lower in the more recent era for infants born at 22-23 weeks, likely reflecting a trend towards initiating intensive care in the smallest and most premature infants more often.

For each gestation group before 28 weeks, SITAR curves for the two eras diverged slightly, with infants in the more recent era exhibiting faster weight gain signified by a steeper SITAR curve. This effect was not so apparent for infants born after 28 weeks PMA. For the more mature infants, the SITAR growth curve for the 2006-2011 infants demonstrated early weight loss, followed by increasing weight. This pattern was much less prominent for the 2014-2018 cohort, with very little early weight loss seen in these infants. Instead, there was a period of static weight followed by acceleration in growth velocity.

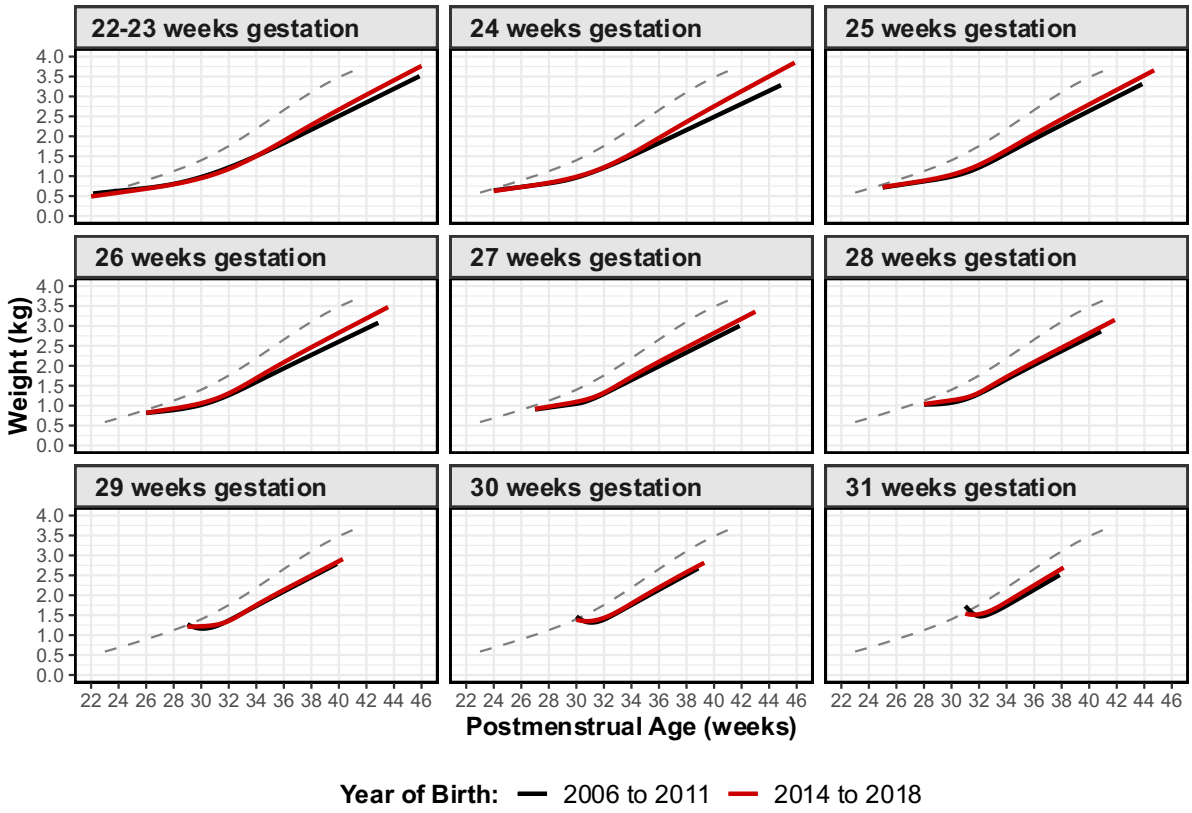


Figure 5-4. SITAR curves by gestational age, comparing infants born during the era of 2006-2011 and the era of 2014-2018. Grey dashed line is the reanalysed *UK1990* growth chart 50th centile line for cweeksison.[115]

Table 5-4 shows the effect of being in the 2014-2018 cohort on the size (in grams), timing (in weeks) and intensity (in percentage change) fixed effects of the SITAR models for each gestation group, with the associated t-statistics where a value exceeding ±2 is statistically significant. The size value was generally higher (although not always significantly so). This indicates that mean weight was consistently greater during the 2014-2018 cohort than the 2006-2011 cohort (see Figure 5-5). The timing parameter reflects the age at peak weight gain, and the negative effect for the later cohort indicates peak weight gain occurring earlier (although not significantly so for most groups). Intensity was significantly greater for every gestation group born before 30 weeks of gestation, indicating more rapid weight gain in the later cohort.

Table 5-4. The effect of birth in the 2014-2018 era (compared to the 2006-2011 era) on mean size, timing and intensity of SITAR models, grouped by gestation.

Gestation group (weeks)	Effect of later era on size, grams (95% CI)	t-statistic	Effect of later era on timing, weeks (95% CI)	t-statistic	Effect of later era on intensity, % (95% CI)	t-statistic
22-23	20 (-35 to 74)	0.7	-0.4 (-1 to 0.3)	-1.2	12 (5 to 19)	3.3
24	62 (31 to 93)	3.9	-0.2 (-0.6 to 0.1)	-1.3	15 (11 to 19)	7
25	11 (-17 to 40)	0.8	-0.4 (-0.7 to -0.2)	-3.2	5 (1 to 8)	2.5
26	61 (38 to 84)	5.1	-0.1 (-0.3 to 0.2)	-0.5	13 (10 to 16)	8
27	53 (33 to 73)	5.1	0.0 (-0.2 to 0.1)	-0.3	9 (7 to 12)	6.7
28	50 (32 to 68)	5.3	0.1 (-0.1 to 0.2)	1.1	6 (4 to 9)	4.8
29	24 (6 to 42)	2.6	0.1 (0 to 0.1)	2.1	5 (3 to 7)	4.3
30	11 (-7 to 29)	1.2	-0.3 (-0.6 to -0.1)	-2.8	-2 (-3 to 0)	-2.2

5.4.3 Influences on Weight Gain in the 2014-2018 National Cohort

For the analysis of influences on weight gain in the 2014-2018 cohort, infants were included who had weight measured at birth and 36 weeks PMA, excluding those born out of hospital. Change in weight SD score between birth to 36 weeks PMA was calculated. Demographic details of included infants are given in Table 5-5. The mean change in weight SD score was -0.95, indicating that weight gain failed to keep pace with fetal growth.

Table 5-5. Summary details of included infants

Number of infants		29,312
Gestational age (weeks+days) (median (IQR))		29^{+3} (27^{+3} to 30^{+6})
Birthweight (g) (median (IQR))		1170 (901 to 1445)
Birthweight z-score (mean (SD))		-0.42 (0.91)
Small for gestational age (% (n))		17% (4931)
Male (% (n))		54% (15843)
Singleton (% (n))		72% (21198)
Neonatal Unit Level at Birth Centre	Level 1 (% (n))	6% (1913)
	Level 2 (% (n))	38% (11009)
	Level 3 (% (n))	56% (16390)
Change in Weight SD score from birth to 36 weeks PMA (mean (SD))		-0.95 (0.66)

Table 5-6 summarises the multiple linear regression of weight SD score change in the 2014-2018 cohort on selected perinatal factors. The effect size column shows the effect of a one unit change in each factor on the change in weight SD score between birth and 36 weeks PMA, and the t-

statistic indicates its statistical significance—all factors are highly significant. Gestational age was positively associated with weight SD score change, and birthweight negatively, indicating that weight gain was generally greater for infants born at later gestations and for smaller infants (after adjustment for gestation). There was a positive interaction between birthweight and gestational age at birth. This indicates that the effect of birthweight (i.e. faster growth for infants with lower birthweight) is greater for more preterm infants.

Infants with higher Apgar scores exhibited faster weight gain and those who required ventilation during the first day of life had slower weight gain, suggesting that infants with a more severe illness phenotype have slower weight gain. Birth in a level 3 unit was associated with an increase in weight SD change of 0.07 (95%CI 0.06 to 0.09), i.e. an infant born in a level 3 unit could be expected to demonstrate weight gain deficit which was less severe than an infant born in a level 1 or level 2 unit by a magnitude of 0.07 SD scores. Initiation of parenteral nutrition during the first day of life was also associated with a positive effect on weight gain (0.08SD, 95%CI 0.06-0.09). Birth to a mother residing in less deprived neighbourhoods, measured by index of multiple deprivation, was associated with slower weight gain.

Table 5-6. Results of multiple linear regression examining the influences of factors around the time of birth on change in weight SD score from birth to 36 weeks postmenstrual age

	Effect size – SD score change (95% CI)	**t-*statistic***
Gestational age (weeks)	0.13 (0.12 to 0.14)	24
Birthweight (kg)	-2.7 (-3.0 to -2.4)	-17
Interaction term for gestational age (weeks) and birthweight (kg)	0.05 (0.04 to 0.06)	10
Sex (female)	-0.04 (-0.05 to -0.02)	-5
IMD decile*	-0.01 (-0.013 to -0.009)	-9
Born in level 3 unit (vs born in level 1 or level 2 unit)	0.07 (0.06 to 0.09)	10
Apgar score at 5 minutes	0.009 (0.005 to 0.012)	4
Ventilated on day of birth (vs not ventilated on day of birth)	-0.05 (-0.07 to -0.04)	-7
Parenteral nutrition during first day of life (vs no parenteral nutrition on first day of life)	0.08 (0.06 to 0.09)	10
Birth year (effect of birth one year later)	0.03 (0.02 to 0.03)	11

*Index of multiple deprivation decile for mother's postcode (higher is less deprived)

5.4.4 Comparison Between England and Southampton Data

The Southampton cohort was comprised of 294 infants, with the number of infants and measurements in each group given in Table 5-7.

Table 5-7. Number of infants and number of weight measurements split by gestational group in data from Southampton 2014-2021

Gestation	Number of Infants	Number of Weight Measurements
24 weeks	38	1530
25 weeks	49	1831
26 weeks	55	1606
27 weeks	49	1261
28 weeks	51	1489
29 weeks	52	1403

Visual comparison of SITAR curves suggested that weight gain was generally similar between the English national cohort and the local Southampton cohort (Figure 5-6). Early weight gain looked slightly faster for extremely preterm infants born before 28 weeks of gestation. However, this faster early growth was followed by slower weight gain later in the neonatal period. Possible explanations for this effect are examined in the discussion section of this chapter.

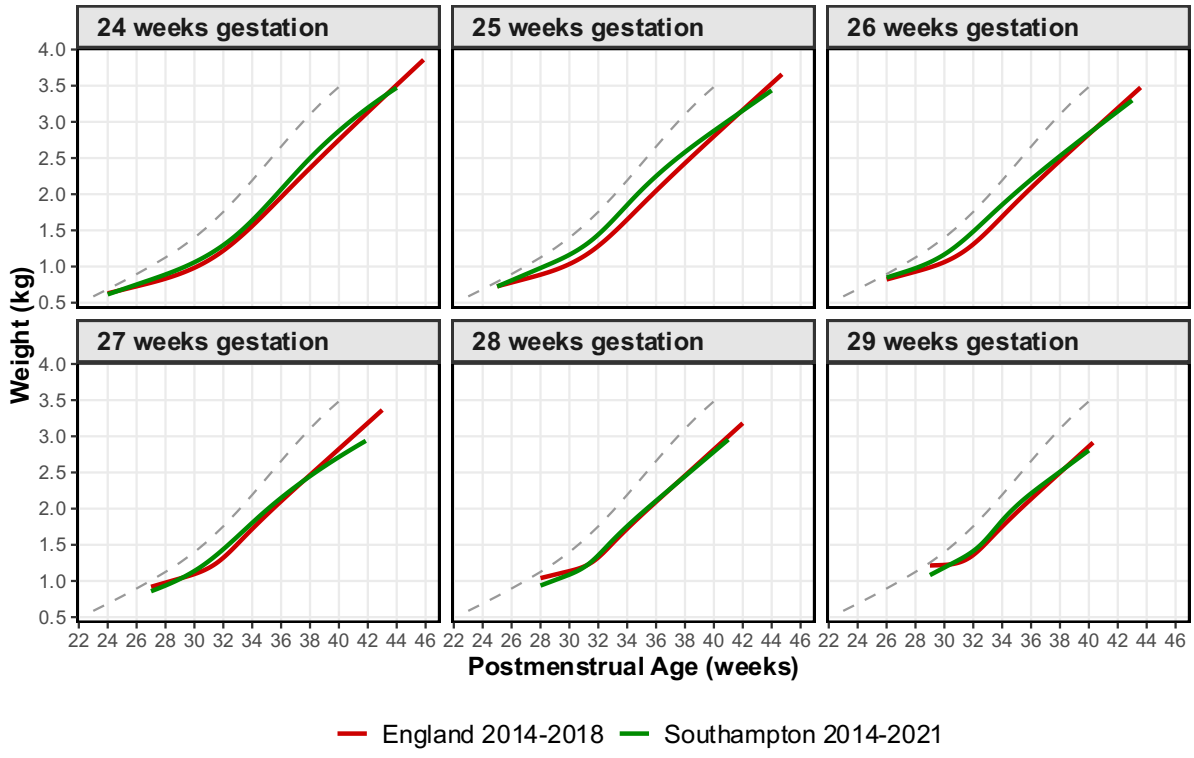


Figure 5-5. Comparison of SITAR curves for English national weight gain and weight gain in Southampton.

Table 5-8 shows the fixed effect of being born in Southampton (compared to England as a whole) on the SITAR parameters.

Table 5-8 The effect of birth in Southampton (2014-2021) compared to English 2014-2018 cohort on mean size, time and intensity of SITAR models, grouped by gestation.

Group	Effect of care in Southampton on size/α grams (95% CI)	t-value	Effect of care in Southampton on timing/β weeks (95% CI)	t-value	Effect of care in Southampton on velocity/γ % (95% CI)	t-value
24w	24 (-50 to 98)	0.6	-0.4 (-1.2 to 0.3)	-1.2	0.03 (-0.07 to 0.12)	0.6
25w	-48 (-117 to 22)	-1.3	-1.3 (-1.9 to -0.7)	-4.2	-0.03 (-0.11 to 0.06)	-0.6
26w	-29 (-93 to 35)	-0.9	-1.0 (-1.6 to -0.5)	-3.7	-0.04 (-0.13 to 0.06)	-0.8
27w	-141 (-209 to -73)	-4.1	-1.5 (-2.0 to -0.9)	-5.3	-0.10 (-0.20 to -0.01)	-2.1
28w	-99 (-158 to -40)	-3.3	-0.7 (-1.1 to -0.3)	-3.2	0.00 (-0.08 to 0.07)	-0.1
29w	-140 (-206 to -76)	-4.2	-1.9 (-2.4 to -1.5)	-8.6	0.00 (-0.05 to 0.06)	0.1

These data demonstrate the small differences between the two cohorts in question. At the more mature gestations, the overall size of the infants (generally reflecting birthweight) was lower for infants born in Southampton. The tempo was earlier in the Southampton group (with the exception of infants born at 24weekss of gestation), reflecting an earlier acceleration in weight gain. The velocity was generally not statistically significantly different between the two groups (being slightly slower overall for Southampton infants born at 27 weeks of gestation).

Data were reanalysed, limiting national data to measurements made during care in a Level 3 neonatal unit (to make these more comparable to Southampton data). This made very little difference to the overall curves, although weight gain was slightly slower in the national group so that the weight gain curves converged at the earlier gestations (Figure 5-7).

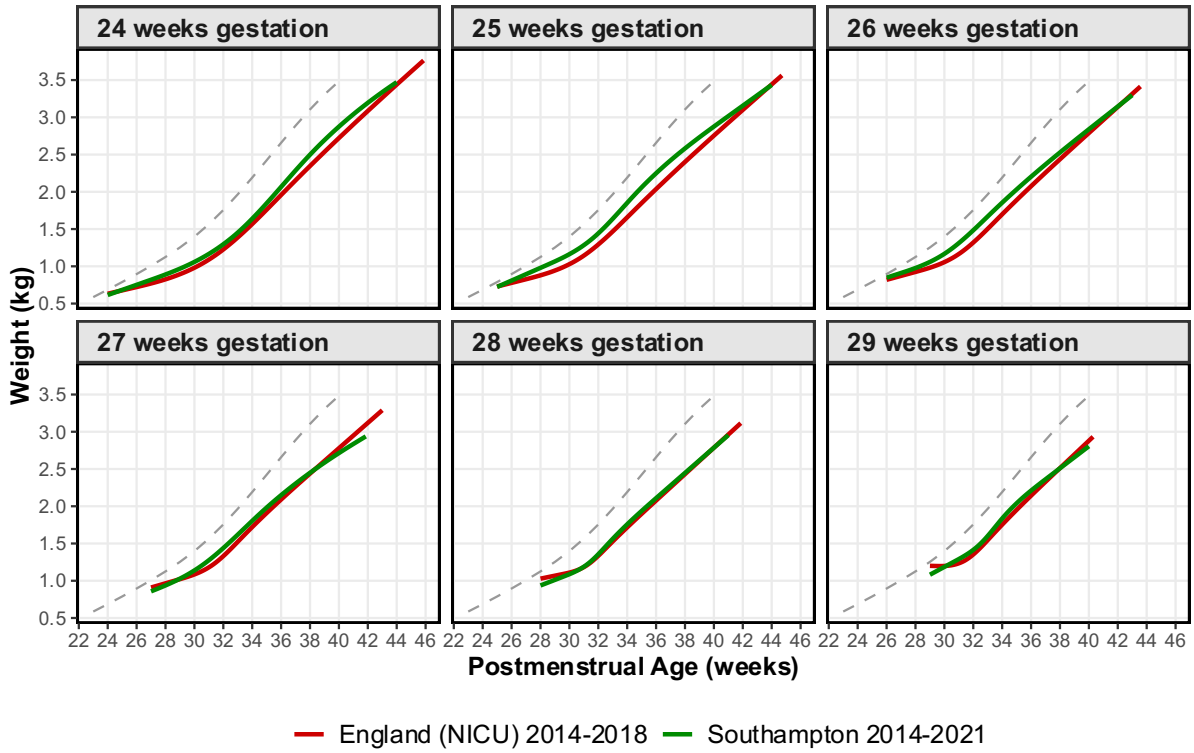


Figure 5-6. Comparison of SITAR curves for English national weight gain (limited to measurements taken in Level 3 neonatal units) and weight gain in Southampton.

Change in weight z-scores were calculated at two-week intervals of postnatal age were calculated. Whilst both Southampton and English infants dropped their weight z-score, that deficit was consistently lower for infants from Southampton. These differences are illustrated in Figure 5-8 and all differences were statistically significant with two-sided student's t-test p-values below 0.05. The absolute change in z-score is difficult to interpret at the higher postnatal ages Is there is a discontinuity in the reanalysed UK1990 data at 40 weeks of postmenstrual age, but comparisons between the groups remain valid.

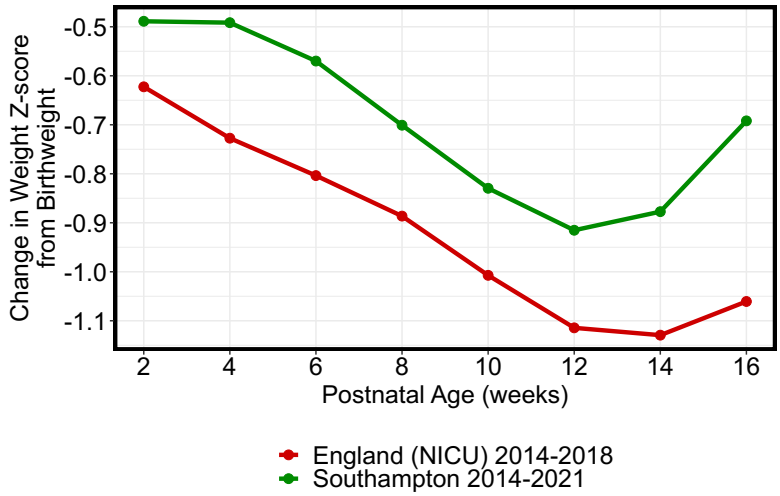


Figure 5-7. Change in weight z-score since birth for infants born in England (limited to measurements in Level 3 neonatal units) and in Southampton.

5.5 Discussion

Weight Gain Patterns in England

This chapter describes the weight gain pattern of very preterm infants in England born during 2014-2018. In comparison to infants born during 2006-2011, the most preterm infants (born before 30 weeks of gestation) exhibited faster weight gain. Early weight loss, which was apparent for later gestations in 2006-2011, was much less pronounced in 2014-2018, being replaced by a short period of static weight.

During the period of time from 2006 to 2018, there have been myriad changes in neonatal care in England. Quality improvement efforts have focused on antenatal transfer to a location with an appropriate level of neonatal care,[155] antenatal steroid provision and prevention of hypothermia on admission. Guidelines for the nutritional care of very preterm infants have emphasised early use of parenteral nutrition and fortification of breastmilk.[23, 54] Despite their increased growth rate, infants in the English 2014-2018 cohort still do not keep pace with fetal growth as represented by current charts based on birthweight. This is reflected by downward deviation of SITAR lines across centiles (Figure 5-4) and by the mean change in SD score being close to -1 (*i.e.* a loss of one standard deviation during neonatal admission).

The findings presented here are in agreement with the recent paper by Greenbury and co-workers.[132] That group found that early postnatal weight loss had decreased over time (as seen in the more mature groups in this analysis) and that subsequent weight gain was faster. Similarly to the data presented here, they identified that weight gain continues to fall below the rate needed to keep pace with the equivalent fetus in utero.

The appropriate pattern of weight change during the first two weeks of life remains disputed. In term-born infants, there is a well-recognised pattern of early weight loss, caused by the term contraction of extracellular spaces (TeCES).[34] The extent to which there is a preterm correlate to this effect is unclear, and it is likely to be influenced by the gestation at birth. Direct measurement of fluid compartment volumes is difficult in preterm infants.[156] In the absence of body composition measurements, it is hard to theorise about the causes of the change in early weight gain patterns in infants born at 30 to 31 weeks of gestation. Changes in fluid management may have reduced the contraction of extracellular space. Alternatively, early weight loss may not be physiological in these infants, and improvements in nutrition (or other changes in care) may have prevented malnutrition and pathological weight loss.

Influences on Growth in English National Cohort

Multiple linear regression demonstrated that numerous demographic and care factors were associated with changes in growth during the period from birth to 36 weeks PMA (Table 5-6). Change in weight SD score was positively associated with later gestation, corroborating the reduced deviation from birthweight centile seen in SITAR models (Figure 5-4). There was also a positive association with year of birth, confirming that weight gain has increased over time. Birth in a hospital offering level 3 neonatal care was associated with faster weight gain, suggesting an impact of antenatal transfer to a specialised setting.[155] Early parenteral nutrition was also associated with a reduced weight gain deficit. Infants born to mothers residing in areas with less deprivation demonstrated slower weight gain.

Weight Gain in Southampton

SITAR analysis identified subtle differences between weight gain patterns in Southampton compared to those in the whole of England. For infants born at 27-29 weeks of gestation, the overall size of infants was smaller in Southampton, likely reflecting lower birthweight. This small difference was similar when national data were limited to weight measurements taken in Level 3 neonatal units. Early weight gain in Southampton exceeded that in other Level 3 units, reflected by SITAR lines and by the SITAR tempo value being higher. However, SITAR lines subsequently converged. This may be due to a true later weight gain deficit in Southampton due to nutritional care or other care factors. However, there are other possible causes. Firstly, there are relatively few infants in the Southampton dataset (approximately 50 in each gestation group). As infants were transferred to local neonatal units, the number of infants included dropped further. As healthier infants were likely to be transferred earlier, those remaining in Southampton were disproportionately drawn from those with the most significant disease. Whereas the SITAR process identified simple curves with just one inflection point for English data, the Southampton curves were more complex. This suggests that the Southampton curves may have been overfitted to the underlying data. Secondly, Southampton provides surgical care to infants, whereas some other Level 3 neonatal units do not provide such care. This may have caused systematic differences between the types of infants included in the datasets, with infants undergoing surgery likely to have significant challenges to growth. The national dataset was anonymised at the level of the neonatal unit and did not capture whether an infant was cared for in a surgical centre, making it impossible to adjust for this effect.

Analysing changes in weight z-score by postnatal age (Figure 5-8) suggested that weight gain deficit was lower for infants born in Southampton than those born in England as a whole. These data suggest that the apparent slower weight gain in Southampton infants at later stages seen in the SITAR curves was an artefact of the small number of infants considered in the models rather

than a real finding. Comparisons between the eras are limited by the inclusion of far more neonatal units in the more recent cohort. Weight data from the English national cohorts were gathered during routine clinical care and were not subject to standardisation or error checking. Data concerning head circumference and length growth were insufficient to assess and so it is difficult to judge whether weight gain was proportionate to linear growth in these cohorts.

5.6 Conclusions

Comparison of weight gain in infants born in England between the two eras suggested that there have been modest increases in the velocity of weight gain, especially for the most preterm infants, during that period. Early weight loss was less common in the later era. Birth in a Level 3 neonatal unit and early use of parenteral nutrition were associated with a greater magnitude of weight gain. Comparisons between weight gain in England and in Southampton were more difficult and were troubled by unstable SITAR models when few infants were available for analysis. Both SITAR and z-score change analysis suggested that early weight gain was significantly faster in the most preterm infants receiving care in Southampton.

A reduction in early weight loss is an unexpected finding of these data. Early weight loss has typically been viewed as physiological rather than pathological, and my findings challenge that assumption. The finding that infants born in Level 3 neonatal units have faster weight gain lends further support to efforts to centralise births of the most vulnerable infants in the most advanced settings. Recently, the practice of early initiation of parenteral nutrition has been shown to be associated with poorer outcomes in critically ill children and term newborns.[79] These data have led some to question the early use of PN in preterm infants. Data from this chapter show that lack of early PN is associated with slower weight gain across the neonatal stay, highlighting the potential downsides of any move to deferring initiation of PN in this patient group.

Chapter 6 Tracking Growth

This chapter examines approaches to the tracking of the growth of preterm infants, especially the development and use of growth charts. Growth charts were created using the LMS method described in Chapter 3 derived from measurements of infants in Southampton. Infants from the national dataset who experienced normal neurodevelopment were used to generate an alternative set of personalised growth charts.

Generation of growth charts and producing web applications was performed by me and parts of this work have been published as a paper in Archives of Disease in Childhood: Fetal and Neonatal Edition (Appendix 1). [1]

6.1 Background

The first chapter in this thesis set out the current understanding of the expectations for the growth of preterm infants, the charts which benchmark normal growth and the shortcomings and misgivings about the current approaches. This chapter describes the work I have performed to form growth charts using data derived from the Southampton Preterm Nutrition Database.

6.1.1 Concepts Guiding Growth Chart Formation

As explored in Chapter 1, current growth charts for preterm infants are formed using measurements of infants at birth.[27] The consequence of this approach is that infants will remain on their birth centile line if they match the in-utero growth of an equivalent fetus. This implied expected growth rate has been criticised as being possibly unphysiological or unhealthy in the preterm infant, especially as weight loss secondary to water loss is commonly seen shortly after birth.

This chapter presents two approaches to addressing these concerns. LMS growth charts were formed using longitudinal measurement values from preterm infants cared for in University Hospital Southampton where nutritional guidelines were in place. As well as forming charts for weight, head circumference and length, novel charts were formed to track the growth of limb circumferences. These charts benefitted from being derived from infants who were subject to a standardised nutritional approach. Charts for the standard measurements was presented as a poster to the Neonatal Society[157] and published as a peer reviewed article.[148] The text of the

methods, results and discussion of these charts is adapted from drafts of that paper prior to significant changes made by my supervisors.

There are a number of problems with the charts I have created. Firstly, the centile lines at each gestational day reflect a combination of birth values for infants born at that gestational age and later measurements for infants born before that gestation. Therefore, it is difficult to understand how a normally growing infant should grow in relation to the centile lines. Furthermore, these charts share the property of current standard growth charts that initial weight loss (or static weight) due to fluid loss is not reflected.

In an attempt to form charts more closely reflecting the actual growth of preterm infants, I also used the English national data from the NNRD to form individualised growth charts. Infants were included only if there was evidence of normal neurodevelopment from the NNRD. This work was inspired by the approach to individualised growth charts by Landau-Crangle and co-workers.[34] As discussed in Chapter 1, that group used physiological models to construct individualised growth trajectories for infants, with those individual growth charts formed dependent on the infant's gestation and birthweight and made available online. This approach relies on an underlying understanding of the appropriate physiological changes (especially early weight loss due to fluid loss). I took a different approach to forming such individualised charts, selecting infants from the English national dataset who went on to exhibit normal neurodevelopment and forming LMS charts which reflected their growth. These charts overcome some problems with current charts but introduce difficulties of their own which are explored in the discussion section of this chapter.

6.2 Aims

This chapter aimed to:

1. Construct growth charts using at-birth and subsequent measures of weight, length, head circumference, upper arm circumference and thigh circumference for infants born in Southampton.
2. Create a web platform so that clinicians can plot measurements of their infants on these charts.
3. Create a mechanism to produce individualised growth charts for preterm infants, defined by their birthweight and birth gestation, based on the growth of infants in the national English dataset who exhibited normal neurodevelopment.
4. Create a web platform to allow clinicians to generate these individualised growth charts for their infants, to plot their growth and to give some guidance about the current trajectory of growth.

6.3 Methods

6.3.1 Growth Charts from Southampton Data

All patients recruited to the GAP study at the time of analysis were included. These patients represent a subset of the infants included in the Southampton Preterm Nutrition Database (SPND) who had undergone regular measurement of their limb circumferences. Follow-up ceased at discharge from Southampton, reaching 36 weeks corrected gestational age (CGA) or death. Infants were eligible for recruitment if they were born before 30 weeks of gestation during this period.

Nutritional Care

Infants were subject to a standardised nutritional approach detailed in the local nutrition practice section of Chapter 1 of this thesis. In brief, this consisted of early provision of parenteral nutrition, clear guidelines concerning the volume of parenteral nutrition required and the means by which to transition to enteral feeding, standardised use of breastmilk fortifier or preterm infant formula and weekly nutrition-focused MDT meetings. See Appendix 5 for details of the guidance.

Statistical Analysis

Data were gathered from the SPND as detailed in the shared methods chapter of this thesis. Nutritional intakes are expressed as mean daily intakes with standard deviations. Centile charts describing the growth of infants were constructed using the LMS method as described in Chapter 3. LMS functions of the gamlss package (v5.1-3, Stasinopoulos et al, 2019) were used in R (R Core Team 2019, v3.5.3). Centile charts were created using lines at the same centiles as those displayed by the standard UK-NICM chart. Boys and girls were pooled for LMS analysis and growth chart creation as female infants were slightly smaller than their male counterparts but demonstrated the same pattern of growth.

Online Package for Southampton-Based Charts

The Shiny package[158] within RStudio was used to form a web application, published on the shinyapps.io platform (RStudio, 2017).

6.3.2 Individualised Growth Charts from National Data

Cohort Selection

Infants were selected from the English national dataset (see Chapter 4), who were marked as having experienced development which was normal or had less than a three month delay compared to normal infants when assessed at two years of age.

Application of the LMS Method to National Data

Weight and head circumference values were plotted in relation to corrected gestational age, outliers were identified visually and were excluded. As identified in Chapter 4, infants born before 29 weeks of gestation demonstrate a different weight gain pattern than those born after 29 weeks of gestation. Therefore, infants were separated into these two gestation groups and all infants from each group were used to create two distinct LMS models for weight. Models were formed using day of life (up to twelve weeks of life) as the independent variable, using the *lms* function of the *gamlss* package for R.[159] Centile lines were smoothed using natural cubic splines with four knots (set at 0, 5, 18 and 70 days).

A single LMS model was created using head circumference as the dependent variable and day of life as the independent variable. Using the normal settings of the *lms* command, the resultant centile lines exhibited early reduction in head circumference. Whilst this may be a real phenomenon (possibly caused by fluid loss from the scalp or by settling of inflammation caused by the birth process), it is not a process which is well-recognised. Therefore, the LMS process was adjusted so that only a monotonic relationship between the dependent and independent variable was allowed. In practice, this used the *gamlss* function with an argument to permit only a monotonic relationship and specifying the Box-Cox Cole and Green distribution described by Cole in the paper setting out the LMS method.[160]

Web Application for Individualised Charts

A web application was created using the Shiny package for R.[158] It allowed users to input an infant's gestation, sex, birthweight or birth head circumference and subsequent measurement values along with the day of life on which they were measured. It then reported the infant's birthweight (or birth head circumference) centile and birthweight category (e.g. "small for gestational age") based on reanalysed UK 1990 data.[161] The app then generated an individualised growth chart for the infant using the charts created from national data described above, selecting the appropriate weight model (i.e. for infants born before or after 29 weeks of gestational age), calculating the centile of the infant's birth value, plotting a solid line representing growth

following that centile line and dashed lines representing centiles at 0.67 z-score intervals above and below the birth centile. Crossing these dashed lines was equivalent to crossing centile spaces up and down on a traditional growth chart.

The app also provided a traffic light system indicating the current growth trajectory of the infant. This assessment was made using regression modelling of the entered values. When two values were entered, the change in z-score (based on the new national data charts) was calculated and a linear regression model was formed using these values. The model was then used to predict the change in z-score from birth to day 84 of life if that trajectory continued, returning a green traffic light if an increase or a drop of less than 0.67 z-scores was predicted. An amber indication was given if a drop of 0.67 to 1.33 was predicted (i.e. a drop of between one and two centile spaces) and a red indication was returned if a drop of more than 1.33 z-scores was predicted. Where there were more than two measurements given, an initial linear regression model was produced to smooth the initial values, and a subsequent predictive model was used to predict z-score change based on the trajectory at the time of entering the data. Default values were set in the app to provide an example of its use.

6.4 Results

6.4.1 Growth Charts from Southampton Data

At the time of publication of the peer-reviewed paper, 212 infants were included in the GAP study and in analysis for creation of the growth charts. By the time of writing this thesis, 306 infants born before 30 weeks of gestation had been recruited to the GAP study and are included in the charts below.

Demographic characteristics are given as Table 6-1. There were a total of 7903, 3797, 3569, 1905 and 1894 measurements for weight, head circumference, length, left mid-upper arm circumference and left mid-thigh circumference respectively.

Table 6-1. Demographic characteristics of infants included in formation of growth charts using Southampton data. HC - head circumference, Data are median (range) unless otherwise stated.

Gestational age at birth	27^{+0} (23^{+0} to 29^{+6})
Gestational age at discharge or death	36^{+3} (25^{+4} to 53^{+3})
Length of follow-up (days), median (IQR)	58 (35 to 88)
Sex (male), n (%)	179 (58%)
Birthweight (g)	900 (450 to 1620)
Birthweight z-score	-0.34 (-3.4 to 2.37)
Birth HC (cm)	24.3 (19.0 to 29.0)
Birth HC z-score	-0.76 (-3.86 to 2.20)
Birth length (cm)	34.6 (21.2 to 51.6)
Large for gestational age (weight >90th centile), n (%)	7 (2.2%)
Small for gestational age (weight <10th centile), n (%)	66 (23%)

Centile charts for weight, length and head circumference (Figure 6-1) and for limb circumference (Figure 6-2) were generated using the LMS method. Centile lines were smoothed using cubic splines with three knots. Sparse data meant that lines before 24 weeks and after 36 weeks were unstable and so charts were restricted to 24 to 36 weeks corrected gestational age. Male and female infants were pooled for LMS analysis and growth chart creation. Visual comparison of the LMS 50th centile lines for males and females separately indicated that, although male infants were generally larger than females, patterns of growth did not substantially differ for any growth parameter (Figure 6-3), meaning that changes in z-score (equivalent to downwards or upwards 'crossing' of centile lines) can be used to track the progress of growth regardless of sex. Further support for this approach can be found in the routine pooling of infants of different sexes in subgroup analysis of a very large cohort of preterm infants by Cole at al.[25] Left and right limb circumference measurements were very similar (Figure 6-4) and so only the left-sided charts are given.

An online tool illustrating the new growth charts was created and published at bit.ly/sotongrowth as the Southampton Preterm Growth Package. This tool allows users to enter weight, length and head circumference growth data for a patient and plots that data on centile charts. It reports and graphically illustrates changes in z-score and allows the changes in weight, length and head circumference z-score to be easily compared. It also provides a basic simulation of a growing preterm infant to illustrate the application's features. Figure 6-5 shows an example of head circumference values plotted on the chart using the web application.

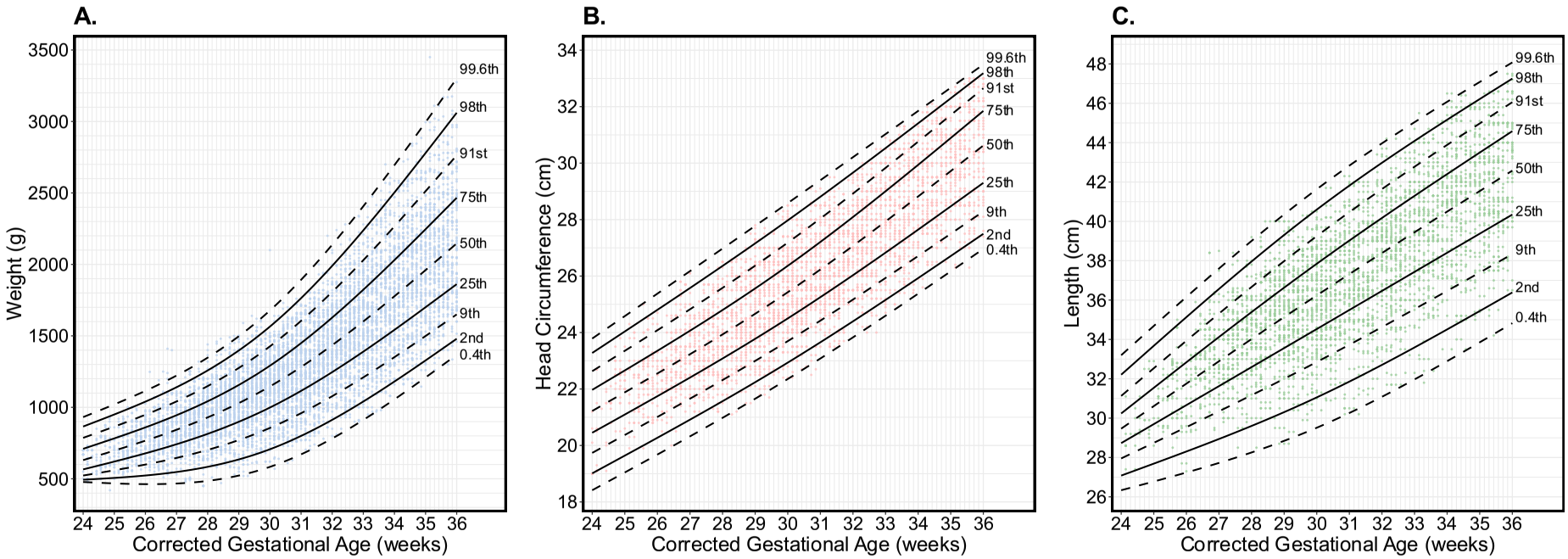


Figure 6-1. Centile charts based on repeated measurements of infants born in Southampton for: A. weight, B. head circumference and C. length from 24 to 36 weeks corrected gestational age. Coloured points are individual measurements.

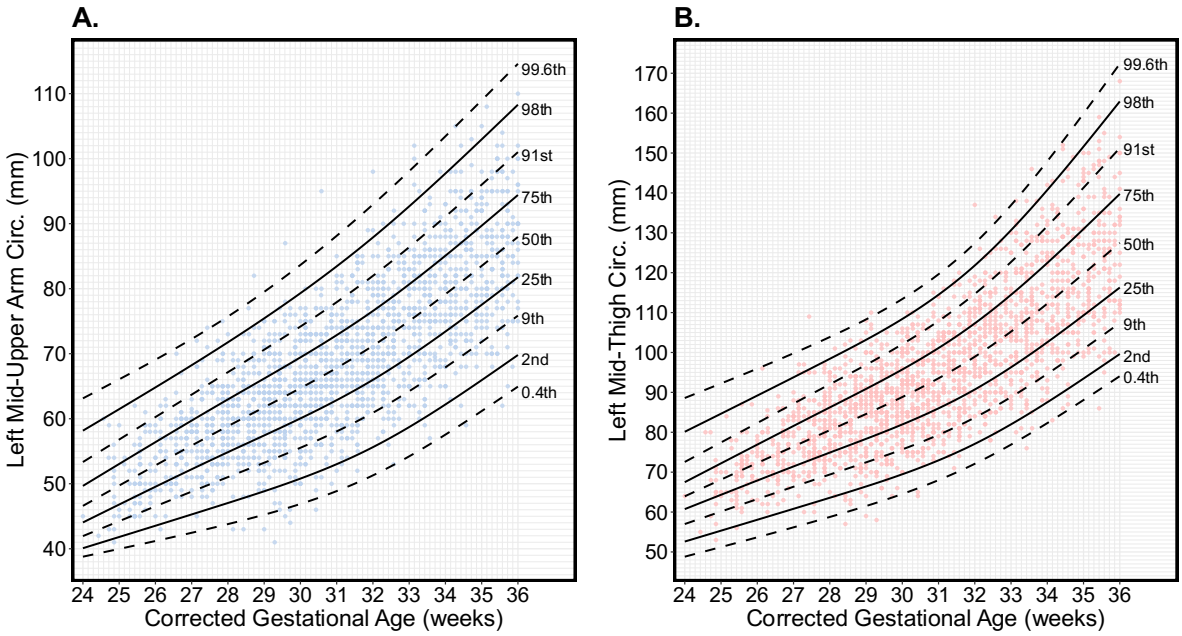


Figure 6-2. Centile charts for: A. left mid-upper arm circumference and B. left mid thigh circumference from 24 to 36 weeks corrected gestational age. Coloured points are individual measurements.

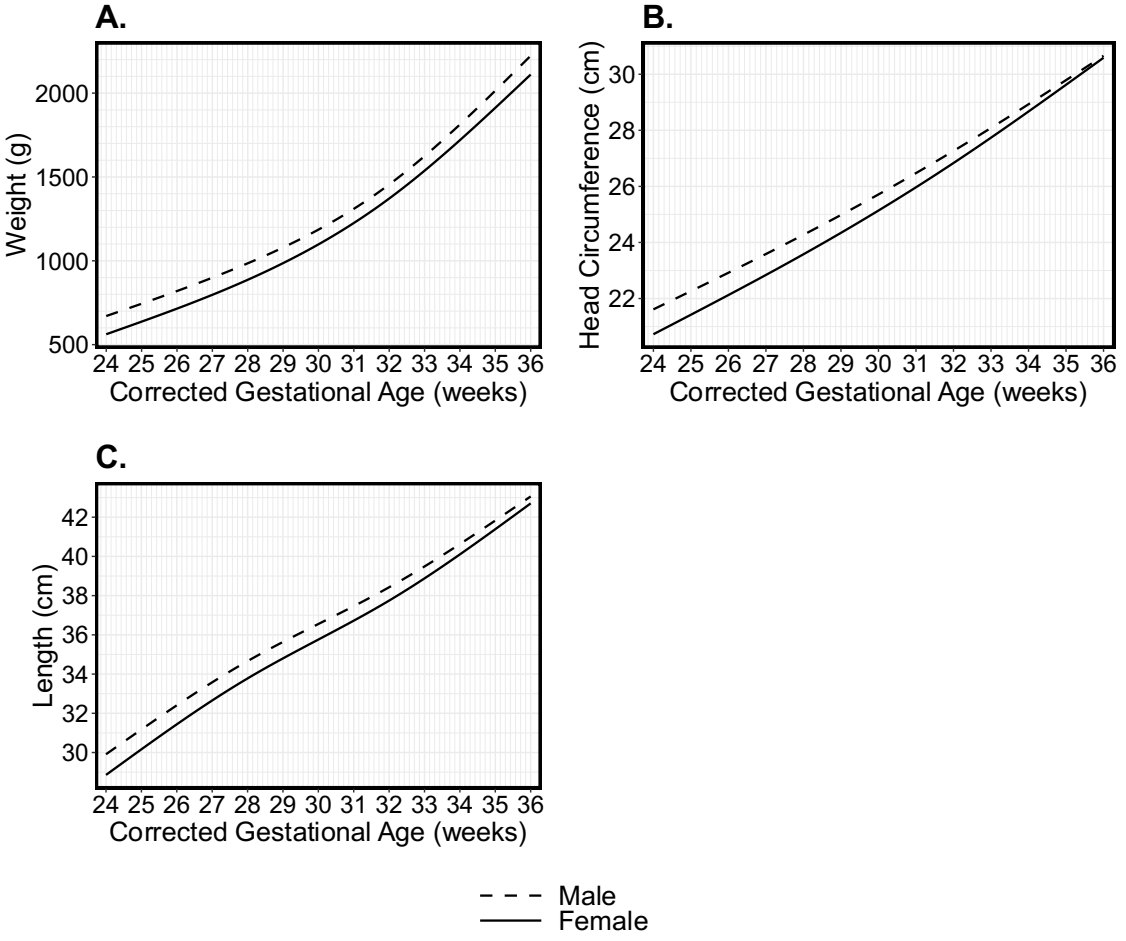


Figure 6-3. 50th centile lines for male and female generated using LMS for A. weight, B. head circumference and C. length.

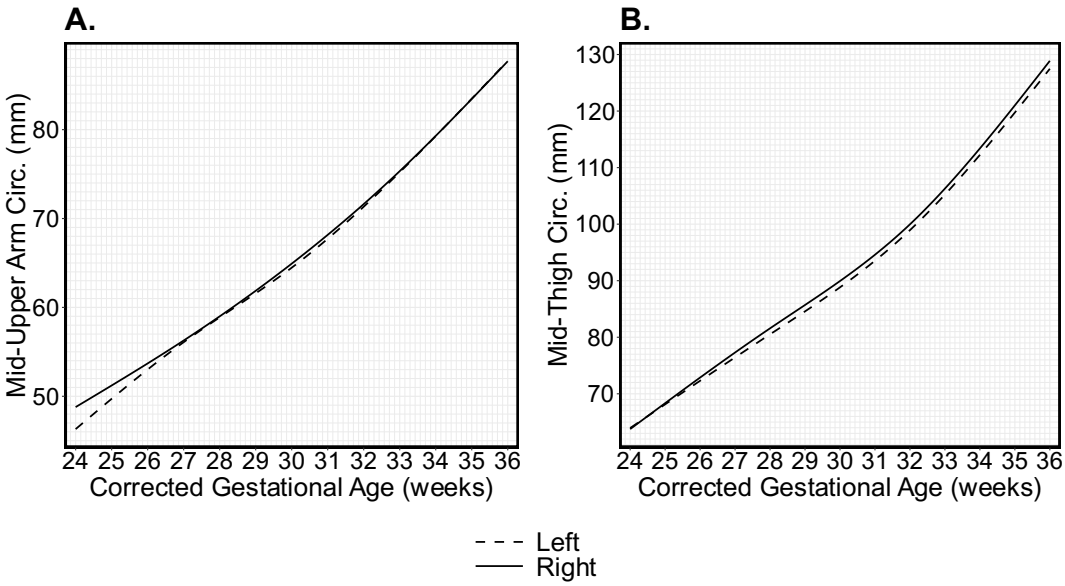
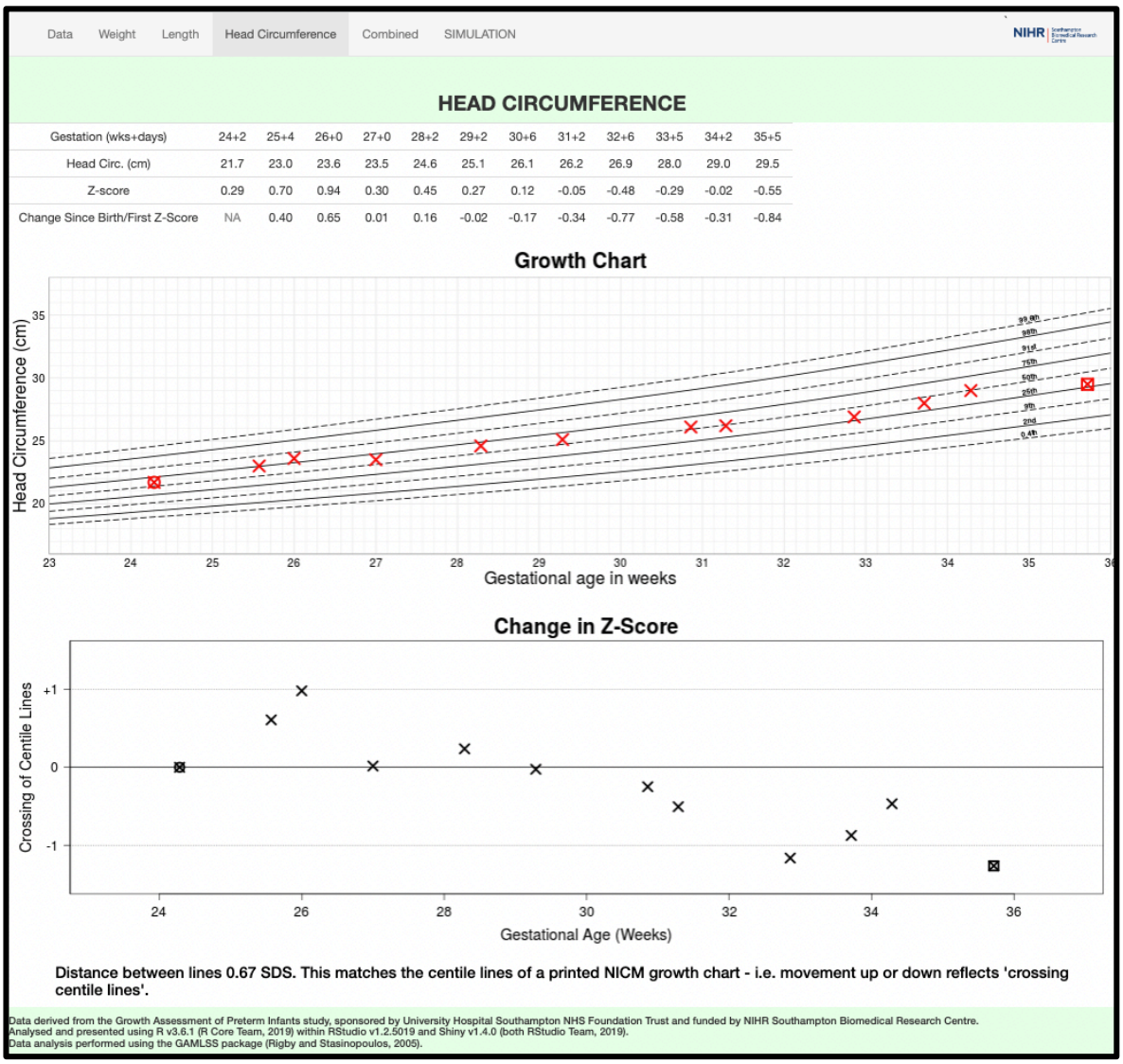


Figure 6-4. 50th centile lines for left-sided and right-sided measurements generated using LMS for A) mid-upper arm circumference, and B) mid-thigh circumference.



Gestation (wks+days)	24+2	25+4	26+0	27+0	28+2	29+2	30+6	31+2	32+6	33+5	34+2	35+5
Head Circ. (cm)	21.7	23.0	23.6	23.5	24.6	25.1	26.1	26.2	26.9	28.0	29.0	29.5
Z-score	0.29	0.70	0.94	0.30	0.45	0.27	0.12	-0.05	-0.48	-0.29	-0.02	-0.55
Change Since Birth/First Z-Score	NA	0.40	0.65	0.01	0.16	-0.02	-0.17	-0.34	-0.77	-0.58	-0.31	-0.84

Figure 6-5. Example of head circumference values plotted on chart using Southampton data.

6.4.2 Individualised Growth Charts from National Data

Cohort Selection

Of the 37700 infants in the English national dataset, 14120 had a documented neurodevelopmental follow-up at two years of age. 9128 infants had normal development (or less than three months of delay) documented and were used to form growth charts. The demographic details of these infants are set out in Table 6-2.

There were some systematic differences between infants for whom neurodevelopmental follow-up was available and those for whom it was not recorded. The median postmenstrual age at birth for infants with follow-up was 28^{+5} weeks (IQR 27^{+0}-29^{+6}) and for those without follow-up it was 30^{+0} weeks (IQR 27^{+4}-31^{+1}). Median birthweight was also lower for included infants.

Table 6-2. Demographic characteristics of infants included in formation of growth charts using national data. HC - head circumference, Data are median (interquartile range) unless otherwise stated.

Postmenstrual age at birth	28^{+6} (27^{+2} to 30^{+0})
Sex (male), n (%)	4575 (50%)
Birthweight (g)	1130 (900 to 1380)
Birthweight z-score	-0.25 (-0.86 to 0.23)
Birth HC (cm)	27.0 (25.0 to 28.5)
Birth HC z-score	-0.43 (-1.14 to 0.33)
Birth length (cm)	37 (34 to 40)
Large for gestational age (weight >90th centile), n (%)	213 (2.3%)
Small for gestational age (weight <10th centile), n (%)	1358 (15%)

LMS Models

For infants born before 29 weeks of gestation, 121152 weight values were available and 67281 weight values were available for infants born after 29 weeks of gestation. 26800 head circumference values were used. Figure 6-6 shows the standard centile lines of the weight and head circumference LMS models.

Web Application

These models were used to form a web app allowing the user to enter measurement values, generate individualised charts and to generate warning traffic lights based on the current trajectory. It was published at www.bit.ly/preterm-plotter. Example screenshots for weight and head circumference are given as Figure 6-7. The web app contained default values to provide an example of how an infant's weight and head circumference could be used.

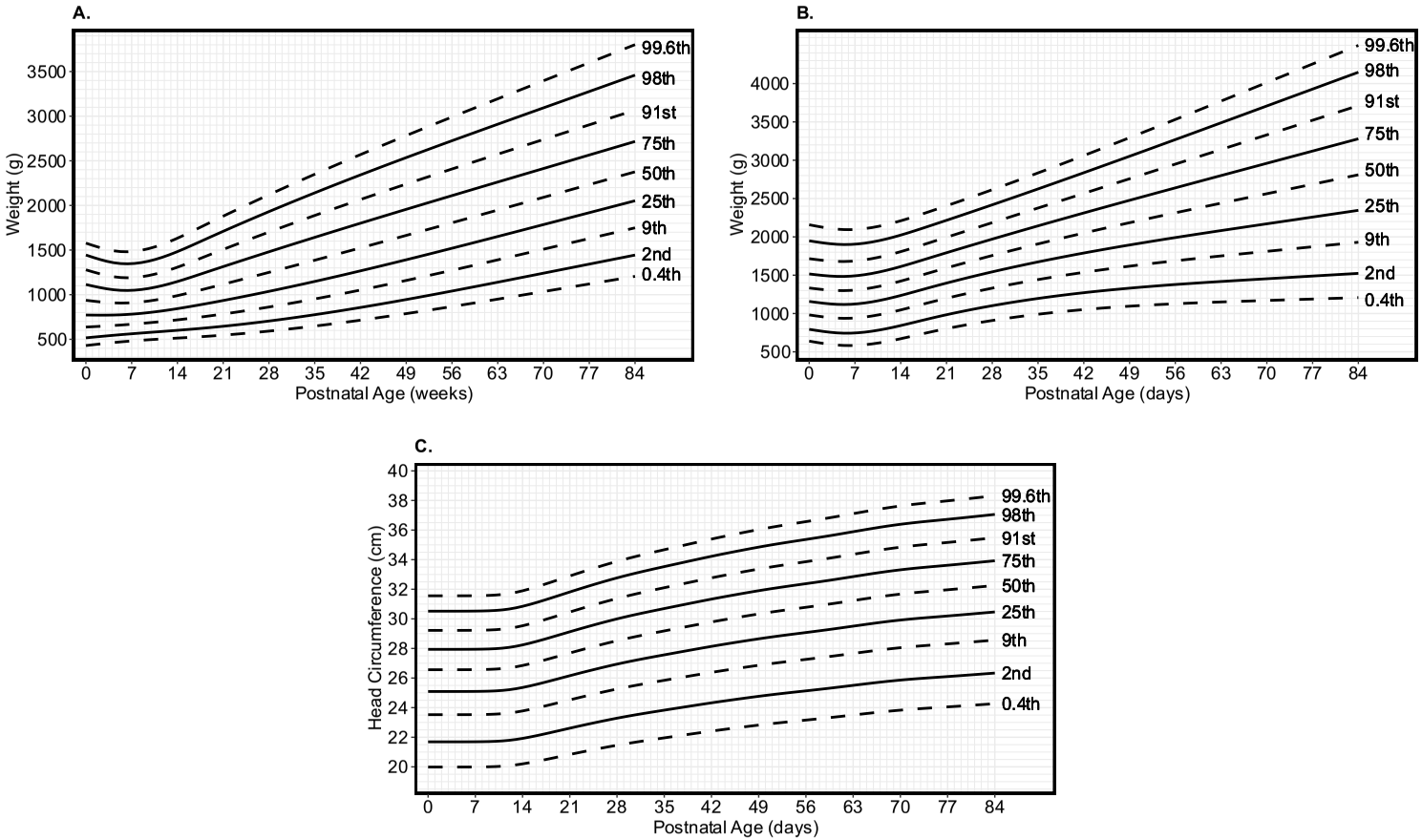


Figure 6-6. LMS models taking day of life as the independent variable for A. weight in infants born before 29 weeks of gestation, B. weight in infants born on or after 29 weeks of gestation and C. head circumference, using national English data restricted to infants exhibiting normal neurodevelopment.

A.

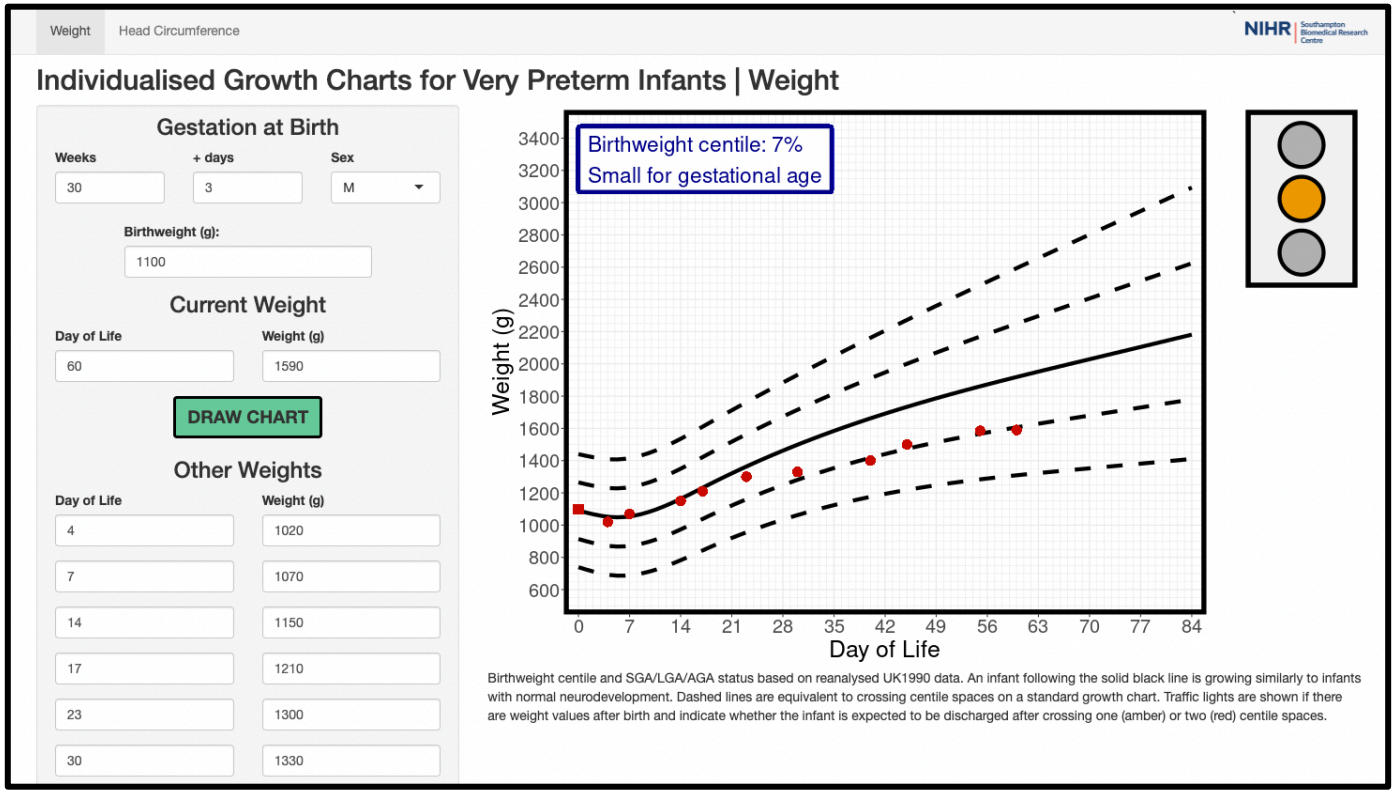
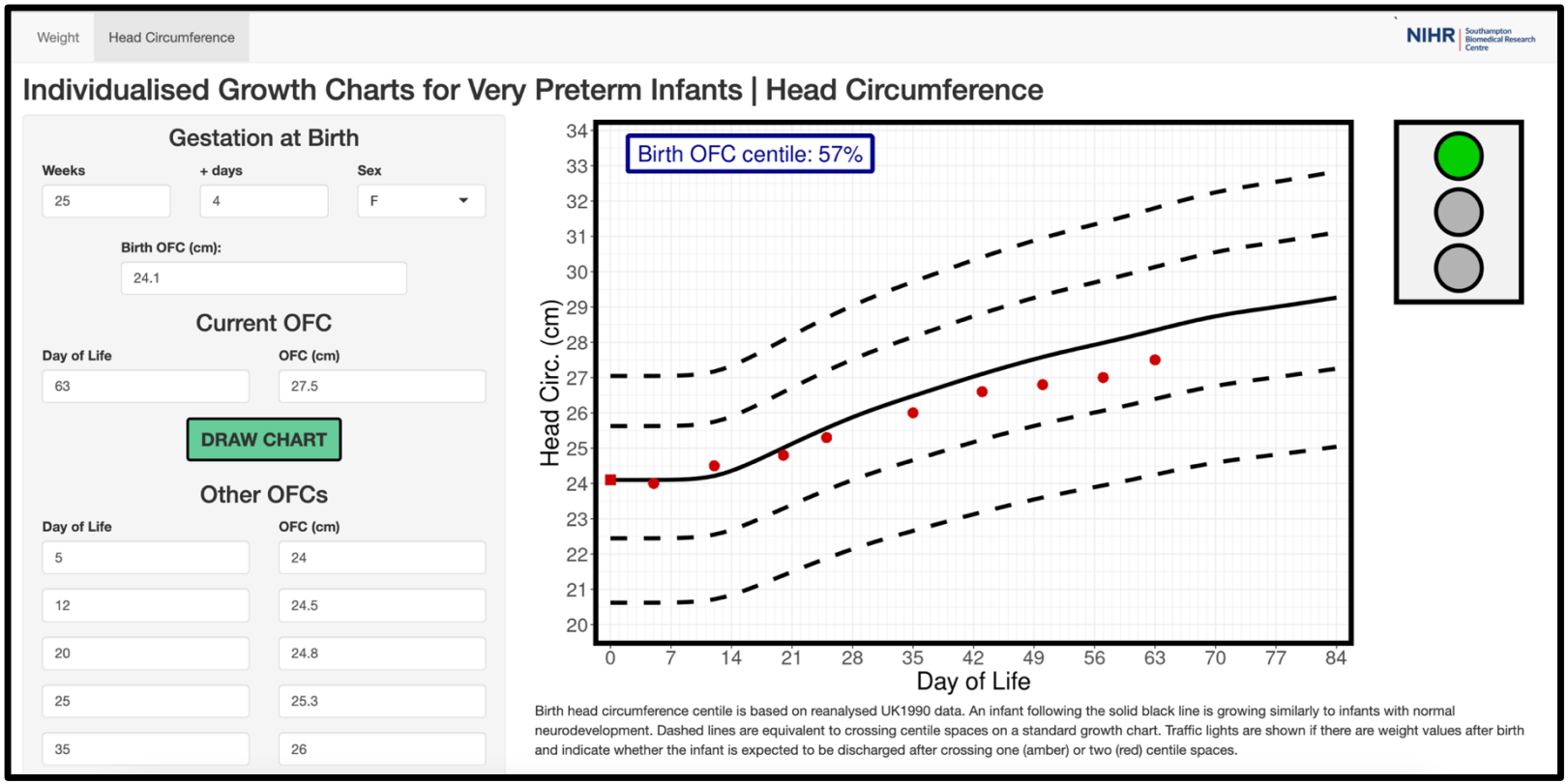


Figure 6-7 Example screenshot of web app for national data (available at www.bit.ly/preterm-plotter),A. with example weight data for an infant born at 30^{+3} weeks demonstrating the birthweight classification using reanalysed UK1990 data in the blue box, an individualised growth chart based on the infants birthweight, subsequent weights plotted and an amber traffic light indicating that modelling predicts that the infant will drop between 0.67 and 1.33 weight z-scores from birth to day 84 of life; and B. (below) with example head circumference data for an infant born at 25 weeks of gestation.

B.


Weight
Head Circumference
NIHR Southampton Biomedical Research Centre
Individualised Growth Charts for Very Preterm Infants | Head Circumference
Gestation at Birth
Weeks
+ days
Sex
25
4
F
Birth OFC (cm):
24.1
Current OFC
Day of Life
OFC (cm)
63
27.5
DRAW CHART
Other OFCs
Day of Life
OFC (cm)
5
24
12
24.5
20
24.8
25
25.3
35
26
Birth OFC centile: 57%
Head Circ. (cm)
Day of Life
Birth head circumference centile is based on reanalysed UK1990 data. An infant following the solid black line is growing similarly to infants with normal neurodevelopment. Dashed lines are equivalent to crossing centile spaces on a standard growth chart. Traffic lights are shown if there are weight values after birth and indicate whether the infant is expected to be discharged after crossing one (amber) or two (red) centile spaces.

6.5 Discussion

Careful monitoring of the growth of preterm infants is a vital element of their care, with frequent measurement of weight alongside head circumference and length recommended by international bodies. In order to be useful, such measurements must be analysed and compared to some standard or benchmark, and this comparison must be rapid and easily understood by clinicians to be practical. Traditional growth charts, such as those produced using reanalysed UK1990 data[27] or by Fenton and co-workers[30], use birthweight values to derive centiles and thus track normal fetal growth. These charts are easy to use but have been criticised for not reflecting the common pattern of growth seen in preterm infants. Specifically, they do not allow for an initial period of weight loss and their centile lines imply that an infant should mimic the growth of a fetus. These issues are considered in greater detail in the first chapter of this thesis.

Repeated Measurements on Southampton Infants

This project generated growth charts using repeated measurements of infants admitted to Southampton's neonatal unit and subjected to a defined nutritional approach. They included anthropometric measurements which were made after birth and therefore do not solely reflect fetal growth patterns. These data were derived from a population of preterm infants for whom detailed nutritional information was available and who mostly received nutritional intakes close to current recommendations. They represented a real-world mixed population of preterm infants, including those with intrauterine growth restriction and those who suffered complications of prematurity. The pattern of growth is similar to that seen in the reanalysed UK1990 data, although the velocity of weight gain is a little slower; Figure 6-8 compares the 2nd, 50th and 98th centile lines of the Southampton charts with the same centile lines of the reanalysed UK1990 charts (taking the mean value for the sexes).

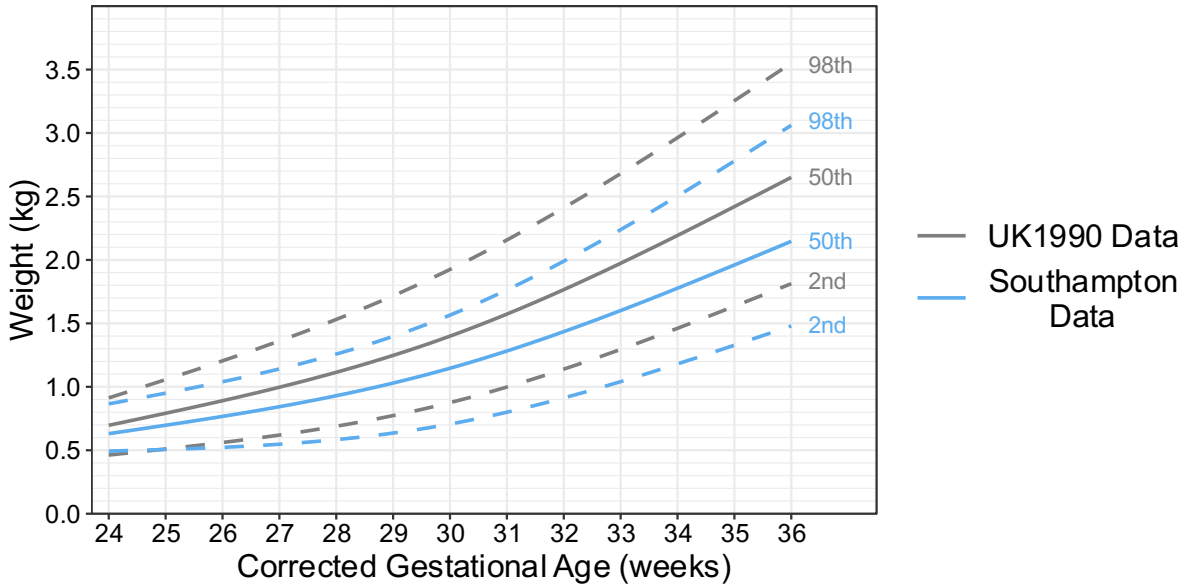


Figure 6-8. Comparison of weight LMS charts generated from reanalysed UK1990 data and infants who received care in Southampton.

These charts are likely to reflect the postnatal growth of preterm infants more closely than those derived purely from birthweight values. More specifically, they illustrate the range of growth which can be expected when infants are fed according to a defined nutritional approach, receiving nutrient intakes which are generally within the current European recommendations (see chapter 4 for nutrient intake information for this cohort of infants). Therefore, growing along the centile lines on these charts could be considered achievable and reflective of generally adequate nutrition. Furthermore, they provide information about the expected pattern of limb growth, which has not previously been described. However, there is no evidence that this would represent ideal growth. In fact, nutrient intake data showed that they did not consistently meet all published nutrient intake guidelines, especially in terms of protein intake.

The concept of a target of ideal growth is a difficult one to define. Growth during the neonatal period is associated with neurodevelopmental outcome (as explored in the introduction to this thesis) and with other health outcomes. Whilst international guidance recommends growth which is similar to the equivalent fetus, there is no strong evidence that such growth is truly optimal, and it is unclear whether neurodevelopmental advantages of rapid growth may be offset by an elevated risk of the metabolic syndrome later in life.[22]

These charts did not reflect the expected early weight loss for preterm infants. This may be due to a real lack of significant early weight loss in these infants (as explored in Chapter 4) or due to the smoothing applied by the LMS algorithm. Early fluid shifts and weight patterns have been investigated in some detail over several decades.[162] The work by Landau-Crangle and co-workers provides an overview of these findings and they define the term “preterm contraction of extracellular spaces” (PreCES) to describe what they consider to be physiological fluid loss (and concomitant weight loss) after birth.[34] That group’s work is based on the belief that early weight

loss is physiological and that the infant's growth trajectory should be considered from the time of maximal weight loss, although they acknowledge that nutritional deficits may contribute to early weight changes. As with much research into very preterm infants, it is difficult to know what a "normal" or "physiological" process entails, as most infants suffer complications of prematurity and all are subjected to medical management of their fluid balance. In the term infant, it has been asserted that a baby born at a normal size after a normal pregnancy, free from perinatal complications and breastfeeding effectively is the model infant. There is no corollary in the very preterm population and so it is difficult to know what constitutes normal early fluid shifts and weight changes. Landau-Crangle's modelling was based on observations made by a related research group, examining weight changes in preterm infants without severe comorbidities.[131] After initial weight loss, Landau-Crangle used growth velocities from Fenton growth charts to plot an expected weight gain pattern, restoring infants to near their birth centile by 42 weeks corrected gestational age. This presents an appealing case for defining an appropriate weight gain pattern. It also benefits from the individualised growth charts approach, which allows the chart to reflect early postnatal weight changes and is supported by a web application to generate the individualised charts. However, the charts rely on an assumption that the early weight loss they permit is healthy, when in reality infants included in that analysis may have been subjected to insufficient nutrition or inappropriate fluid management.

Individualised Charts from National Data

An alternative to this modelling approach is to find a group of infants who have a desirable outcome and to base charts on their growth. Parents consistently identify that neurodevelopmental outcomes are the most important target for research into preterm infants.[163] Therefore, I identified a cohort of very preterm infants from the English national dataset who had confirmed normal (or near-normal) development when assessed at two years of age. Only around half of infants in the cohort had those data available (some had missing data and some had not reached two years old by the time data were extracted). Detailed neurodevelopmental indices were available for even fewer, and so analysis was limited to the overall clinician assessment.

The use of these data allowed individualised growth charts to be created, which provide a guide to the typical weight gain and head circumference growth exhibited by infants who had a good neurodevelopmental outcome. Development of the web application allowed these charts to be generated for individuals on demand and for measurements to be plotted electronically. In practice, this step would likely be implemented within a clinical information system rather than as a standalone website. A simple system of alerting the clinician to the current trajectory was also implemented, although it has not yet been validated against important outcomes. These

individualised charts provide a more outcome-focused alternative to those formulated by Landau-Crangle. They are also more agnostic in their approach, simply reflecting the growth of normally developing infants without applying assumptions about early weight changes. They are also agnostic as to the ultimate centile of the infant and do not assume that optimal outcomes will be achieved when an infant reaches his or her birth centile at 42 weeks corrected gestational age.

Despite these advantages, there are some limitations of the charts. Firstly, they use only neurodevelopment as an outcome and do not consider other important outcomes. They are likely to include infants with insufficient nutrition or inappropriate fluid management if they nevertheless had a good developmental outcome. Infants for whom there was missing follow-up data were excluded from analysis. This may have introduced some bias, especially as follow-up is significantly less common for infants born after 30 weeks postmenstrual age, as national audit standards do not oblige clinicians to follow up these patients. Considering the high rate of missingness in these data, it is possible that other unmeasured imbalances are present.

The data used for these charts includes infants who were small for their postmenstrual age at birth (SGA). Fifteen percent of included infants were SGA as judged by the UK national standards. This cohort must represent a mixture of infants who are appropriately small (i.e. their genetic and familial background means that they are simply small infants) and those whose growth has been restricted by an adverse in utero environment (IUGR). SGA infants have been shown to exhibit a different pattern of growth to larger infants.[164] In that study, SGA infants had a lower average downward deviation from their birthweight-derived birth centile and were more likely to grow faster than their centile. Most of the SGA infants who experienced an increase in weight z-score in this study were born after pregnancies complicated by pre-eclampsia (a condition which is liable to cause intrauterine growth restriction). These charts do not differentiate between infants who are appropriately SGA and those who are growth restricted, despite the evidence that their growth trajectories may be expected to differ.

6.6 Conclusions

This chapter presents two distinct approaches to generating charts to track the growth of preterm infants: traditional LMS charts with the postmenstrual age on the x-axis generated from measurements during the growth of preterm infants in Southampton, and individualised charts which can be generated for an infant at birth with age along the x-axis and based on infants from the national dataset with a favourable neurodevelopmental outcome.

The first approach produced charts which reflect an achievable growth pattern of preterm infants. However, these charts did not truly reflect the growth patterns of real infants who often

experience early weight loss or slow weight gain. They are also difficult to interpret fully because each timepoint reflects the growth of infants who are at different postnatal ages.

With the advent of electronic clinical information systems, there is a potential to form growth charts which more properly reflect the actual growth pattern expected of a preterm infant. The individualised growth charts presented here are generated from infants with a favourable outcome. However, identifying a “healthy” cohort of preterm infants is difficult and the measure of favourable neurodevelopmental outcome used in this chapter does not exclude infants who had other poor outcomes. Furthermore, the classification of infants into a “healthy” cohort is based on a one-time clinical assessment in which infants are broadly defined as having “normal development or less than three months delay” at around two years old. This is a subjective judgment made by one clinician and is not subject to standardisation between clinicians, neonatal units or regions. Whilst the individualised growth chart technique holds promise and I consider it to be superior to charts which cannot take account of early weight loss, defining a cohort of infants with optimal outcomes will continue to present a serious challenge. This question depends on which outcomes are considered important. For example, if severe cerebral palsy were to be considered the only important outcome, assessment at two years would likely suffice. If neurodevelopmental and educational attainment were valued, some assessment of the school-age child would be required. However, if it is considered that the growth pattern in early life is likely to influence cardiometabolic outcomes throughout the life course, follow-up into middle age may be needed! In practice, follow-up into adolescence, when emerging metabolic derangement can be detected, may facilitate the identification of a truly “healthy” cohort. As with so many current questions in neonatal medicine, efforts to extend and refine follow-up will present opportunities to identify a cohort of infants with good outcomes, whose growth during the neonatal period may act as a template for personalised growth chart creation.

Chapter 7 Protein Intake and Serum Urea

Serum urea level has been used to assess sufficiency of protein intake in preterm infants. This chapter aims to assess the relationships between protein intake, serum urea and growth to explore the utility of this approach in nutritional management.

The work in this chapter was carried out by Aneurin Young. I am grateful for the expert advice of Dr Rodney Gilbert, Consultant in Paediatric Renal Medicine. This work was presented to the spring meeting of the Neonatal Society in March 2022.

7.1 Background

As discussed in the opening chapter of this thesis, international guidelines recommend that the preterm infant should exhibit growth which matches both the weight gain pattern and the body composition trajectory of the gestational age-equivalent fetus in utero.[23, 24] In practice, this is rarely achieved, with infants born preterm demonstrating both a deficit in weight gain[25] and abnormalities of body composition[35] compared to term-born infants when they reach term-equivalent age. Protein (or amino acid) intake is especially important to promote lean growth and studies have demonstrated a persistent cumulative protein deficit during the initial hospital admission of preterm infants.[165]

The metabolism of nitrogenous compounds can be conceptualised as a 'pool' of free amino acids in the cells and extracellular fluid, with amino acids entering the pool from dietary sources or protein breakdown and exiting by incorporation into protein (or other nitrogenous compounds) or by catabolism, forming urea. Figure 7-1 highlights the relationships between protein intake, urea production and growth in the context of this amino acid pool. This figure demonstrates that plasma urea is likely to be influenced by dietary protein intake but that several other factors, especially the mass of protein used for growth, will also influence urea production. Plasma urea will be determined by the balance between production of urea, its elimination by the kidneys and the hydration status of the infant.

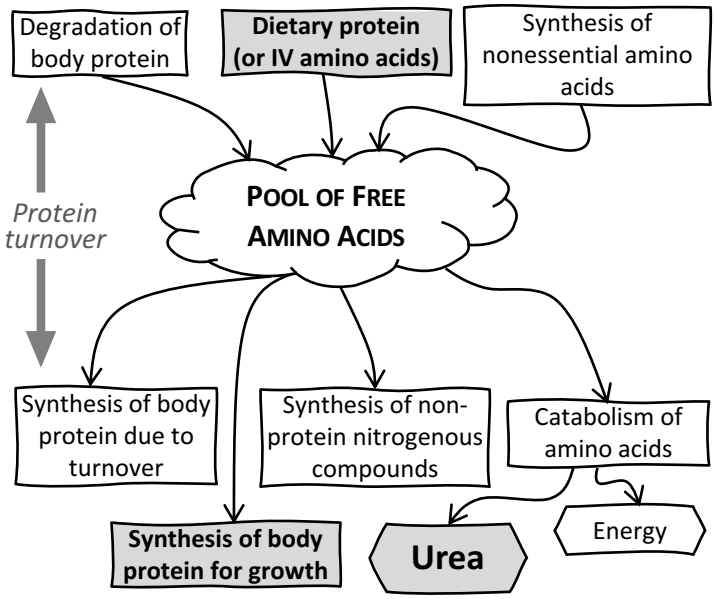


Figure 7-1. The relationship between protein intake, urea production and growth, conceptualised as their interactions with the pool of total body free amino acids.

Recent guidelines have recommended that parenteral amino acid provision should be 2.5-3.5g/kg/day[166] and enteral protein supply should be around 3.5-4.5g/kg/day.[23] Despite these established guidelines, the protein requirement of an individual infant on a given day may be influenced by numerous factors. Whilst some guidelines differentiate between infants weighing more than 1kg and those weighing less than 1kg (with the larger infants requiring less protein in g/kg/day terms), there is no physiological reason to believe that there is a clear cut-off at the 1kg threshold. Gestation, postnatal age and growth restriction are likely to influence protein requirements. Furthermore, sepsis, inflammation or surgical intervention (all of which are common in preterm infants) may place the infant in a state of obligatory catabolism, with a significantly reduced capacity to process protein or amino acids.

As an alternative to simply targeting a set protein intake range, there has been interest for several decades in using blood urea as a marker of sufficient protein intake. Polberger and co-workers examined the relationship between serum urea and protein intake in the context of a trial of enhanced protein provision.[89] They suggested that a serum urea level less than 1.6mmol/l was likely to represent protein intake deficiency. Subsequent investigators have been more guarded in their conclusions, stating only that poor growth in combination with a low urea level (<3mmol/l) is likely to reflect inadequate protein intake.[144] Two small trials demonstrated increased weight gain when protein supplementation of human milk was adjusted in response to blood urea.[85, 167] However, in both of these studies total protein provision was significantly higher in the urea-adjusted group, rendering it unclear whether urea-targeting *per se* was responsible for the increased growth, or whether the effect was simply due to an overall increase in protein provision.

This chapter aims to define the relationship between protein (or amino acid) intake and plasma urea measurement in very preterm infants during the period from 14 to 42 days postnatal age. Specifically, the use of a urea threshold to assess for protein adequacy was investigated. This period was selected as it is a key phase of growth, after initial fluid shifts but prior to discharge from the neonatal unit.

7.2 Methods

Patients

Data were extracted from the Southampton Preterm Nutrition Database (SPND) and were limited to infants born before 30 weeks of gestation and born before September 2021 (when this analysis was performed).

Calculation of Dependent and Independent Variables

Glomerular filtration rate (GFR) was estimated using the method for preterm infants described by Brian and co-workers, using the body length measurement nearest in time to the relevant measurement of plasma creatinine.[168, 169] Protein or amino acid intake for the three days prior to plasma urea sampling was expressed as a mean value. Non-protein energy intake (in kcal), parenteral nutrition volume and enteral nutrition volume were similarly calculated for the three days prior to urea measurement and averaged. C-reactive protein (CRP) was included as a covariate to reflect the impact of inflammation on the balance between catabolism and anabolism, with values accepted from up to two days prior to the relevant urea sample if CRP was not measured on the index day. Infants were classified as parenterally fed if the mean intake of parenteral nutrition for the three days prior to blood sampling exceeding 80ml/kg, and were classified as enterally fed if enteral feeds averaged more than 140ml/kg/day. Nutritional guidelines in place on the neonatal unit meant that these groups were distinct.

7.2.1 Statistical Analysis

Data were manipulated and analysed in R,[111] using the *lme4* package to perform mixed effects regression.[170] The relationship between plasma urea and protein intake during the preceding three days was first assessed by simple linear regression, taking plasma urea as the dependent variable (with a p value <0.05 interpreted as significant). Mixed effects models were then tested, taking participant identity as a random effect, and nutritional, demographic and clinical status variables as fixed effects. Fixed effects are reported as effect sizes with 95% confidence intervals, along with p values.

The definition of renal impairment in preterm infants is complicated by rapidly changing GFR and by the dual influence of gestation and postnatal age on GFR.[171] For this analysis, the LMS method[28] was used to define centile values for eGFR based on postnatal age and measurements falling more than two standard deviation scores below the mean were discarded as likely to be affected by significant renal impairment. Where infants had more than one plasma urea value during a single day, the mean average of all values was taken.

The performance of plasma urea as a marker of sufficient protein intake (at least 3.5g/kg/day, as recommended by ESPGHAN)[23] was assessed by the receiver operator characteristic (ROC) curve using the *ROCit* package for R. A linear regression model, including each of the demographic and clinical factors, was used to generate a protein intake prediction algorithm. The performance of this value as a predictor of actual protein intake was also assessed.

For analyses examining influences on growth between 14 and 42 days postnatal age, the most recent measurement on each of those days was taken and the change in standard deviation score was calculated using UK-WHO standards.[27] Nutrient intakes during the period were summarised by taking the mean daily intake value. For days during which plasma urea was not measured, an estimated urea value was interpolated by linear regression using the preceding and following urea values, with the mean of all actual and interpolated values taken as the summary figure for plasma urea during the period. This method of interpolating urea values is commonly used in clinical practice and research applications.[172]

7.3 Results

Urea results were available for 352 infants. Demographic details of included infants, biochemical data and daily nutrition data are summarised in Table 7-1.

Table 7-1. A. Demographic features of infants - mean (range) unless specified. B. Descriptive features of serum urea measurements; and C. Daily macronutrient intakes. Median (IQR) unless specified IUGR – intrauterine growth restriction <10th centile; eGFR – estimated glomerular filtration rate by Brian method.

A. Infants	Gestational age (weeks)	27+0 (23+2 - 29+6)
	Male, n(%)	201 (57)
	Birth weight (kg)	0.91 (0.45 - 1.5)
	Birth weight SD score	-0.6 (-3.4 - +2.2)
	IUGR, n(%)	89 (25)
B. Plasma urea measurements	Plasma urea (mmol/l)	3.6 (2.3 - 5.5)
	eGFR (ml/min/1.73m2)	30 (23 - 38)
	Full parenteral nutrition, n(%)	506 (17)
	Full enteral feeding, n(%)	830 (27)
C. Daily Macronutrient Intakes	Total energy intake (kcal/kg/day)	126 (111 - 139)
	Combined protein and amino acid intake (g/kg/day)	3.4 (2.8 – 3.9)
	Carbohydrate intake (g/kg/day)	15 (13 – 16)
	Fat intake (g/kg/day)	6.2 (4.8 – 7.0)

7.3.1 Assessment of Renal Impairment

Simple linear regression identified a positive relationship between postnatal age and eGFR during the age range in question (each additional week of life being associated with an increase in eGFR of 6.0ml/min/1.73m^2, 95%CI 5.6-6.3, p<0.001). A centile chart illustrating the LMS model for eGFR is given as Figure 7-2. Estimated GFR values for 83 samples (2.4%) were more than 2 SD scores below the median and data for these samples were discarded.

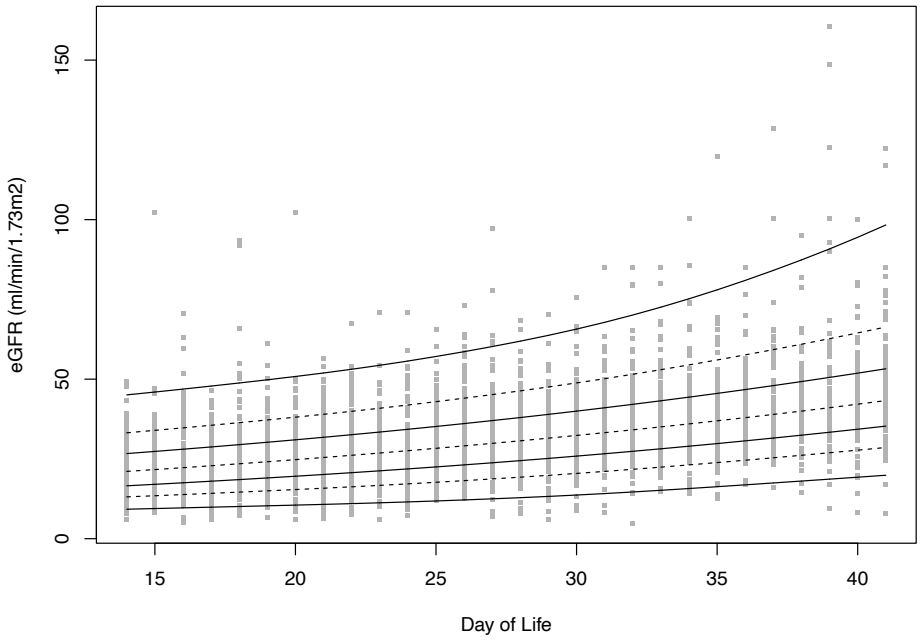


Figure 7-2. LMS centile chart for eGFR by day of life (points are individual measurements, centile lines are at the 2nd, 10th, 25th, 50th, 75th, 90th and 98th percentiles). eGFR – estimated glomerular filtration rate.

Effect of Recent Protein Intake on Plasma Urea

Simple linear regression identified a positive association between protein intake and plasma urea, with each increase of 1g/kg/day protein intake associated with an increase in plasma urea of 0.6mmol/l (95% confidence interval: 0.5 to 0.8, p<0.001) (Figure 7-3-A). This effect was strongest for infants receiving full parenteral nutrition (defined as at least 80ml/kg/day) with a 1g/kg/day increase in protein associated with increase in plasma urea of 1.4mmol/l (95%CI: 0.7 to 2.0, p<0.001) compared to a urea increase of 0.9 (95% CI: 0.7 to 1.1, p<0.001) in the enterally fed group (Figure 7-3-B). A transitional feeding group, receiving 30ml/kg/day enteral feed but not meeting the full PN or full enteral feed criteria showed an intermediate response (1g/kg/day protein associated with increase in plasma urea of 1mmol/l (95% CI: 0.8-1.2 ,p<0.001) (Figure 7-3-C).

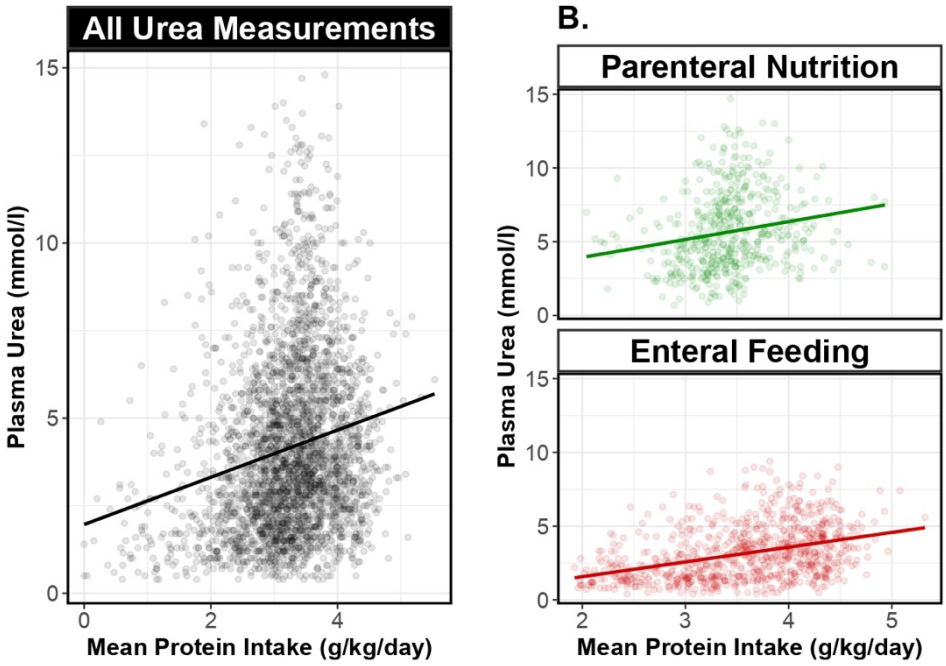


Figure 7-3. Association between plasma urea and mean protein intake during the three days prior to urea sampling during A. all nutritional intake approaches (n=3209 samples from 352 infants) and B. when receiving at least 80ml/kg/day parenteral nutrition (n=542 samples from 134 infants) or full enteral feeds of at least 140ml/kg/day (n=837 samples from 244 infants). Lines indicate linear regression.

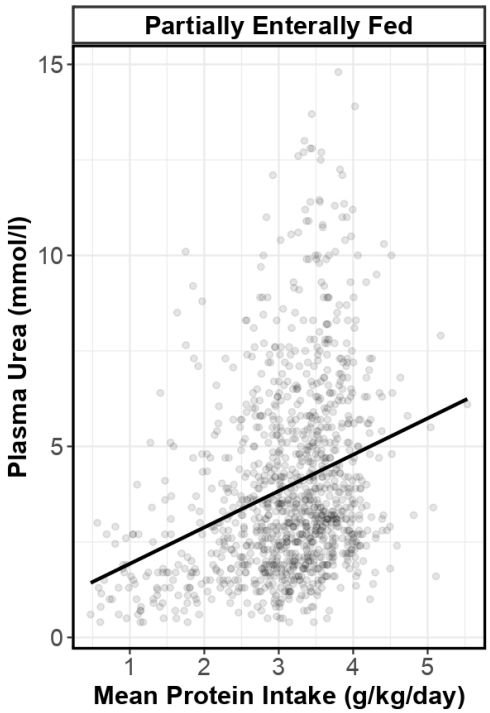


Figure 7-4. Association between plasma urea and mean protein intake during the three days prior to urea sampling during feeding with at least 30ml/kg/day enteral feed but not meeting criteria for full PN or full enteral feeds (n=1208 samples from 274 infants)

7.3.1.1 Mixed Effects Regression

Mixed effects regression was used to adjust for demographic and clinical features, as well as the repeated measures undertaken for each subject. Results of mixed effects regression are given in Table 7-2, along with models formed for infants receiving full parenteral nutrition and full enteral feeds separately. Infants undergoing transition from parenteral to enteral feeding demonstrated an intermediate response (Table 7-3).

Table 7-2. Effects of nutritional intakes and demographic factors on plasma urea by mixed effects regression with participant identity taken as a random effect (with 95% confidence intervals of the effect size) for A. all urea measurements, B. urea measurements taken in the context of full parenteral nutrition and C. urea measurements taken in the context of full enteral feeding. eGFR – estimated glomerular filtration rate.

A. All samples

3209 samples from 352 infants

Variable	Effect on plasma urea mmol/l (95% CI)	p-value
Preceding protein intake (g/kg/day)	1.6 (1.4 to 1.8)	<0.001
Preceding non-protein energy intake (kcal/kg/day)	-0.040 (-0.047 to -0.033)	<0.001
Gestational age (weeks)	-0.40 (-0.51 to -0.28)	<0.001
Sex (female)	-0.44 (-0.80 to -0.08)	0.02
Birth weight SD score	0.33 (0.14 to 0.51)	<0.001
Postnatal age (weeks)	-0.35 (-0.43 to -0.28)	<0.001
eGFR (ml/min/1.73m^2)	-0.039 (-0.047 to -0.030)	<0.001
C-reactive protein (mg/l)	0.011 (0.009 to 0.013)	<0.001

B. Samples preceded by full parenteral nutrition

542 samples from 134 infants

Variable	Effect on plasma urea mmol/l (95% CI)	p-value
Preceding protein intake (g/kg/day)	3.0 (2.2 to 3.9)	<0.001
Preceding non-protein energy intake (kcal/kg/day)	-0.10 (-0.12 to -0.07)	<0.001
Gestational age (weeks)	-0.42 (-0.67 to -0.17)	0.002
Sex (female)	-1.1 (-1.8 to -0.4)	0.004
Birth weight SD score	-0.19 (-0.17 to 0.55)	0.3
Postnatal Age (weeks)	0.14 (-0.13 to 0.40)	0.3
eGFR (ml/min/1.73m^2)	-0.076 (-0.099 to -0.052)	<0.001
C-reactive protein (mg/l)	0.0055 (-0.0015 to 0.012)	0.13

C. Samples preceded by full enteral feeding

837 samples from 244 infants

Variable	Effect on plasma urea mmol/l (95% CI)	p-value
Preceding protein intake (g/kg/day)	1.3 (1.1 to 1.5)	<0.001
Preceding non-protein energy intake (kcal/kg/day)	-0.030 (-0.044 to -0.018)	<0.001
Gestational age (weeks)	-0.12 (-0.23 to -0.01)	0.03
Sex (female)	-0.05 (-0.37 to 0.27)	0.8
Birth weight SD score	0.30 (0.13 to 0.48)	<0.001
Postnatal Age (weeks)	-0.64 (-0.75 to -0.54)	<0.001
eGFR (ml/min/1.73m^2)	-0.018 (-0.032 to -0.005)	0.008
C-reactive protein (mg/l)	0.012 (0.005 to 0.019)	<0.001

Table 7-3. Effects of nutritional intakes and demographic factors on plasma urea by mixed effects regression with participant identity taken as a random effect (with 95% confidence intervals of the effect size) for urea measurements taken in the context of feeding with at least 30ml/kg/day enteral feed but not meeting criteria for full PN or full enteral feeds. eGFR – estimated glomerular filtration rate.

1208 samples from 274 infants

Variable	Effect on plasma urea mmol/l (95% CI)	p-value
Preceding protein intake (g/kg/day)	1.7 (1.4 to 2.1)	<0.001
Preceding non-protein energy intake (kcal/kg/day)	-0.042 (-0.055 to -0.028)	<0.001
Gestational age (weeks)	-0.39 (-0.55 to -0.23)	<0.001
Sex (female)	-0.33 (-0.82 to 0.16)	0.2
Birth weight SD score	0.29 (0.04 to 0.54)	0.03
Postnatal age (weeks)	-0.13 (-0.28 to 0.02)	0.07
eGFR (ml/min/1.73m^2)	-0.050 (-0.064 to -0.036)	<0.001
C-reactive protein (mg/l)	0.012 (0.008 to 0.015)	<0.001

Receiver Operator Characteristic Assessment of Plasma Urea and Protein Intake

The area under the ROC curve for plasma urea as a predictor of adequate protein intake (at least 3.5g/kg/day) was 0.57 (Figure 7-5-A). The area under the ROC curve for a score derived from regression analysis including gestational age, postnatal age, sex, birthweight SD score, estimated GFR and CRP was little better at 0.71 (Figure 7-5-B). Use of a definition of protein adequacy of 3g/kg/day did not significantly alter the AUC values (Figure 7-6).

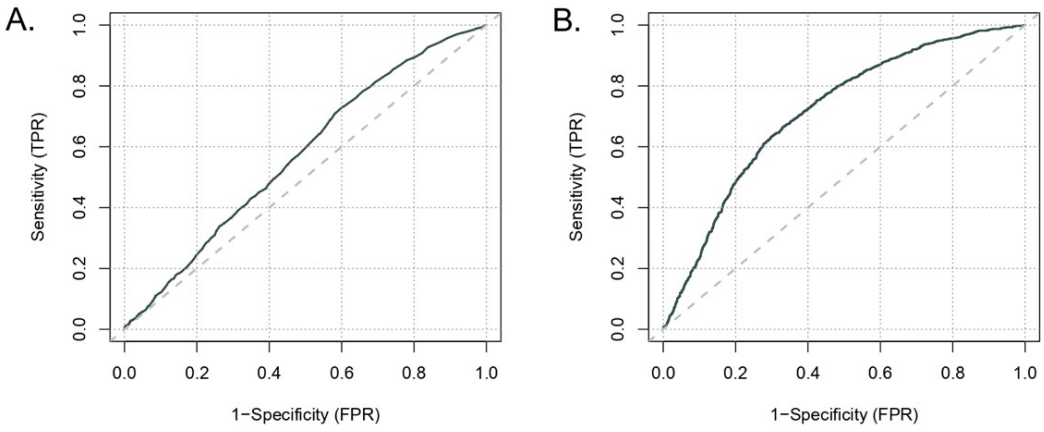


Figure 7-5. ROC curve analysis of the performance of A. plasma urea to predict adequate protein intake (at least 3.5g/kg/day; AUC 0.57), and B. a clinical score derived from multiple linear regression of urea, gestational age, postnatal age, sex, birthweight SD score, estimated GFR and CRP to predict adequate protein intake (at least 3.5g/kg/day; AUC 0.71).

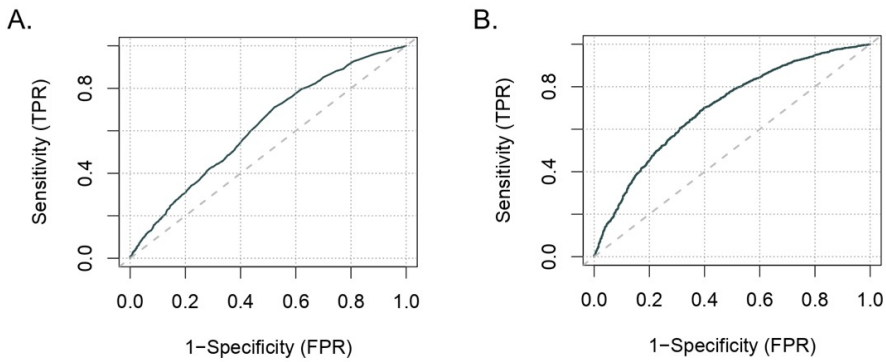


Figure 7-6. ROC curve analysis of the performance of A. plasma urea to predict adequate protein intake (at least 3g/kg/day; AUC 0.62), and B. a clinical score derived from multiple linear regression of urea, gestational age, postnatal age, sex, birthweight SD score, estimated GFR and CRP to predict adequate protein intake (at least 3/kg/day; AUC 0.70).

7.3.2 Influences on Weight Gain During the Third to Sixth Week of Life

246 infants had weight values available at day 14 and day 42 and had complete nutritional intake data, after exclusion of infants with a minimum estimated GFR SD score of less than -2. 227 infants were included in analysis of head circumference.

After adjustment for gestational age and birthweight SD score, higher mean protein intakes were associated with higher plasma urea values, with a rise of 1g/kg/day protein intake associated with an increase in mean plasma urea of 0.8mmol/l (95% CI 0.5 to 1.1) (Figure 7-7-A). Change in weight SD score from 14 to 42 days postnatal age was positively associated with protein intake after adjustment for gestational age, birthweight SD score and non-protein energy intake, with each increase in protein intake of 1g/kg/day associated with a +0.3 change in weight SD score (95%CI: 0.1 to 0.5) (Figure 7-7-B). However, an association between plasma urea and change in weight SD score was not demonstrated (adjusted 95% confidence interval: -0.06 to 0.02) (Figure 7-7-C). Analyses examining associations with growth of head circumference similarly showed an association with protein intake (increase in mean protein intake of 1g/kg/day associated with adjusted head circumference SD score change of +0.5; 95% CI: 0.2 to 0.9) but not with plasma urea (adjusted 95% CI: -0.07 to 0.07) (Figure 7-7-D-E).

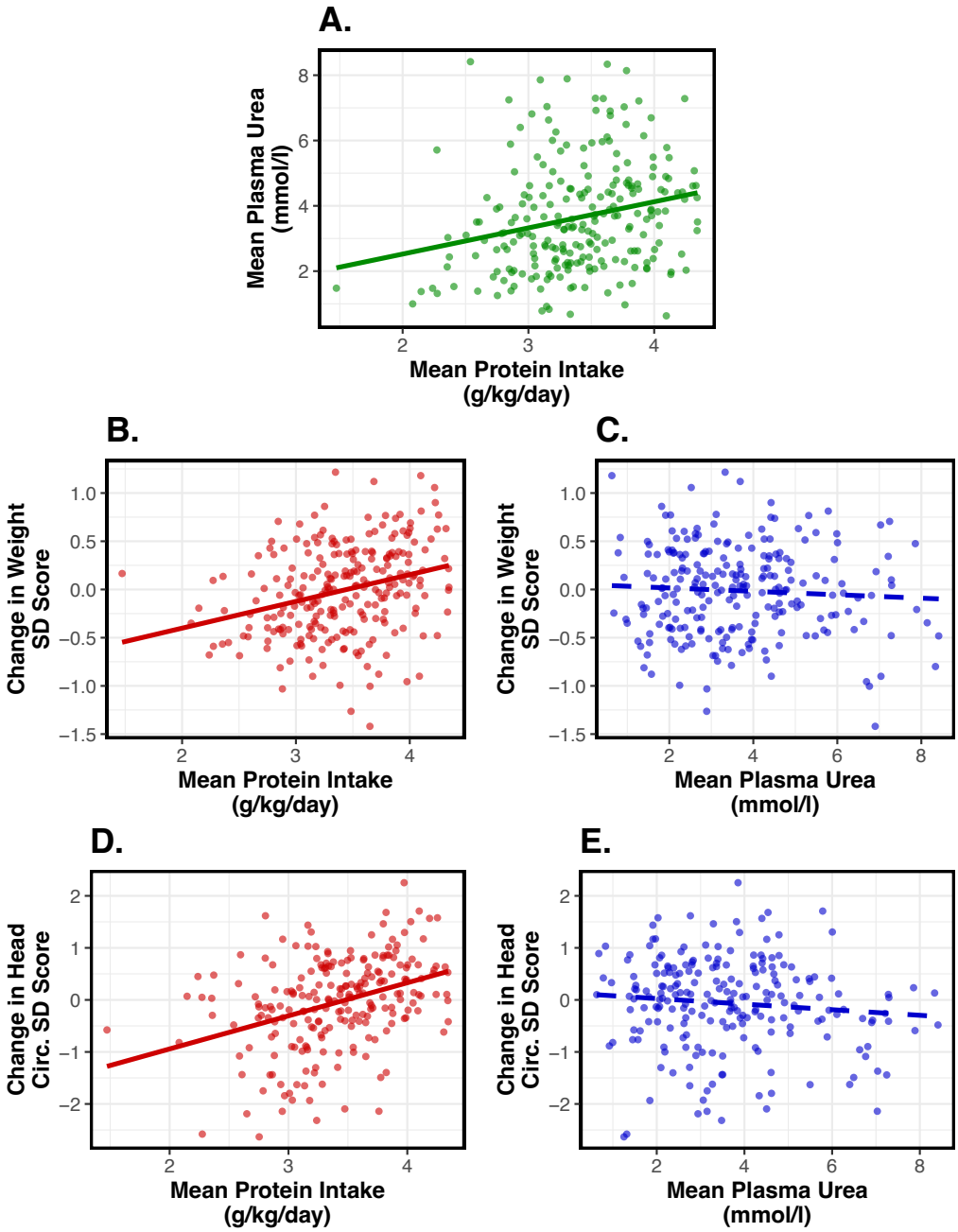


Figure 7-7. Unadjusted linear regression from 14 to 42 days postnatal age for A. mean protein intake and mean plasma urea (regression line p-value <0.001); B. mean protein intake and change in weight SD score (p<0.001); C. mean plasma urea and change in weight SD score (p=0.3); D. mean protein intake and change in head circumference SD score (p<0.001); and E. mean plasma urea and change in head circumference SD score (p=0.9).

7.4 Discussion

The work in this chapter demonstrates that plasma urea is positively associated with protein intake during the critical growth period from the third to the seventh week of postnatal life in very preterm infants. This effect is most pronounced during parenteral nutrition, although it is also observed during full enteral feeding. Whilst there is a relationship between protein intake

and plasma urea, it was not possible to identify a threshold beyond which adequate protein intake can be assumed, as suggested by earlier researchers.[89] Testing for associations between mean protein intake and mean plasma urea during the whole time period confirmed these findings, and mean protein intake was positively associated with change in weight SD score during that period, even after adjustment for non-protein energy intake. However, there was no detectable association between mean plasma urea and weight gain. Taken together, these data suggest that monitoring plasma urea may be of limited use when targeting weight gain and head circumference growth.

This study is limited by the number of infants included, especially in the assessments of the impact of protein and plasma urea on indices of growth. Estimated glomerular filtration rate was used to adjust for renal function, but these estimates have only received limited validation in preterm infants[168]. Whilst eGFR went some way to excluding infants who were in frank renal failure, the residual cohort includes infants who were significantly unwell at the time of blood sampling. The inclusion of CRP as a variable in analysis should adjust the results for concurrent inflammation (which may prompt a catabolic reaction), but an alternative approach would have been to exclude infants who were known to be suffering from significant infection or inflammation at the time of blood sampling. Infants who have non-inflammatory illness are more difficult to identify algorithmically and it is unclear to what extent those diseases would interfere with the relationship between protein intake and plasma urea. It is notable that some infants had urea measurements made on more than one occasion during a single day. Whilst this may represent acuity of illness, timing of "daily" bloods just before or just after midnight may also cause this and it is difficult to separate out these causes from one another.

The SPND (and its source databases) treats intravenous amino acid preparations and dietary protein intake as entirely equivalent, but there is evidence to show that this equivalence cannot be safely assumed.[173] Separation of infants into those receiving full parenteral nutrition and full enteral nutrition (as in Table 7.2) helps to decrease the impact of this phenomenon on the findings of this chapter, but it remains a source of imprecision and possible bias in the pooled results.

The study is strengthened by the careful recording of daily nutrient intakes alongside anthropometric measurements (including regular measurement of head circumference in addition to weight), demographic information and laboratory test values.

Amino Acid and Urea Metabolism

The metabolism of dietary protein and intravenous amino acids is complex, as demonstrated in Figure 7-1.Urea is produced by the catabolism of amino acids (along with other nitrogen-

containing compounds), the greatest source being the hepatic processing of glutamate (derived from aminotransferase reactions of other amino acids) in the urea cycle.[174]

The degree to which amino acids in the pool are diverted into catabolism will also influence urea production. In most cases, proteins which are no longer required can be broken down to their constituent amino acids, which are then used to form new proteins. However, any amino acids which cannot be immediately used to build new proteins cannot be stored and must be catabolised, and catabolic states prompted by infection or inflammation may prevent the reintegration of amino acids into protein, causing them to be oxidised with a resultant increased urea production. Furthermore, synthesis of new proteins is a profoundly energy-intensive activity, so growth can only occur when there is sufficient non-protein energy to support protein synthesis.[144] Urea clearance from the blood will also exert a profound effect on the serum urea concentration. Urea is freely filtered at the glomerulus and its excretion is critically dependent on GFR (although there is some absorption and secretion from the renal tubules, meaning that only 40-50% of filtered urea is excreted).[175] GFR is known to be lower in preterm infants than in term-born infants and increases with age in both groups.[171]

In this context, it may not be surprising that there is not a simple relationship between growth and plasma urea. The concentration of urea in the blood is influenced by a complex balance between dietary protein intake (or intravenous amino acid provision), *de novo* synthesis of non-essential amino acids, the amount of protein laid down during growth, turnover of protein, utilisation of amino acids as an energy source and the capability of the kidneys to filter and excrete urea. Therefore, if protein intake is closely matched to amino acid demands, there may be little circulating urea despite appropriate growth. Analysis of covariates in the present study demonstrates that a higher eGFR or greater non-protein energy provision was associated with lower plasma urea values, as expected from analysis of these metabolic pathways. Inflammation (using CRP as a proxy) was associated with elevated urea levels as expected.

Protein Requirements

A fetus receives around 3.5-4g/kg/day amino acid via the placenta during the third trimester.[176] Stable isotope studies have demonstrated fetal protein accretion of around 2g/kg/day during this period, with the remainder of the transferred amino acid (around 1g/kg/day) being obligately wasted or oxidised for energy provision.[60] Studies of nitrogen balance in parenterally fed preterm infants have shown that more than 3g/kg/day amino acid intake is required to achieve protein accretion matching that of a fetus in the third trimester.[177] Trials randomizing infants to differing protein or amino acid intakes have used heterogenous protocols, limiting their suitability for

meta-analysis. However, a recent systematic review identified a trend towards greater weight gain when protein intakes of >3g/kg/day were used.[178]

Previous studies have focused on the relationship between protein intake and serum urea during early parenteral nutrition or when full enteral feeds had been achieved. These studies reached differing conclusions, with Giretti and coworkers[91] concluding that serum urea was profoundly influenced by protein intake during early parenteral nutrition, and Roggero and co-workers[179] finding that such a relationship only existed during full enteral feeding. A recent secondary analysis of a randomized controlled trial also found that actual protein intake was positively associated with plasma urea during enteral feeding.[180] This chapter examines relationships between protein intake and urea during the critical growth phase after initial weight loss, and finds that an association between protein intake and plasma urea exists in response to enteral or parenteral nutrition during this period (albeit with a more pronounced effect during parenteral nutrition).

This study did not detect a linear relationship between plasma urea and change in weight or head circumference SD score. Previous studies have found that altering protein intake in response to blood urea promotes more rapid growth, although such studies were weakened by the higher rates of protein given to the intervention group.[85] As discussed above, when protein intake matches the demand for metabolism and growth, there is likely to be little urea production, as amino acids are incorporated into growing tissues and are not metabolised to release nitrogen as urea. It would be expected that a rise in plasma urea would only be seen when the body enters a catabolic state or when protein intake outstrips amino acid demands for growth and metabolism. Therefore, urea levels may remain low during growth, as long as there is not a gross mismatch between protein intake and amino acid demands.

Conclusions

Plasma urea is influenced by protein intake during the stable growth phase of very preterm infants. This relationship is seen during both enteral and parenteral feeding (although is more pronounced during parenteral nutrition provision). Protein intake is positively associated with more rapid weight gain and head growth. Despite the relationship between plasma urea and protein intake, and the relationship between protein intake and growth, a direct predicting effect of plasma urea on growth has not been shown. Taken together, these data suggest that monitoring of blood urea is likely to provide only limited guidance when seeking to achieve adequate protein intake, and that targeting of plasma urea may not promote more rapid growth.

Chapter 8 Systematic Review of Total Body Water in Preterm Infants

---

This chapter sets out a systematic review performed by me to assess the current understanding of the patterns of total body water in preterm infants. It is adapted from a peer-reviewed paper written by me and published in Archives of Disease in Childhood: Fetal and Neonatal Edition[181] (Appendix 2).

I am grateful to Dr Lisa Brown for acting as a second reviewer.

8.1 Introduction

Postnatal changes in total body water are influenced by a complex and interacting set of factors. During early adaptation there is a weight loss which is mainly mediated by loss of body water (the "preterm contraction of extracellular spaces" or PreCES).[131, 182] However, this is also frequently a period of cumulative nutritional deficit. Therefore, TBW (total body water) changes reflect a composite of early loss of body water, subsequent abnormalities of fluid and electrolyte balance and alterations in body composition. A fuller understanding of normal values and patterns of TBW would inform investigation and clinical management of fluid status, electrolyte management and nutritional approaches.

TBW contributes to the fat free mass (FFM) component of body composition, and measurement of TBW can be used to help derive estimates of FFM (especially after early fluid loss and fluid balance disturbances), using assumptions relating to the water content of lean and fat-containing tissues.[156] Therefore, a greater understanding of TBW patterns may be used as one avenue to explore and monitor elements of the derangements of body composition seen in preterm infants.

It is increasingly recognised that changes in weight are inadequate to understand these abnormalities of body composition, even when taken together with length. Observational data have demonstrated an association between greater early gains in fat-free mass (FFM) (but not fat mass) with improved neurodevelopmental markers at twelve months corrected age. These findings were robust to adjustment for known clinical confounding factors.[36]

This systematic review aims to assess:

1. The normal percentage TBW of the term-born infant at birth
2. The percentage TBW of preterm infants at birth, and

3. The change in TBW as preterm infants grow and mature.

8.2 Methods

This systematic review was prospectively registered with PROSPERO (CRD42019111436). It is reported in line with the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) guidelines.[183]

Eligibility Criteria

The initial search strategy identified studies using a range of methods to assess total body water percentage. Results arising from different methods could not be combined in meta-analysis and so retained studies were limited to those using dilution techniques (deuterium oxide [^2H_2O), doubly labelled water [$^2H_2{}^{18}O$] or antipyrine dilution).

Studies were included if they used dilutional methods to measure total body water in term-born or preterm infants within two weeks of birth. In addition, selected studies were required to report TBW percentage and sufficient information to define gestational age and corrected gestational age of infants at the time of TBW assessment. Studies were excluded if they concerned infants with congenital abnormalities (e.g. congenital cardiac disease) or infants during the postoperative period. Animal studies, case reports, studies not published in English and review articles without primary data were also excluded.

Search Strategy

Structured searches were made of Medline (Ovid MEDLINE and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions), Web of Science Core Collection and EBSCO-CINAHL (January 1946 to January 2020). Ovid MEDLINE search terms are included as a supplementary file for the published paper. Bibliographies of selected papers were reviewed to find additional papers.

Study Selection

Titles and abstracts were downloaded to EndNote X9 reference management software and duplicates removed. Retained titles and abstracts were uploaded to the Rayyan QCRI web application.[184] Titles and abstracts were independently screened and selected by two reviewers (the doctoral candidate and another research fellow), with disagreements resolved by consensus. The full text of selected articles was retrieved and reviewed for inclusion and risk of bias, and data were extracted to a custom data extraction form and compiled into a database.

Risk of Bias Assessment

Risk of bias assessment was carried out according to the Joanna Biggs Institute checklist for Analytical Cross-Sectional Studies[185].

Data Analysis

The meta package (version 4.11-0)[186] for R (version 3.6.1) was used for all analyses. Where data were reported as median and range, mean and standard deviation were estimated for comparison and meta-analysis using methods described by *Wan et al*.[187] Random effects meta-analysis was selected due to anticipated high inter-study heterogeneity (I^2 >75%) and variations in the methodological approaches. Where weight for gestational age groups were listed categorically, author definition of SGA/AGA/LGA status was accepted. Where individual patient data were provided, percentile values were calculated from UK-WHO data[27] (with infants described only as 'term' treated as being born at 40 weeks of gestation). Subgroup differences were tested with the Q test for heterogeneity. Random effects meta-regression of the influence of gestational age on TBW percentage at birth of preterm infants was selected due to different methodologies and ranges of gestational ages studied. Narrative synthesis and graphical comparison were used where meta-analysis or meta-regression was not possible due to the heterogeneity of the data.

8.3 Results

Study Identification

Searches identified 2349 articles after deduplication. Following screening of titles and abstracts, 128 full text articles were retrieved and assessed for eligibility. From this group, 22 studies met inclusion criteria (Figure 8-1). Fifteen papers concerned full term infants[44-48, 188-197] (Table 8-1-A) and 9 concerned preterm infants[189, 197-204] (Table 8-1-B), with two providing information on both term-born and preterm infants[189, 197].

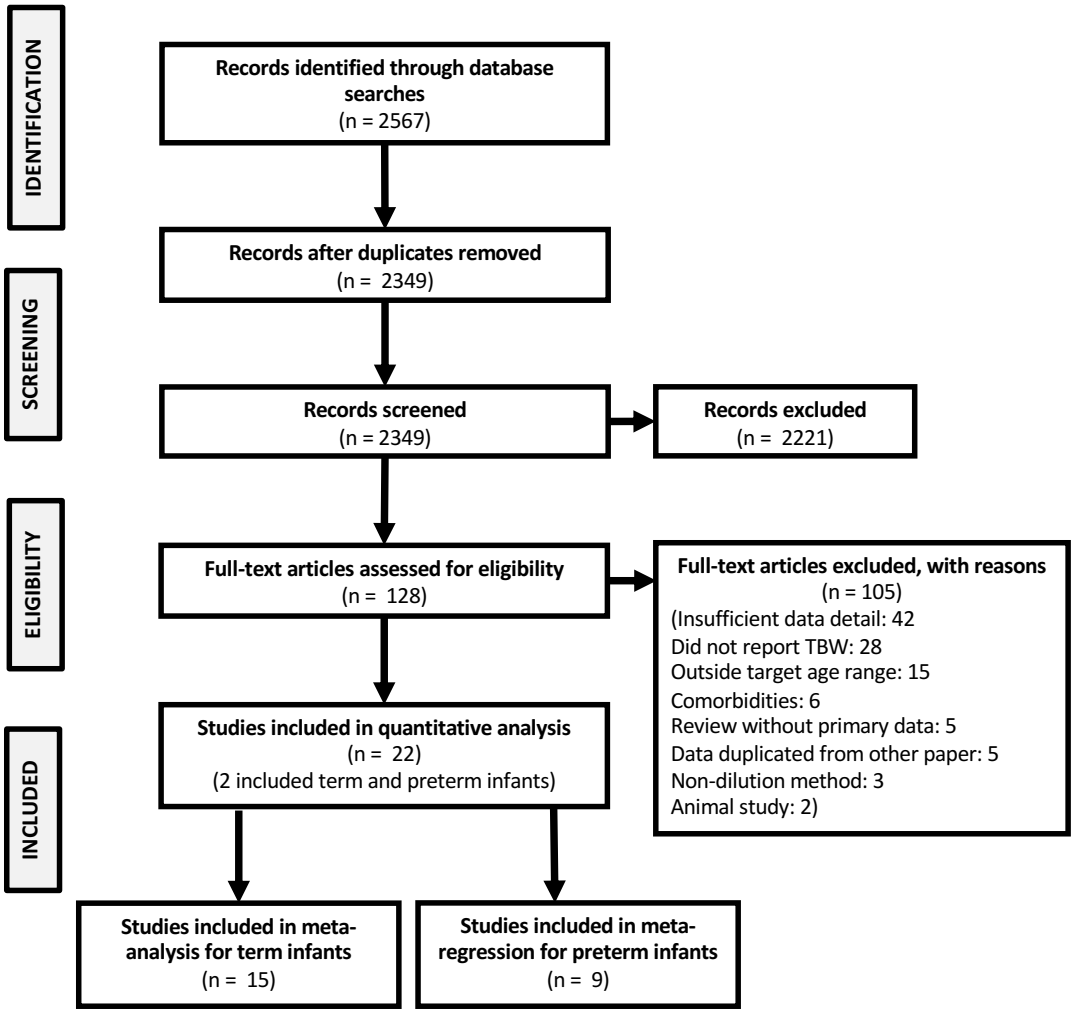


Figure 8-1. PRISMA flow chart of screening and selection of studies, with reasons for exclusion based on full text review.

Table 8-1. Characteristics of studies included in meta-analysis of total body water percentage in A. term-born neonates and B. in meta-regression of total body water percentage in preterm infants. *Insufficient data to calculate or estimate mean and standard deviation. † Study measuring total body water percentage multiple times for each infant. ††Mean and standard deviation estimated from median and range using method described by Wan et al.[187]

A. INFANTS BORN AT FULL TERM

Study	*n*	Method of TBW Assessment	TBW Percentage (mean ±SD)	Birthweight (g) (mean ±SD)	Age at Assessment	Comments
J. B. Andersen (1970)[18]	16	Deuterium oxide dilution	73.93 ± 4.02	3551 ± 622	2-4 hours after birth (9 hours for one infant)	Some mothers received diuretics during labour
Y. W. Brans (1983)[19]	109	Antipyrine space	71.72 ± 8.74	3600 ± 576	0.5 to 12 hours of life	Some infants of diabetic mothers and some macrosomic infants
G. Cassady (1971)[19,20]	26	Antipyrine space	72.95 ± 6.26	3215*	Within 19 hours of birth	Compared normal delivery and Caesarean section
D. B. Cheek, et al. (1961)[21]	11	Antipyrine space	73.18 ± 2.49	Not available	Day 3 of life	Compared diabetic and non-diabetic mothers
D. B. Cheek, et al. (1984)[22]	13	Deuterium oxide dilution	75.70 ± 3.20	2938 ± 563	12 hours of life	Compared preterm, term AGA and term SGA
D. B. Cheek, et al. (1982)[23]	56	Deuterium oxide dilution	75.79 ± 3.48	3931 ± 449	6-24 hours of life	Compared normal delivery and Caesarean section
J. R. Christian, et al. (1956)[24]	17	Antipyrine space	66.21 ± 5.00	3290 ± 330	6-42 hours of life	

W. M. Clapp, et al. (1962)[25]	12	Deuterium oxide dilution	77.40 ± 3.10	range: 2590-4985*	1-6 days of life	Compared term and preterm infants. Found lower total body water in term infants of diabetic mothers.
W. J. Cochran, et al. (1986)[26]	4	Oxygen-18 dilution	79.35 ± 6.38	2335 ± 381	2-3 days of life	Compared total body electrical conductivity with oxygen-18 technique
A. Llanos, et al. (1976)[27]	36	Deuterium oxide dilution	78.50 ± 4.62	2728*	First 30 hours of life	Included AGA and SGA infants
P. J. Offringa, et al. (1990)[28]	13	Deuterium oxide dilution	75.10 ± 5.00	3456 ± 591	First day of life	
E. Olhager, et al. (2003)[29]	9	Doubly labelled water	68.10 ± 4.10	3895 ± 565	4-11 days of life	Compared term and preterm infants
M. Osler (1960)[30]	38	Deuterium oxide dilution	73.37 ± 5.53	3476 ± 523	First day of life	Compared diabetic and non-diabetic mothers
S. C. Singhi, et al. (1995)[31]	55	Tritiated water dilution	73.70 ± 2.60	2836 ± 231	Within 6 hours of birth	Compared term and preterm infants
C. J. Thornton, et al. (1983)[32]	18	Antipyrine space	73.09 ± 2.47	3096 ± 560	Within 12 hours of birth	Compared normocythaemic and polycythaemic infants (no difference)

B. PRETERM INFANTS

Study	*n*	Mean GA (±SD)	Method of TBW Assessment	TBW Percentage (mean ±SD)	Birthweight (g) (mean ±SD)	Age at Assessment	Comments
Baarsma et al. (1992)[33] †	8	31.63 ± 3.66	Deuterium oxide dilution	83.24 ± 4.74	1683 ± 584	2-7 days (one infant at 15 days)	
Bauer et al. (1991)[34] †	8	28.00 ± 1.39††	Deuterium oxide dilution	79.20 ± 11.40	1057 ± 213	First 12 hours of life	
Bhatia et al. (1988)[35]	17	34.60 ± 1.90	Deuterium oxide dilution	74.23 ± 4.32	1990 ± 82	First 7 days of life	
Cheek et al. (1984)[22] †	5	32.50 ± 1.74††	Deuterium oxide dilution	81.90 ± 2.80	2012 ± 397	12 hours of life	Compared preterm, term AGA and term SGA
Hartnoll et al. (2000)[36]	42	27.52 ± 1.01††	Oxygen-18 dilution	83.39 ± 7.75††	1011 ± 238††	Within 18 hours of birth	Compared AGA and SGA infants
Heimler et al. (1993)[37] †	14	30.70 ± 2.40	Deuterium oxide dilution	86.60 ± 6.10	1473 ± 342	Day 1 of life	
Raghavan et al. (1988)[38]	18	27.22 ± 2.24	Oxygen-18 dilution	90.65 ± 2.82	838 ± 161	1-4 days	Validation study for bioelectrical impedance
Singhi et al. (1995)a[39] †	23	33.71 ± 1.00††	Tritiated water dilution	77.34 ± 16.80	1902 ± 242	Within 6 hours of birth	Tracked changes in TBW
Singhi et al (1995)b[31]	44	34.00 ± 1.57††	Tritiated water dilution	77.70 ± 2.60	1978 ± 412	Within 6 hours of birth	Compared term and preterm infants

Total Body Water Percentage of Newborn Term Infants

Fifteen studies assessed TBW in a total of 433 newborn term infants (Figure 8-3-A). The estimated total body water percentage was 73.8% (95% confidence interval 72.47-75.06%, I^2 = 90%). The antipyrine method yielded significantly lower estimates of total body water than the deuterium oxide method (random effects Q test for subgroup differences: $p<0.01$). There were also significant differences between the total body water estimates for SGA infants (78.3%), AGA infants (73.9%) and LGA infants (68.7%) ($p<0.01$) (Figure 8-3-B-C)

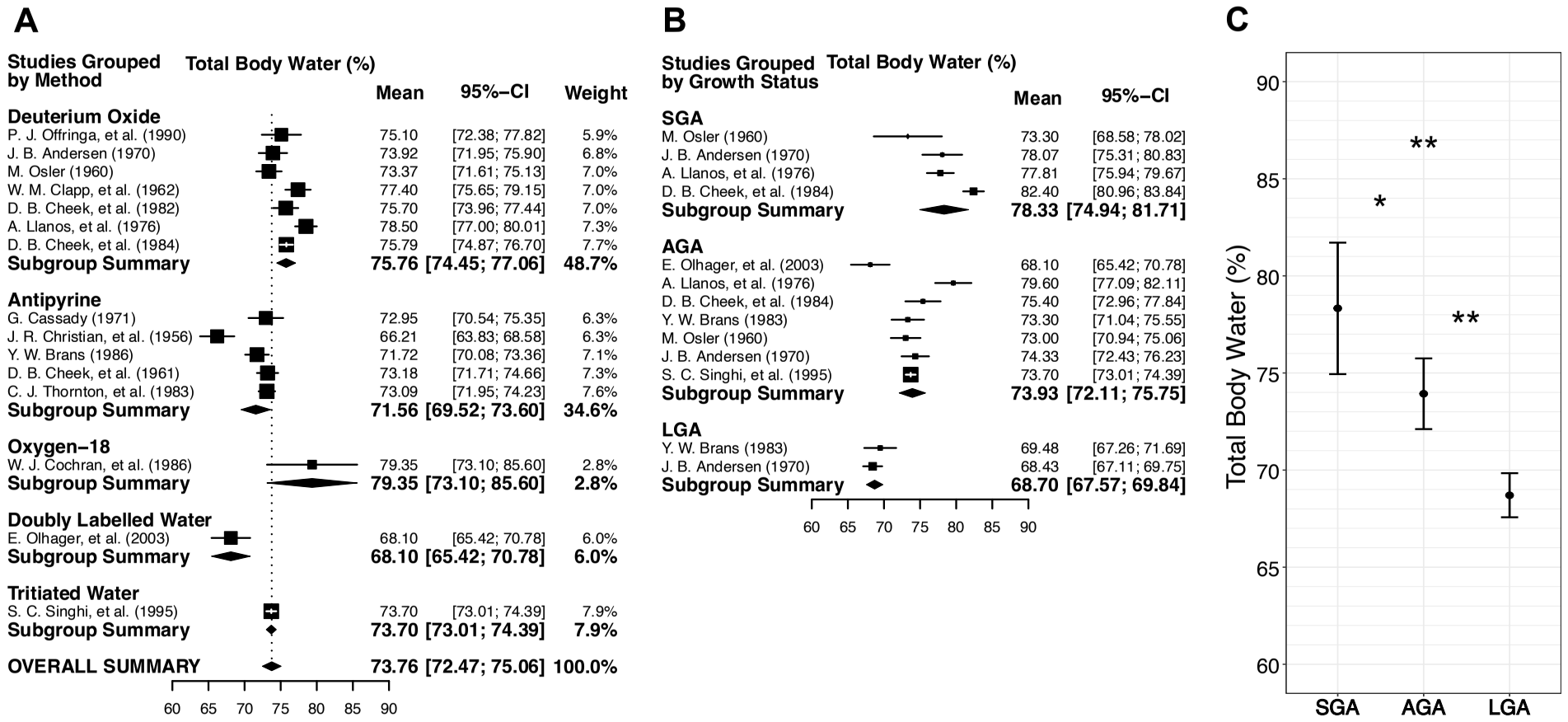


Figure 8-2. Total body water percentage in newborn full term infants. A. Forest plot with methods of TBW assessment as subgroups. B. Forest plot with weight for gestational age subgroups (overall test for subgroup differences: p<0.01) and C. Mean total body water percentage (±95% confidence interval). SGA, small for gestational age; AGA, appropriate for gestational age; LGA, large for gestational age. *p<0.05, **p<0.01.

Total Body Water Percentage of Preterm Infants at Birth

Mixed effects meta-regression of nine studies measuring total body water in 179 preterm newborns identified a progressive decline in TBW from 85-90% between 26 and 28 weeks completed gestational age, to around 75% at 36 weeks gestational age. The estimated decline in percentage TBW was 1.44% per week (95% confidence interval 0.63-2.24%, p<0.001) (regression equation where $y = TBW$ and $x = gestational\ age\ in\ weeks$: $y = 127 - 1.45x$) (Figure 8-4).

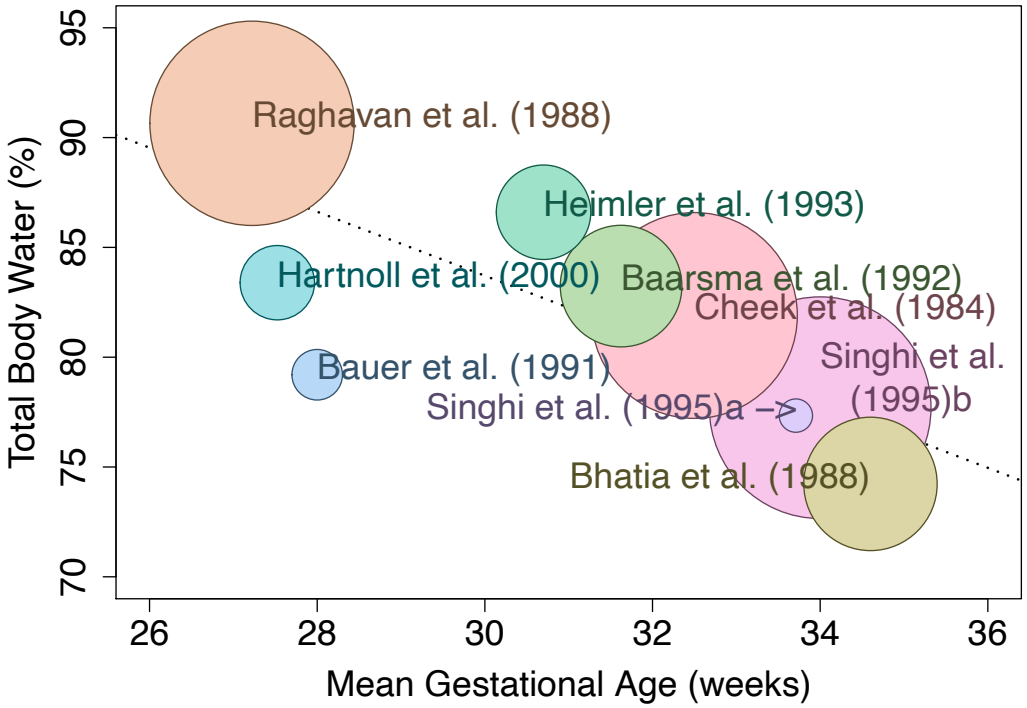


Figure 8-3. Bubble plot with fitted meta-regression line of total body water percentage at birth and mean gestational age in infants included in each study. Circles are sized according to precision of the estimate, larger circles indicating greater precision. Regression equation: y=127-1.45x.

Changes in Total Body Water as Preterm Infants Grow

Five studies measured TBW multiple times for each preterm infant (marked with † in Table 8-1).[189, 198, 199, 202, 204] These studies assessed 56 infants at heterogenous gestational ages and the repeated measurements were taken on different days of life (Figure 8-5-A). None of the studies gave detailed information on nutritional intake. *Baarsma et al*.[198] reported two measurements each for two infants, only one of whom had a measurement taken shortly after birth. The remaining four studies did not report individual patient data but summarised the gestation at birth and the day of life when measurements were taken. *Bauer et al*.[199] and *Singhi et al*.[204] focused on early body composition changes as they relate to initial weight loss in preterm infants. *Cheek et al*.[189] and *Heimler et al*.[202] measured body composition at weekly intervals. *Cheek et*

al.,[189] *Heimler et al*.[202] and *Singhi et al*.[204] all identified that TBW percentage fell over time, with the one patient for whom early assessment was made in *Baarsma et al*.[198] also exhibited a falling total body water percentage. In contrast, *Bauer et al*.[199] found an initial drop in total body water followed by a rise almost back to the first measurement.

Data from these five papers were too heterogenous to perform formal meta-analysis. Data were adjusted to plot the change in total body water against the day of life. A linear regression model, weighted for the number of infants assessed in each study, confirmed that percentage TBW fell over time (1.31% per week; 95% CI 0.53-2.09%; p<0.01) (Figure 8-5-B).

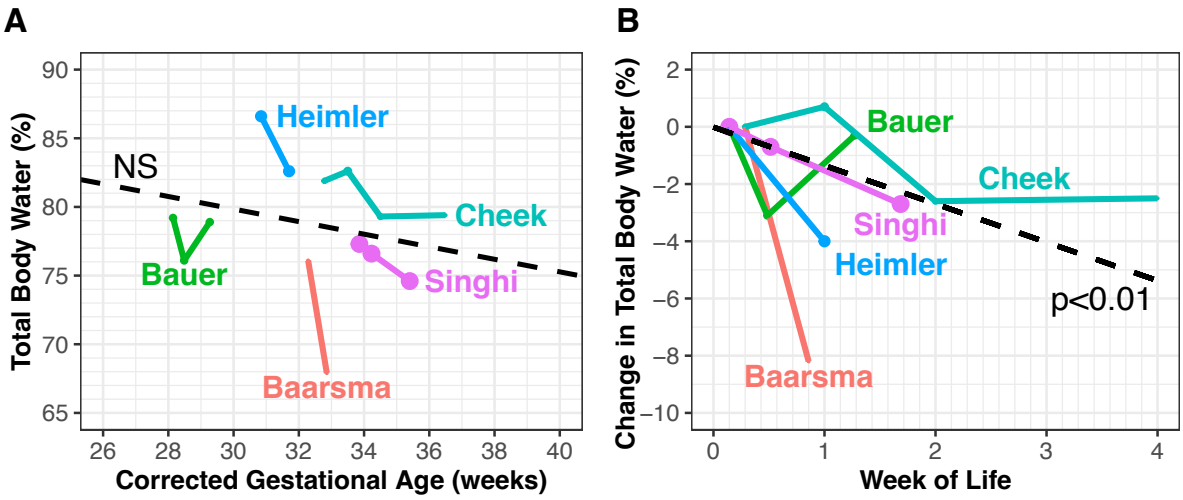


Figure 8-4. A. Total body water percentage in 58 preterm infants in 5 studies where repeated measurements of total body water were taken during their growth and maturation. Sizes of points indicate number of infants assessed. Points are plotted at the mean gestational age for infants in each study. Dashed line: linear regression line weighted for study size (slope value not significant). B. Change in total body water percentage in preterm infants after birth, grouped by study. Sizes of points indicate number of infants assessed. Dashed line: linear regression line weighted for study size (y=-0.13-1.31x).

8.4 Discussion

This systematic review and meta-analysis identified the normal percentage TBW for term-born infants. Meta-regression of studies examining percentage TBW in newborn preterm infants demonstrated a negative correlation between gestation at birth and percentage TBW. Studies with repeated measurements of preterm infants were too heterogenous for formal meta-analysis but most studies identified a fall in total body water over time.

Studies included in this systematic review were mostly performed at least twenty years ago with many being published forty to sixty years ago. This is likely to limit the generalisability of these findings to modern cohorts of preterm infants. Advances in neonatal care have led to a rapidly changing phenotype of preterm infants during this period. Despite uncertainty surrounding the

impact of respiratory distress syndrome (RDS) on early fluid loss[205], the surfactant era has led to significantly improved early morbidity in very preterm infants, with severe RDS now mostly limited to those born near the limits of viability. Antenatal corticosteroids have a profound influence on early fluid shifts[206] and their widespread use is likely to have significantly altered the "normal" pattern of TBW changes since publication of most included studies.

Term Infants

A strength of this review is that it includes over 400 term-born infants, whereas individual studies have rarely recruited more than 100 infants. Meta-analysis of the percentage TBW of term-born infants is limited by heterogeneity between studies. This may be due to real differences in studied populations (including demographic and nutritional differences between mothers, and differences in antenatal care) or due to differences in study protocols and methods. Furthermore, there was some variability in the timing of TBW assessment. In most cases, measurements were made within the first two days of life, but it has previously been noted that there can be rapid fluid shifts even within this limited time period.[182] Subgroup analysis confirmed that measured values of TBW are influenced by the use of antipyrine compared to deuterium oxide. In addition, there were significant differences in TBW identified between SGA, AGA and LGA infants.

Preterm Infants at Birth

This review is limited by the presence of relatively few primary research articles measuring TBW in preterm newborn infants (with only 179 subjects in this section of the review). Furthermore, data were generally reported only as the mean or median gestation at birth. Therefore, meta-analysis was performed taking this mean as the independent variable. This reduced the power and precision of the meta-regression to identify the effect of gestation at birth. This shortcoming emphasises the importance of published individual patient data.[207] However, most included studies were more than two decades old and it was unfeasible to obtain individual patient data. Like the data for term infants, there was heterogeneity in methods, antenatal care and timing of TBW assessment.

Preterm Infants During Growth

There was a general paucity of data tracking TBW of preterm infants during their growth and maturation, with only 59 patients considered in this section of the meta-analysis. In addition to the sources of heterogeneity listed above, there was also very limited information on the postnatal care of these infants. It was not possible to perform formal meta-analysis of these results, although narrative synthesis identified that most papers reported a fall in percentage TBW over time. A weighted linear regression model discounting gestational age and taking week

of life as the independent variable confirmed this finding. Of note, the fall in TBW of approximately 1.3% per week was just under the 1.4% figure found in the meta-regression of preterm newborn infants.

Implications for Practice

Bearing in mind international guidance[23, 208] that the body composition of preterm infants at term-equivalent age should mimic that of the term-born newborn infant, the value of TBW in term-born infants identified in this study may act as a guide. However, it should be remembered that TBW reflects only one element of body composition and that derangement of fluid and electrolyte imbalance may render this comparison unreliable. This study also defines the expected TBW value for newborn preterm infants at birth in relation to their gestation. It may be expected that preterm infants would be expected to track along this trend if they are to reach a normal percentage TBW at term-corrected age.

The small amount of data pertaining to repeated measurements of TBW in preterm infants during growth makes it difficult for current *ex utero* body water changes to be assessed. Even if more such data were available, it is likely that widespread differences in medical and nutritional management of preterm infants (as well as differences in antenatal care) would frustrate attempts to meaningfully combine data to provide an overview of the *ex utero* changes occurring in the preterm infant during growth. Significant changes in body water soon after birth may mean that preterm infants deviate from the trend in TBW identified in cross-sectional assessments of newborn infants at different gestations.

Routine and reliable measurement of body composition in clinical practice currently remains out of reach. Non-invasive measures of body composition, such as bioelectrical impedance have been used but have stalled at the validation stage.[209] Our group have shown that limb circumference measurements are feasible in preterm infants and demonstrate a predictable pattern of growth,[52] but further work is needed to assess their clinical utility in reflecting changes in body composition.

Conclusion

This systematic review and meta-analysis identified potential normal values for TBW in newborn preterm and term-born infants. These findings may be useful for clinicians trying to assess changes in fluid status, electrolyte balance, growth and body composition after birth in preterm infants. Future work should focus on tracking the body composition of preterm infants as they grow and should be contextualised by detailed monitoring of fluid management, nutritional intake and growth. Reliable and validated bedside measures of body composition in preterm infants will be required before real-time targeting of body composition can be integrated into clinical

practice. In particular, the paucity of longitudinal tracking data concerning total body water in preterm infants demonstrates a gap in current knowledge.

Chapter 9 Assessment of Total Body Water

This chapter sets out the results of a programme of measuring total body water by deuterium dilution in Southampton.

The protocol for this chapter was written by me, building on preparatory work carried out by my supervisor, Mark Johnson, prior to my starting this doctoral programme. Recruitment, administration of deuterium oxide and collection of urine samples was led by me with the assistance of Pip Crowley, research nurse. This element of my project was impacted by the COVID-19 pandemic in that recruitment was halted whilst research nurses were recalled to clinical practice, and the laboratory which processed the samples was closed for a prolonged period.

9.1 Introduction

As set out in the systematic review above, there are significant differences in body composition between preterm infants and term infants at the time of birth. Furthermore, differences in body composition persist during these infants' initial hospital stay so that preterm infants have a lower body weight but with preserved fat mass at discharge compared to their term-born counterparts. Their body fat percentage is therefore higher and their lean mass percentage lower than their term counterparts. The introduction chapter to this thesis sets out the current understanding of the importance of body composition in preterm infants.

Measuring Total Body Water

Considering the paucity of data concerning longitudinal changes in total body water in preterm infants found in the systematic review, this doctoral project aimed to track total body water in growing preterm infants. Nine infants were included in a pilot phase of this arm of the project, for which results are available. Whilst more infants have undergone sample collection since the interim analysis, that analysis has not been performed by our external collaborator due to the COVID-19 pandemic and by problems with their equipment.

Method Development

Conceptually, the method for assessing total body water by stable isotope dilution is simple: an exogenous substance is given to the subject, the substance freely mixes into the entire pool of body water and is then excreted unchanged in the urine, with the concentration in the urine being used to derive a value for the total body water volume. In addition, the substance used must be safe and pharmacologically inert, and would ideally be inexpensive to buy and to detect in the urine. In practice, deuterium oxide is the currently preferred substance for this process, meeting all the obligatory requirements but representing a significant cost both for acquisition and analysis.

An appropriate preparation of deuterium oxide was prepared by a specialist pharmacy prior to the start of this doctoral project. The method used for deuterium oxide analysis of total body water was based on guidance by the International Atomic Energy Agency (IAEA).[49, 210]

9.2 Methods

Deuterium Oxide Delivery and Specimen Collection

Measurements of total body water were made up to once every three weeks. A baseline urine sample was collected prior to administration of the deuterium oxide solution (by means of absorbing urine into a ball of cotton wool placed in the nappy and then squeezing using a syringe into a sample container – see Appendix 10). A dose of 1ml/kg bodyweight of deuterium oxide 70mg/ml in 0.9% sodium chloride was drawn up from a vial containing 5ml of the solution into a syringe, labelled and capped with a plastic bung. This syringe was weighed on calibrated scales and the weight recorded to the nearest 10 micrograms. The dose was then delivered to the infant by:

1. Slow intravenous injection followed by a flush of 1ml 0.9% sodium chloride *or*
2. By nasogastric or orogastric tube preceding a planned feed *or*
3. Orally by mixing with 10ml milk and delivering through a bottle teat.

The bung was then replaced and the syringe reweighed so that the difference between the weights would reflect the precise mass of the deuterium oxide solution given. Excess deuterium oxide solution was retained.

Specimen Management and Analysis

Urine samples were collected at 6-10 hours after administration and then daily for five days. During the pilot phase, samples were collected for seven days, but interim analysis demonstrated

that later samples were not informative and therefore they were not carried out during subsequent phases.

Urine samples were refrigerated for up to seven days on the neonatal unit before being transferred to a laboratory where they were centrifuged. The pellet was discarded and supernatant was retained and frozen at -80°C.

Samples were batched and an external collaborator (Prof Simon Eaton, University College London) carried out mass spectrometry to determine the concentration of deuterium oxide in each urine sample and each of the retained deuterium oxide solutions.

Statistical Analysis

Due to the high water turnover of infants and the unpredictability of voiding interval, the back-extrapolation method is preferred.[49] In compliance with International Union of Pure and Applied Chemistry guidance,[211] the relative abundance of deuterium (^{2}H) is reported by the laboratory in relation to two internationally accepted standards (the Vienna Standard Mean Ocean Water sample and the Standard Light Antarctic Precipitation sample) producing a 'δ 2H-V-SMOW/SLAP' (or δ^2H) value. This value must then be converted to parts per million ([^{2}H]) using the following formula:

$$[^2H] = \frac{\delta^2H + 1000}{6.420135}$$

The enrichment of ^{2}H in response to deuterium dosing can then be calculated by subtracting the baseline concentration from each subsequent urine sample. The natural logarithm of this enrichment value can be plotted against the elapsed time following dosing, a linear regression line can be applied to these data and the y-intercept of the line can be used to estimate the natural logarithm of the theoretical concentration at the moment of dosing. A graphical example of this using data from the pilot phase is given as Figure 9-1. This value can then be exponentiated to determine the enrichment value at the time of dosing (e.g. 66.8ppm in the example given in Figure 9-1).

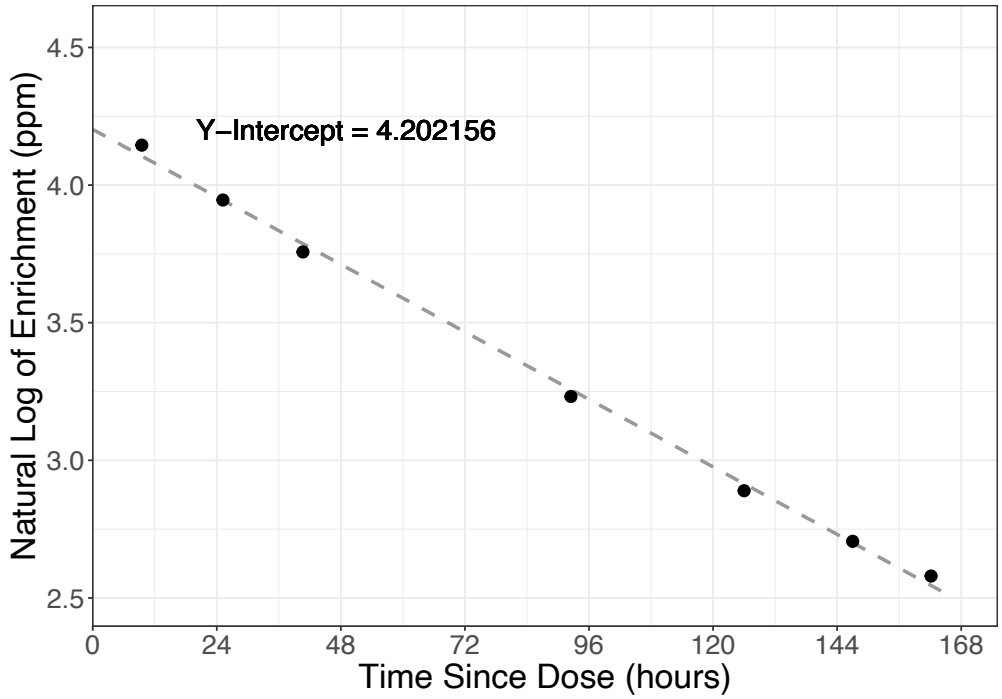


Figure 9-1. Example of back-regression method used to estimate nominal ^{2}H enrichment at the time of dosing.

The following formula can then be used to calculate the total body water at the time of ^{2}H dosing:

$$TBW(kg) = \frac{WA}{a} \times \frac{\Delta DD}{\Delta BW} \times \frac{1}{1000 \times 1.041}$$

Where:

W = weight of water used to dilute dose sample for analysis (g)

A = weight of dose taken by subject (g)

a = weight of dose sample in dilution used for sample analysis (g)

ΔDD = enrichment of ^{2}H in the diluted dose compared to the water used for dilution (ppm)

ΔBW = enrichment of ^{2}H in body water dose (i.e. the final exponentiated value from the back extrapolation calculations above)

1.041 = correction factor for non-aqueous exchange of hydrogen ions

Body water percentage can then be calculated by comparing this value to the infant's weight in kg.

9.3 Results

Table 9-1 sets out total body water percentage measurements for nine infants during the first phase of total body water assessments for this doctoral project. The final column of the table demonstrates that the derived total body water mass was greater than the total weight of the infant at the time of assessment for some measurements (i.e. the total body water percentage is

greater than 100%). This shows that there were significant problems with the methods (or their application) which are discussed below.

Table 9-1. Total body water assessments for nine infants.

Patient	Gestational Age at Birth (weeks)	Corrected Gestational Age at Assessment (weeks)	Day of Life	Weight (g)	TBW Percentage
1	24.86	25.57	5	710	90.32
		28.29	24	930	83.60
2	29.00	30.00	7	900	71.31
3	26.71	27.86	8	650	76.04
		31.00	30	900	65.16
		34.14	52	1540	78.24
		38.00	79	1775	76.82
4	28.71	29.57	6	920	81.49
		32.14	24	1340	83.36
5	28.71	29.57	6	920	113.46
		32.14	24	1340	108.20
6	29.00	29.57	4	1040	82.85
	29.00	35.71	47	1760	100.17
7	25.71	26.86	8	755	92.09
		29.71	28	1020	84.56
		32.86	50	1540	132.36
8	28.29	29.43	8	1445	59.83
9	28.43	29.00	4	810	155.51
		32.14	26	1350	89.56

9.4 Discussion

Despite removal of extreme outliers from back extrapolation calculations, several clearly erroneous total body water measurements remain. There is no systematic difference in known variables between those assessments which yielded plausible results and those with implausible results. For example, some implausible values were derived from oral dosing and others from intravenous dosing. Baseline urine ^{2}H concentration varied little between infants. Furthermore, progressive clearance of ^{2}H from the urine proceeded linearly after logarithmic transformation as expected. Figure 27 shows the back extrapolation for the first assessment of patient 9 – the most egregiously faulty total body water percentage value. There were no outliers and the adjusted R squared value for the linear regression model is 0.994, reflecting the predictable elimination of ^{2}H

in the urine. However, the y-intercept is so low that the total body water percentage is predicted at 156% of total body mass.

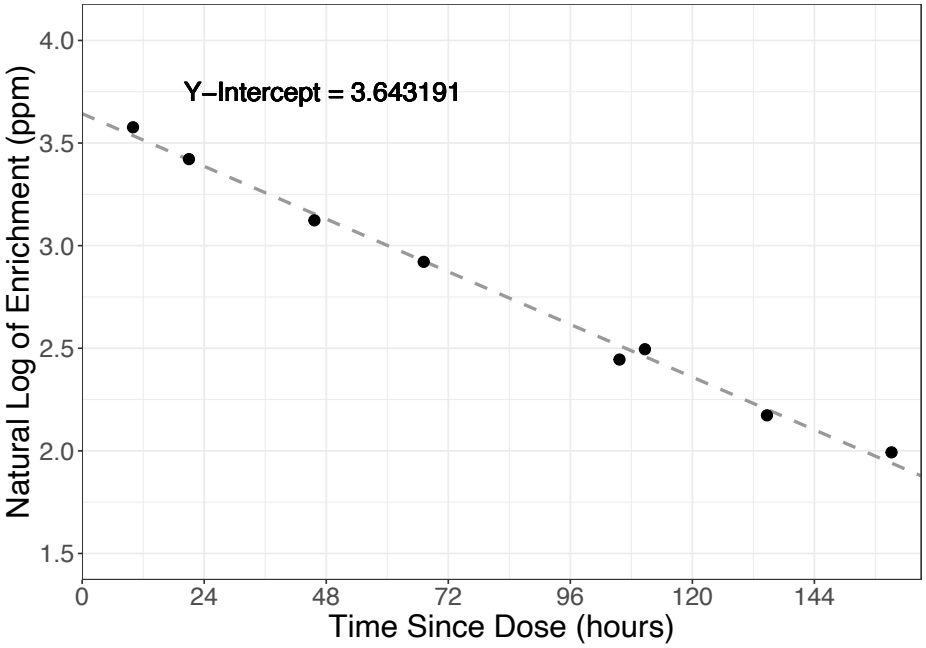


Figure 9-2. Example of back-regression used to estimate nominal ^{2}H enrichment at the time of dosing for infant 9, for whom a total body water percentage of 156% was calculated.

In the absence of clear errors in the data, the most likely explanation for these results is that administration of the dose was incomplete or that the dose failed to equilibrate with the total body water pool.

During the second phase of this work, extra care was taken to ensure that the entire measured dose was delivered to each infant. In addition, any problems with dose administration or with subsequent vomiting have been recorded. I have personally supervised each delivery of a dose of deuterium oxide to an infant.

Due to the COVID-19 pandemic and unexpected maintenance issues with the mass spectrometer used by our collaborator, the results of the second phase of body water measurement have not been received at the time of submitting this thesis. The data will be presented at conferences when they have been received. The current data are not suitable to attempt to judge whether limb circumference measurements provide a reliable marker of body composition in preterm infants.

The results of this work highlight the difficulties of measuring body composition in preterm infants. Even if the deuterium dilution method could be adapted to become more reliable, deriving the total body water content of an infant provides only limited information about that infant's body composition, with estimation of lean mass and fat mass relying on the application of formulae which may not be reliable in this patient group. True assessment of lean and fat mass

would be better judged by performing MRI scans (either of the whole body or of a representative part). New MRI technologies which make MRI more feasible for unstable infants and which can be located on neonatal units may provide a more reliable and sustainable opportunity to directly measure body composition in these infants as they grow.

Chapter 10 Genomic Analysis

Chapter Summary: *This chapter introduces the applications of genomic testing in newborns and preterm infants. There is then a description of the aims of the exome analysis of a subset of 13 infants from the local Southampton cohort, the methods employed, a summary of the results and an exploration of the way these results will inform future work.*

I applied for ethical approval, coordinated collection of blood samples and interpreted the identified variants. Recruitment of infants into genomic testing was performed by me and by Research Nurses Philippa Crowley and Sarah McKay. DNA extractions and plating for external sequencing provision (Novogene Ltd.) were performed by Nikki Graham with my assistance. Quality control and pre-processing of genomic data were carried out by me, mostly using scripts written by Dr Guo Cheng and Dr Eleanor Seaby.

10.1 Background

Since the publication of the first draft of the human genome in 2003, techniques for the acquisition and analysis of human genetic material have advanced rapidly and have become progressively cheaper and more readily available. This background section explores the technology of next generation sequencing and sets out the current state of implementation of genomic approaches in clinical care and research into neonates, with specific emphasis on nutritional investigations. This understanding underpinned the aims and approaches used to integrate genomic information into the exploration of the influences on the growth of preterm infants in this doctoral project.

10.1.1 Fundamentals of Next Generation Sequencing (NGS)

Next generation sequencing describes high-throughput approaches to generating large amounts of genetic data, utilising technologies which record the sequence of individual DNA molecules. Whole genome sequencing provides information on both the coding and non-coding regions of the genome. Whole exome sequencing provides a more targeted approach, selecting only coding regions (around1-2% of the total genome) prior to sequencing. This method typically affords deeper sequencing in the coding regions (i.e. more individual reads informing each base pair call) at the expense of not providing coverage of the non-sequencing regions. Whilst most mutations

causing human disease directly impact coding regions, variation in non-coding regions can also affect gene activity and protein production, for example by influencing promoter regions and regulatory elements. Such mutations are missed by exome sequencing.[212] Nevertheless, exome sequencing has a number of benefits. It provides very detailed and reliable information about coding regions, which are much more likely to be implicated in disease or physiological differences than non-coding regions. In addition, the data files which result from this analysis are much smaller than those for whole genomes (generally less than one tenth the size),[213] reducing the burden of storage and analysis. The lower cost of whole exome sequencing also provides an opportunity to analyse more subjects within a set costing envelope.

10.1.2 NGS Technologies

Short-read Sequencing

Next generation sequencing starts with exome library preparation. This involves the fragmentation of DNA into short molecules (250-500 base pairs)[214], clonal amplification, parallel sequencing of these small molecules and reconstruction of genomic regions from the resultant small reads.[215] This process is known as short-read sequencing.

In practice, several methods have been employed to implement these concepts. Library preparation can include selection of DNA fragments for regions of interest using complementary strands (for example, to select all coding regions in exome analysis), followed by addition of specialised adapters to render the fragments compatible with the sequencing technology employed. Labels can also be added to samples at this stage, so that sequencing of DNA from multiple subjects can be performed in parallel. The selected fragments must then undergo clonal amplification (i.e. copying) and then sequencing.

Different sequencing methods have been developed, although sequencing by synthesis is the current dominant method for short-read sequencing. In this method, a complementary strand of DNA is synthesised and the integration of each nucleotide is detected. In some implementations (for example Ion Torrent), fragments are sequentially exposed to each of the four different nucleotides (A, T, C and G), and the production of a hydrogen ion is detected whenever a complementary base is integrated. Other technologies (including Illumina) use fluorescent marking of nucleotides, with sequential addition of fluorescent nucleotides detected by optical sensors.

Short-read sequencing generates 20-30 million reads, requiring careful data handing and quality control. Analysis of these data requires skilled application of a defined analysis pipeline, running on high powered computing assets. These steps are explored in the methods subsection below.

Novogene Sequencing Protocols

This project used Novogene as an external contractor to sequence exomes for enrolled subjects. The following section explains the various steps involved in Novogene's workflow.

Library Preparation

Novogene uses physical fragmentation to generate short read molecules. They then attach adaptors to the fragments. These adaptors contain three elements: the binding site for the sequencing primer which attaches to the fragment of DNA, an indexing barcode, and a sequence complementary to the oligonucleotides which are attached to the flow cell (described below). Indexing barcodes allow multiple samples to be run during the same analysis, with the resultant data separated using the indexing barcodes (known as multiplexing). SureSelect Human All Exon V6 probes (Agilent Technologies) were used in library preparation to enrich the library for the 20,000 human genes over an approximately 60 megabase target region.[216] This library preparation step allows for the selective sequencing of known coding regions of the genome (i.e. the exome).

Illumina Sequencing

After library preparation, DNA fragments are loaded onto flow cells. Illumina flow cells are glass slides containing a number of channels. Each channel contains millions of nanowells – tiny wells with oligonucleotides at their base, designed to capture prepared DNA fragments. As fragments are washed over the solid surface of the flow cell, the complementary sequences on their adaptors bind to oligonucleotides in the nanowells and they are captured. Capture technology is designed to maximise the chance of only a single fragment being captured in each flow cell, allowing for easier cluster analysis during later stages.

After capture of DNA fragments, bridge amplification is used to form clusters of identical DNA fragments. In this process, polymerases move along the DNA strands, creating a complementary strand. The original strand is then washed away, leaving only the complementary strand in place. An adaptor sequence is left at the top end of the fragment and this bends to attach to another flow cell surface oligonucleotide. The resulting double-stranded DNA is denatured, leaving two complementary strands attached to the flow cell. This process can be repeated many times to form clusters of identical DNA fragments each nanowell, in a process known as clonal amplification. All reverse strands are then washed off the cell surface so that only forward strands are left in the clusters.

After clonally amplified clusters have been formed in the flow cell nanowells, sequencing by synthesis can begin. The flow cell is washed with dNTPs (deoxynucleoside triphosphates, i.e. DNA bases) which have been labelled with a fluorescent dye (each different nucleoside having been

labelled with a combination of two dyes so that four options are available: red, green, yellow-appearing with both red and green labels, and unlabelled). The dye molecule prevents further polymerisation of the DNA strand so that only one base is added to each captured DNA strand. Images are then taken of the flow cell and the fluorescence of each cluster can be interpreted as having incorporated a dNTP labelled with the corresponding colour. After these images have been captured, the fluorescent tag is washed away, preparing the strand for addition of a further dNTP. This cycle is repeated many times, with the sequence of fluorescence colours of each cluster interpreted as the genetic code of the strand of DNA. Whilst this process generally produces reliable results, base reads typically become less reliable as the position along the read increases. This is primarily due to phasing, the failure to remove the fluorescent marker from the previous base or the failure to add a fluorescent marker to the new base. The read on this strand is then out of phase with the other strands in the cluster, a phenomenon which increases with the length of read.[217] Illumina technologies use paired-end sequencing, in which both ends of a strand of DNA are read in sequence (around 150bp from each end), improving the coverage and reducing certain errors within the reads.

Data Provision

Short read data are then stored as FASTQ files. The first line of a FASTQ file contains information on the equipment used, the cluster position within the flow cell and the indexing barcode. The subsequent lines provide the sequence of bases identified and a character representing the Phred quality score for each base, indicating the risk of the base having been called incorrectly.

Long-read Sequencing

In recent years, technologies have emerged allowing for the sequencing of much longer fragments of DNA (often more than 10,000 bp). In these approaches, complementary strands of DNA are synthesised within pores and the identity of each incorporated nucleotide is defined by a fluorescent signature (in SMRT [Single Molecule, Real-Time] sequencing) or by a characteristic disruption of an ion current (Oxford Nanopore). These approaches avoid the need for amplification of DNA samples, can provide more structural information and typically have faster turnaround times that short-read methods. Error rates are generally higher, although error rates are improving over time. Long reads capture hard-to-sequence sections which may be poorly mapped by short read techniques.

Joint Calling and Annotation

Joint calling is a process which leverages pooled information from a number of samples to identify variants in individuals with greater confidence, and to define the certainty with which the variant is called based on the base pair calls of the other individuals in the joint calling cohort.

Seqr is a web-based tool which presents a user interface to filter variants.[218] Annotations are added to the called variants to assist with interpreting the clinical or research significance of the variants discovered. These annotations include the reported clinical implications of the mutation and various scores aiming to predict the deleteriousness of the variants.

ClinVar Clinical Significance Score

The National Institute for Health ClinVar database collates reports about the relationships between genetic variants and disease states.[219] The database defines a clinical significance for each of the variants it contains. In the case of Mendelian diseases with high penetrance, these classifications range from “benign” to “pathogenic”, with degrees of uncertainty in between as set out in an American consensus document.[220] Each classification is also defined by a gold star rating, with significance scores based on multiple submissions receiving more stars than those with a single source. The reports contributing to the ClinVar significance score can be from diverse sources, ranging from population data (e.g. a variant being marked as likely benign if the population incidence is much more frequent than the disease which it may be associated with) to computational data (e.g. evidence from modelling that the variant will cause the same amino acid change as a confirmed cause of disease).

CADD Score

The Combined Annotation–Dependent Depletion (CADD) score is a method for predicting the deleteriousness of a variation in the genome.[221] It is based on the observation that random mutation will introduce variants into the human genome. Those mutations which are advantageous will become more common by natural selection. Neutral mutations will remain present in the population and, over time, will establish a baseline rate. Deleterious mutations will be depleted by natural selection and will therefore be much less common than would be expected for non-pathogenic variants. Scores were generated by comparing the actual mutation rate in a large group of individuals against the expected simulation rate using a model of sequence evolution. The results of this analysis were compared by the CADD designers to a broad range of annotations to the variants, and machine learning was used to build a model to classify variants by their likely deliriousness. This deleteriousness score was scaled to a range from 1 to 99. The producers of the CADD score recommend a cut-off between 10 (meaning that the variant is within the 10% most deleterious substitutions) and 20 (within the 1% most deleterious).

10.1.3 Applications of Genomics in Neonatal Medicine

Interest in the use of genomic approaches in neonatal research and clinical practice has expanded as technologies have become more available. The 100,000 Genomes project in the UK identified that, in probands with a likely genetic condition but for whom normal clinical investigation had failed to identify a cause, 25% could receive a diagnosis by genomic analysis.[222] The following subsections will explore genomic applications relevant to neonatal medicine.

Clinical Diagnosis of Rare Disease and Screening

The findings from the 100,000 genomes project contributed to the rapid integration of genomic approaches into routine clinical investigation for rare diseases in sick neonates, with trio exome analysis being generally favoured (the simultaneous sequencing of the exome of the ill child and his or her parents).[223] Trio exome testing allows analysis of the inheritance pattern of variants including identification of *de novo* mutations. However, trio data may not be available in single parent families, when a parent has died or when a child has been adopted. It may also reveal unexpected non-paternity of the father. Following on from the success of targeted genomic testing, Genomics England (a company owned by a UK government department) has proposed the implementation of routine genomic screening for all infants born in the nation.

The current UK newborn screening programme includes physical examination, hearing testing and biochemical analysis of a bloodspot to detect nine rare inherited diseases of childhood. The inclusion of these screening methods is based on the application of the UK National Screening Committee's published criteria for appraising the viability, effectiveness and appropriateness of a screening programme.[224] In short, this requires that a reliable test should be available for an important condition for which an effective treatment is available.

As set out in Genomics England's vision document for the Newborn Genomes Programme[225], a pilot intervention is planned to assess the genome of 200,000 infants to identify a "set of actionable genetic conditions which may affect their health in early years". Work to identify the list of target diseases, and to define protocols for information governance and management, are currently in progress. If this programme were to progress to universal implementation, a vast amount of genetic information would be available for most infants born in England (with likely expansion to the other home nations). Most of these infants would be apparently well and the use of their genetic information would be subject to intense ethical scrutiny. Furthermore, the use of the term "actionable genetic conditions" opens the door to a much looser application of the normal criteria for screening programme selection, with advocacy groups encouraging the inclusion of genetic diseases for which no effective treatment is available. Whatever the outcome of the Newborn Genomes Programme, the availability of genomic information for infants is likely

to expand very greatly in the coming decade, posing a challenge for clinicians to manage, interpret and communicate the insights this will bring for individual children and families.

Genomic Research in Preterm Infants

Prematurity is the leading cause of death in childhood in developed countries, including the UK.[226] Therefore, significant research effort has been expended in investigating the causes of preterm birth. Preterm birth is known to be significantly heritable[227] and several recent studies have sought to identify genomic markers of risk of preterm delivery, both in the pregnant person and the prematurely born infant. Genome-wide association studies (GWAS) have been fruitful in mothers, identifying an association between preterm delivery and genes involved in uterine development, inflammatory response and other developmental and physiological processes with a plausible link to preterm delivery.[228-230]

There are few publications which explore fetal genomic contributions to premature delivery. When such studies have been performed, findings have often not been replicated in other datasets and have not involved a biological pathway likely to be important in prematurity. For example, a genome-wide association study of 1,349 cases of preterm birth and 12,595 controls identified only two loci statistically associated with preterm birth.[227] The association was present in only two of the five studied ethnic subpopulations. Validation was attempted using three other large datasets in which genomic information had been gathered for other reasons, but gestation had been recorded by the study. Neither of the two loci identified in the index study were found to show a significant effect (in the same direction) in the validation datasets.

An earlier GWAS by Zhang and co-workers had identified 20 loci of interest in fetal genomes (within 14 genes), two of which met the genome-wide significance threshold.[231] Each of these identified loci were located in non-coding regions. Some were introns in genes and others were in intergenic DNA regions (or in pseudogenes, DNA regions which do not code for a protein but are similar to other regions which do code for proteins), as defined by referring to the NIH Variation Viewer.[232] Several of the loci of interest were found in (introns of) the same genes, and they were often in genes associated with immune functioning, a plausible pathway for triggering preterm birth. However, none of the identified loci were replicated in a validation set. GWAS studies identify common variations which are associated with a phenotype. As these variations are common, they are unlikely to the cause of rare diseases or phenotypes, but they identify regions of the genome which may contain pathological variants. Therefore, genes in the same region as the common variants identified by Zhang and co-workers may be useful targets to investigate for causes of preterm birth.

The use of genomic techniques in the study of the nutrition and growth of preterm infants is much more limited. Focused genetic testing was used to investigate an association between parenteral nutrition-associated cholestasis and a gene involved in bile transport (*mdr3*).[233] Genome-wide association studies (GWAS) have identified genetic loci which are associated with birthweight in full term infants (and presumably on growth *in utero*).[92-94] Some identified genes have known effects on metabolic processes relevant to growth, including *ADCY5*'s implication in glucose regulation,[93] and some have been shown to influence birthweight in the preterm population.[95] It is unclear to what extent genetic factors influencing *in utero* growth will be implicated in the growth of preterm infants in the *ex utero* environment. Genomic techniques have not yet been applied to the *ex utero* growth of preterm infants. Exome sequencing has proven fruitful in identifying genetic determinants of other elements of preterm health, including respiratory pathology,[96] suggesting that genomic approaches may shed light on other problems of prematurity.

Preterm infants frequently exhibit disturbances of glucose metabolism at glucose delivery rates which are required to provide adequate energy intake. Insulin, c-peptide and glucagon-like peptide (GLP) have been implicated in this phenomenon and have been shown to display aberrant levels during the preterm period.[97] Animal models have confirmed changes in expression of multiple factors linked to glucose metabolism in primate fetuses, including insulin receptor (*IR*)-β, Akt and glucose transporters (*GLUT-1* and *GLUT-4*).[98] Targeted analysis of polymorphisms in these genes may reveal a role in the glucose intolerance seen in preterm infants.

Genomics of Glucose Handling

Significant heterogeneity is seen in the capacity for preterm infants to process glucose. Many demonstrate marked glucose intolerance, with blood glucose levels rising above the recommended threshold in response to nutrient intakes required to support growth and development. Several environmental factors are associated with hyperglycaemia, especially earlier gestation, lower birthweight and acute illness.[234, 235] However, these factors do not explain all of the variation in glucose tolerance and genomic factors are likely to also be implicated. One previous study has aimed to assess the influence of genetic loci on hyperglycaemia in low-birthweight infants, focusing on genes known to be implicated in type 2 diabetes.[236] This article is published in Russian and is not indexed on PubMed, but the English-language abstract identifies a SNP in one gene (*PPARG*) which was associated with hyperglycaemia in their cohort. *PPARG* is the gene for a nuclear receptor which is known to influence adipocyte differentiation and is associated with diabetes.[237]

Aberrant glucose handling is seen in patients with diabetes. Type 1 diabetes is an autoimmune process associated with loss of pancreatic beta cells and failure of insulin production. Type 2

diabetes is associated with insulin resistance, making it more similar to neonatal hyperglycaemic states. Therefore, validated panels of genes were sought which aim to list genes implicated in type 2 diabetes as these genes will likely be involved in glucose handling. The NHS PanelApp is hosted by Genomics England and lists panels of genes which are associated with diseases (or groups of diseases). The lists are collated by experts in the diseases in question and each listed gene is associated with an assessment of the certainty of its association with disease.

The NHS PanelApp was used to identify a validated set of genes implicated in glucose control. PanelApp does not feature a panel specific to glucose handling but includes a number of panels related to diabetes. The Neonatal Onset Diabetes and Monogenic Diabetes panels are designed to identify unusual diabetes phenotypes. Patients in our study did not exhibit clear signs of neonatal diabetes but some had glucose intolerance. The panel most suited for exploring changes in glucose handling was the Familial Diabetes panel. This panel focuses on non-insulin dependent diabetes without autoimmune associations, a disease phenotype more closely related to neonatal glucose intolerance.[238]

10.2 Hypothesis

I hypothesised that:

- Blood for DNA analysis could be obtained from preterm infants at the time they needed blood tests for clinical monitoring, and that DNA could be extracted for exome analysis.
- That gene panels could be applied to the resultant data which would identify genetic variants associated with inappropriate glucose handling as identified by blood glucose measurements stored on the SPND.

10.3 Aims

Due to limitations of time, budget and availability of subjects during this doctoral research project (exacerbated by the covid pandemic), it has not been possible to gather samples from a sufficient number of preterm infants to apply GWAS techniques to discovering associations between genetic loci and growth. Therefore, the aims of this section are limited and focus on preparation, feasibility testing and method familiarisation to support more extensive genomic work in the future.

Therefore, the aims of this section are:

1. To establish the feasibility of collecting DNA and analysing exome data in the context of preterm infants.
2. To apply an *in silico* panel of genes to identify any mutations within genes:

- which have been previously identified as associated with preterm birth.
- which are implicated in glycaemic control.
- which are known to cause inborn errors of metabolism.

10.4 Methods

10.4.1 Gene Panel Selection

Having identified that there would be an insufficient number of subjects to perform genome-wide assessments for variants associated with growth, a decision was made to identify a set of genes (or genomic loci) to focus the search for variants of interest in the sequenced group of infants. Genes were selected within the three categories in the aims section above (section 10.2).

Genes Associated with Preterm Birth

The 14 genes identified by Zhang and co-workers were included in the list of genes to be assessed (Table 10-1). The paper identified some loci as being situated within genes, whereas reference to the NIH Variation Viewer suggested these were intergenic loci (albeit near the genes listed in the paper). It should be noted that only one of the SNPs identified in the Zhang paper was captured in the exome sequences taken for this project, as, by definition, an exome examines coding sequences only. Loci were provided as GRCh37 positions in the Zhang paper and were converted to GRCh38 positions using the UCSC Genome Browser LiftOver program.[239] These loci were checked for overlap with the padded Agilent SureSelect Human All Exon V6 probes using the intersect function of the bedtools program (v2.3).[240]

Table 10-1. List of gene loci identified to be associated with preterm birth. MAF – minor allele frequency in whole cohort studied in Zhang paper[231]; rsID – reference single nucleotide polymorphism identity; SSV6 - Agilent SureSelect Human All Exon V6 probes plus 100bp padding.

Gene Symbol	Minor Allele	rsID	MAF	Class	Captured by SSV6 + padding
ADGRL2	A	rs480745	0.09	Intergenic	No
INPP1	C	rs11892526	0.23	Intergenic	No
KCNH7	T	rs184270	0.03	Intergenic	No
L3MBTL3	T	rs4429972	0.03	Intron	No
LOC100128365	G	rs17527054	0.08	Intergenic (pseudogene)	No
LOC100652953	A	rs6927568	0.03	Intergenic	No
MAN1A1	A	rs2794256	0.27	Intron	No
NOL10	G	rs1651151	0.3	Intergenic	No
NOL10	A	rs266174	0.3	Intergenic	No
NOL10	G	rs266236	0.3	Intergenic	No
PPP2R2C7	G	rs13130860	0.27	Intron	No
RNASET2	A	rs3777722	0.21	Intron	Yes
RREB1	G	rs560131	0.41	Intergenic	No
RREB1	A	rs3863225	0.41	Intergenic	No
RSPO2	A	rs16877149	0.21	Intron	No
RSPO2	T	rs11780793	0.19	Intergenic	No
SMAD9	T	rs563538	0.04	Intron	No
SORL1	C	rs10892761	0.36	Intron	No
SPOCK3	C	rs17703512	0.28	Intron	No
TMEM229A	G	rs17553718	0.31	Intergenic	No

10.4.1.1 Genes Associated with Glycaemic Control

Having identified that the NHS PanelApp panel for familial diabetes was most suitable for identifying genes of interest in glucose handling, this panel was selected for application to the neonatal dataset. The panel included 55 genes and one genomic region in which deletions and duplications have been implicated in maturity-onset diabetes of the young (MODY). One gene was mitochondrial and so was not sequenced in the exomes used for this study. This set of genes included the PPARG gene identified in the Russian study discussed above. Table 10-2 lists the genes included in this panel and explored in this study.

Table 10-2. Genes included in the NHS PanelApp panel for familial diabetes. AR – autosomal recessive; AD – autosomal dominant; XLR – x-linked recessive.

Gene Name	Entity type	Inheritance
ABCC8	gene	AR/AD
AGPAT2	gene	AR
AKT2	gene	AD
APPL1	gene	AD
BSCL2	gene	AR
CEL	gene	AD
CISD2	gene	AR
DCAF17	gene	AR
DNAJC3	gene	AR
DYRK1B	gene	AD
EIF2AK3	gene	AR
FOXP3	gene	XLR
GATA4	gene	AD
GATA6	gene	AD
GCK	gene	AR/AD
GLIS3	gene	AR
HNF1A	gene	AD
HNF1B	gene	AD
HNF4A	gene	AD
IER3IP1	gene	AR
INS	gene	AR/AD
INSR	gene	AD
KCNJ11	gene	AD
LMNA	gene	AD
LRBA	gene	AR
MNX1	gene	AR
MT-TL1	gene	MIT
NEUROD1	gene	AR/AD
NEUROG3	gene	AR
NKX2-2	gene	AR
PAX6	gene	AD
PCBD1	gene	AR
PDX1	gene	AR
PIK3R1	gene	AD
PLIN1	gene	AD
POLD1	gene	AD
PPARG	gene	AD
PPP1R15B	gene	AR
PTF1A	gene	AR
RFX6	gene	AR
SLC19A2	gene	AR
SLC29A3	gene	AR
SLC2A2	gene	AR
STAT3	gene	AD
TRMT10A	gene	AR
WFS1	gene	AR/AD
ZBTB20	gene	AD
ZFP57	gene	AR
ZMPSTE24	gene	AR
BLK	gene	AD
IL2RA	gene	AR
KLF11	gene	AD
LIPC	gene	AR
PAX4	gene	AD
STAT1	gene	AD
ISCA-37432-Loss	region	AD

Genes Associated with Inborn Errors of Metabolism

As discussed in the introduction to this chapter, a trial programme for genomic screening for inherited disease is currently being formulated by Genomics England. Analysis of genomic data from this doctoral project provides an opportunity to explore the kind of genomic information which will be available if this programme is rolled out more widely. Genomics England is currently deciding on a list of target genes but they are likely to include many inborn errors of metabolism, some of which are currently detected using biochemical methods and others which are not currently screened for. The Genomics England PanelApp includes a comprehensive panel for inborn errors of metabolism which were assessed for this project (version 2.3 of the panel).[241] The panel contained 895 genes (of which 37 were mitochondrial). It also includes one genomic region and two areas of tandem repeats.

Pan-Genomic Screening

As the final structure of the analysis of any universal genomic screening has not yet been defined, a further pan-genomic search was carried out to explore whether the genomic data generated in such a screening programme might provide further diagnoses. Therefore, exomes were searched for variants which were flagged as being pathogenic, likely pathogenic or of uncertain significance by ClinVar and were found with a frequency of less than 5% in the GnomAD reference population. The inclusion of variants of uncertain significance allowed these to be analysed individually to assess whether they were likely to be significant in the context of this study. These results were then further filtered to retain only those with a gene quality score of over 30 (to remove variants with very uncertain calls).

Sample Acquisition, Sample Storage and DNA Extraction

Blood samples for exome analysis were opportunistically drawn from infants alongside clinical samples as set out in the GAP Study protocol and ethical approval (14/SC/1275). Whole blood samples of 0.75-1ml were drawn directly into 1ml BD Microtainer blood tubes coated with EDTA. Some samples were drawn by direct venesection whereas others were acquired by accessing indwelling arterial catheters. Whole blood samples were immediately refrigerated and were frozen at -80°C within 48 hours of sampling. The agreed workflow for sample management is given in Appendix 10.

DNA extraction was performed by a senior laboratory technician and assisted by the doctoral candidate. The HigherPurity™ Blood Genomic DNA Extraction Kit (Canvax) was used. The detailed protocol is given in Appendix 10. In summary, whole blood samples were treated with a buffer

and then centrifuged to obtain a pellet containing the white blood cells. The cells were then lysed by incubation with proteinase K. After cell lysis, DNA was resuspended and samples were centrifuged to remove cellular debris. DNA was precipitated with isopropanol and centrifuged to form a pellet of purified DNA, which was then washed with ethanol before being resuspended in a buffer for frozen storage at -80°C.

DNA Quality Assessment

DNA yield was assessed using Qubit fluorometric quantification (Thermo Fisher Scientific). In this method, a selective dye is bound to DNA and fluoresces when it is bound. This fluorescence can be measured and its intensity is proportionate to the concentration of DNA in the sample.

A NanoDrop spectrophotometer (Thermo Fisher Scientific) was used to assess the purity of DNA samples. This technique measures the absorbance of the sample at a range of wavelengths. The relative absorbance of the sample at 260nm and 280nm is referred to as the 260 / 280 ratio. A ratio around 1.8 typically indicates a pure DNA sample. The 260 / 230 ratio can also be used to provide sensitivity to a broader range of likely contaminants, with high purity indicated by a ratio greater than 2.

Sequencing

High-throughput exome sequencing was carried out by an external service provider (Novogene Corporation). The service provider physically fragmented DNA samples and attached adaptors to the fragments. Agilent SureSelect Human All Exon V6 probes were used to select the regions of interest. A NovaSeq 6000 machine (Illumina, Inc) was used to sequence the prepared exome library. FASTQ files were returned to the research team.

Pre-Processing

Pre-processing and quality control of the exome data was performed by me and by members of the Human Genetics & Genomic Medicine department of the University of Southampton, mostly using scripts written by those collaborators and using the resources of the IRIDIS High Performance Computing Facility at the University.

Internal QC of the raw reads (detailed in FASTQ files) was performed using FASTQC. Alignment of the raw reads to the reference genome applied Burrows-Wheeler Alignment (the *mem* command within the bwa program[242]). This produced a SAM (sequence alignment map) file which was compressed to create a BAM (binary alignment map) file indicating the aligned position of each of the raw reads (along with the sequence, the quality score and metadata).

Further pre-processing steps used Picard tools,[243] a command-line computer program using Java to manipulate genomic data. First, the BAM file reads were sorted by their genomic coordinate. Duplicates (formed by reading the same DNA fragment multiple times) were then marked and re-sorted. Base recalibration was then performed to refine the quality scores of individual bases by comparing them to the reference genome at the site of their alignment. This produced new quality scores for each base call. BAM files resulting from these steps were then used for further quality control (QC) analysis and for variant calling.

10.4.2 Quality Control

Coverage

The contractor provided basic QC metrics, focused on the number of base reads provided and coverage of the target exome region. Picard tools were used in-house to assess the coverage provided by the reads (i.e. the proportion of the target exome sequence for which base calls were available). This step required a file defining the consensus coding sequence of the genome (the coordinates of all exons). QC performed by the contractor used the GRCh37 reference genome whereas in-house analysis used the more recent GRCh38 reference genome. Alignment analysis was also performed using a "padded" version of the targeted capture region to optimise coverage across the CCDS. Whilst Agilent probes targeted a region which the company had defined as coding, data were often obtained in the regions 100bp upstream and downstream of these regions. These data were included in alignment to improve coverage of the current CCDS.

Validation Against Known Sex and Ethnicity Information

The verifyBamID program was used to identify whether any samples had been contaminated by DNA from other subjects.[244] The peddy package[245] for the Python programming language was used to predict the sex and ethnicity of the subjects and to define any relationships between subjects, from their sequencing data. Sex was predicted by examining the heterozygosity of reads on the X chromosome. Polymorphic sites outside the pseudo-autosomic region of the X chromosome (i.e. the small region in which the X and Y chromosomes share genes) were selected by peddy and heterozygosity at those sites was assessed. Females may be expected to be heterozygous at many of these sites, whereas males should not be heterozygous at any (although a small number may appear heterozygous due to misreads). Peddy used genomic data from the 1000 genomes project to produce a principal component analysis (PCA) model of the distribution of variants amongst populations with different ethnicities. This PCA model is then used to predict the ethnicity of the subjects. The number of sites where individuals share an allele was used to assess for relatedness of subjects in the study. The predictions of sex, ethnicity and relatedness

were then compared to the known true values for each individual subject to identify any errors in sample handling.

Variant Calling

The pre-processing step produced a processed BAM file for each subject containing each read, checked for quality and aligned to the reference genome (analysis ready reads). The GATK HaplotypeCaller function was used to identify single nucleotide variants (SNVs), insertions and deletions (indels) for each individual. This resulted in a GVCF (genomic variant call format) file for each subject. The GVCF files described every variant identified in the exome of an individual's sample (and explicitly defined the regions which were non-variant from the reference genome), along with a quality assessment of the variant call.

Joint Calling

Samples were joint called within a larger group of individuals across several studies hosted at the University of Southampton.

Firstly, each individual GVCF file was combined using the GATK GenomicsDBImport command to form a GenomicsDB workspace containing the variant information for every individual. This file was then analysed using the GenotypeGVCFs command to joint call the entire group. This step resulted in a final VCF file containing the joint-called variant data for every individual in the group.

10.4.3 Annotation

The joint called VCF file arising from the joint calling step was uploaded to seqr.[218] Annotation was performed automatically by the seqr program. Variant searches were performed with the results limited to the subjects of this study (as opposed to the whole dataset which included individuals from other studies) and to variants found in the genes identified in the panels described above. Variants were further filtered by their ClinVar status and their CADD score.

Variants which were classed as "benign" or "likely benign" within ClinVar were excluded whereas those which were "uncertain", "likely pathogenic" or "pathogenic were included". In the case of prematurity, genes were selected on the basis of their proximity to non-coding variants identified in the Zhang and co-workers paper.[231] An alternative strategy was implemented incorporating CADD scores (see below) and restricting the results to those occurring in less than 10% of the gnomAD exomes project.[246] This excluded variants which are extremely common in the population. Only variants with a CADD score >15 were included.

10.5 Results

Blood samples from 13 infants were acquired for DNA analysis. Basic demographic information for each of the infants is listed in Table 10-3. All were singletons.

Table 10-3. Basic demographic details of infants who underwent whole exome analysis

Subject Number	Sex	Birthweight (g)	Gestation (weeks + days)	Ethnicity
1	F	1180	31+1	White European
2	F	1310	29+1	White European
3	M	610	27+2	White European
4	F	790	24+4	White European
5	M	1590	30+3	White European
6	F	800	26+6	White European
7	M	1110	26+5	White European
8	M	550	26+5	White European
9	F	760	27+5	White European
10	M	2000	31+6	White European
11	M	920	24+5	White European
12	M	520	23+6	Mixed white European and black African
13	M	700	24+1	White European

10.5.1 Quality Control

DNA Extraction Quality

DNA was successfully extracted from all 13 samples using the method described above. Table 10-4 shows the DNA yield and the NanoDrop absorbance ratios for each of the samples. The 260/280 ratio was used to assess purity of the extracted DNA, with a ratio of 1.8 or higher taken to indicate pure DNA. Therefore, slight impurity was detected in sample 12. The 260/230 ratio also assesses purity, with a ratio of at least 2 indicating pure DNA.

Table 10-4. Sample yield and absorbance ratios for the 13 DNA extractions.

Subject Number	Concentration (ng/μl)*	Yield (μg)	260/280 Ratio**	260/230 Ratio**
1	390	39	1.89	2.3
2	415	41.5	1.88	2.2
3	550	55	1.9	2.3
4	442	44.2	1.89	2.3
5	243	24.3	1.88	2.2
6	497	49.7	1.89	2.3
7	301	30.1	1.89	2.3
8	451	45.1	1.89	2.2
9	248	24.8	1.89	2.2
10	128	12.8	1.85	1.98
11	117	11.7	1.87	2.2
12	610	61	1.71	2.1
13	400	40	1.89	2.3

* Concentration determined by Qubit Analyser. **Absorbance determined by NanoDrop analyser

At least 400ng of DNA at a concentration of at least 20ng/μl was required for exome sequencing and so Table 9-4 indicates that the yield of DNA was significantly in excess of the required mass in every case. Two samples (subjects 10 and 12) contained small clots within the whole blood samples and this led to slightly lower purity of the retrieved DNA as indicated by the absorbance ratio values for these samples.

Quality Control by commercial service provider

Novogene also provided further quality information about the read data, indicating that each infant had 6-10 billion raw base pairs sequenced and that Phred quality score was at least 20 (i.e. at least 99% confidence in the base call) for more than 98% of base calls, and a quality score of at least 30 (at least 99.9% confidence) for over 94% of base calls for every infant meeting Novogene's internal criteria for sequencing quality. After alignment, Novogene estimated that at least 99.4% of the target (exome) loci were captured and more than 98.7% of target bases had at least four reads (with 20x coverage at least 87%). Sequencing quality was rechecked using in-house techniques as set out below.

In-House Sequencing Quality Control

Coverage

In-house coverage statistics using Picard are shown in Table 10-5. Table 10-5-A demonstrates the coverage using a region of interest which is limited to the target coding regions of the SureSelect Human All Exon V6 probes. For each subject, 94-95% of target bases (i.e. bases within the identified coding region) had some coverage. Coverage of at least 20 reads was provided for 71.3% to 86.4% of target bases. Table 10-5-B includes the padding regions 100bp on either side of the target coding regions, which are less richly sequenced, with 55% to 73% of bases in this target region sequenced to a depth of at least 20 reads. The padding regions contain information about the flanking regions of exons, areas which are not specifically targeted by the sequencing capture kit but which frequently influence transcription and gene activity. Novogene uses the sixth version of the Agilent SureSelect probes, which was released in 2015, meaning that some changes in the consensus coding region of the genome are not reflected in the probe coverage.

Validation

The verifyBamID tool did not identify any mixing of DNA within the samples, returning a zero value for every subject. The sex of each subject was confidently predicted by peddy and was concordant with the known phenotypic sex of each of the infants. All but one of the subjects was identified as white European, which was consistent with the known ethnicity of the subjects (one of these subjects identified by peddy as white European had some white South African ancestry). One subject (listed as subject 12 in the tables) was identified as black African by peddy. This infant's listed actual ethnicity was mixed, with one parent having white European ethnicity and the other parent having black African ethnicity. Examination of the PCA plot provided by peddy confirms that this infant's PCA features place him halfway between the white European and black African clusters, as expected for his mixed ethnicity (Figure 10-1).

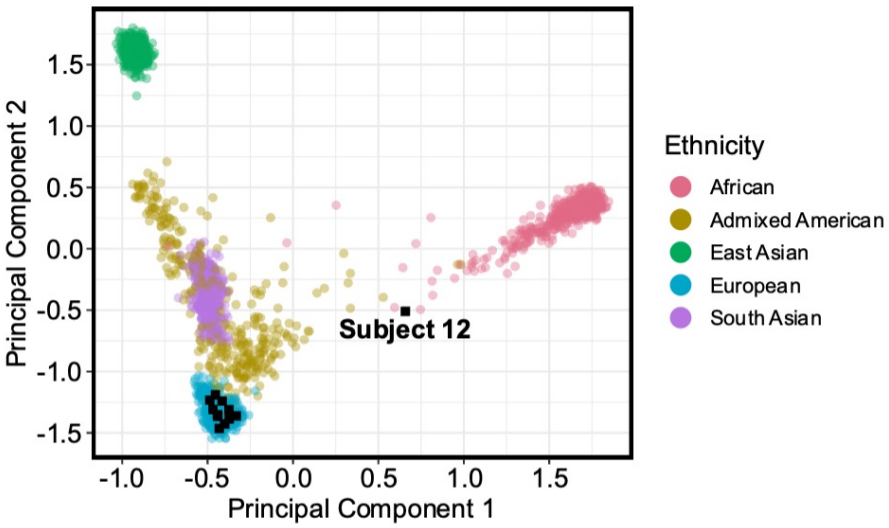


Figure 10-1. Principal component analysis of ethnicity using peddy with ethnicity clusters from 1000 Genomes project illustrated by coloured dots and subjects from this study as black squares.

Table 10-5. In-house Picard coverage statistics, A. using the target region for Agilent SureSelect V6 target region, and B. using the SureSelect target region with the addition of a 100bp padding region before and after each interval.

A.

Subject	Bait Set	Total Reads	Mean Target Coverage	Median Target Coverage	Target bases at 1x (%)	Target bases at 10x (%)	Target bases at 20x (%)	Target bases at 30x (%)	Target bases at 40x (%)	Target bases at 50x (%)	Target bases at 100x (%)
1	SSV6_hg38_gc	65911592	63.3	54	94.85	91.82	85.40	76.19	65.68	55.05	18.08
2	SSV6_hg38_gc	53684592	51.5	44	94.79	90.50	81.35	69.08	56.16	44.07	10.39
3	SSV6_hg38_gc	47682758	46.1	38	94.82	89.05	76.96	62.54	48.69	36.88	8.25
4	SSV6_hg38_gc	63104490	57.8	50	94.82	91.19	83.54	73.05	61.50	50.24	14.49
5	SSV6_hg38_gc	46667872	44.6	38	94.76	88.97	76.89	62.13	47.87	35.68	6.95
6	SSV6_hg38_gc	42472972	42.0	36	94.70	88.72	75.83	60.07	45.00	32.44	5.40
7	SSV6_hg38_gc	40638298	40.1	33	94.72	86.86	71.29	54.77	40.51	29.53	5.74
8	SSV6_hg38_gc	74248466	70.4	60	94.99	92.16	86.41	78.30	68.93	59.38	23.15
9	SSV6_hg38_gc	47679074	45.6	39	94.64	89.42	78.31	64.10	49.89	37.32	7.03
10	SSV6_hg38_gc	49239036	46.2	40	94.74	89.56	78.76	64.86	50.77	38.30	7.25
11	SSV6_hg38_gc	43942312	42.8	34	94.65	87.10	72.40	56.85	43.22	32.43	7.52
12	SSV6_hg38_gc	40391804	39.8	34	94.79	88.18	74.13	57.42	42.09	29.71	4.41
13	SSV6_hg38_gc	50027106	47.9	41	94.85	89.91	79.07	65.26	51.48	39.34	8.72

B.

Subject	Bait Set	Total Reads	Mean Target Coverage	Median Target Coverage	Target bases at 1x (%)	Target bases at 10x (%)	Target bases at 20x (%)	Target bases at 30x (%)	Target bases at 40x (%)	Target bases at 50x (%)	Target bases at 100x (%)
1	SSV6_hg38_pad100_gc	65911592	48.6	37	94.53	85.52	71.63	58.75	47.73	38.40	11.64
2	SSV6_hg38_pad100_gc	53684592	39.6	31	94.39	82.58	65.85	51.31	39.43	29.84	6.59
3	SSV6_hg38_pad100_gc	47682758	35.0	26	94.21	78.27	59.76	45.04	33.54	24.72	5.20
4	SSV6_hg38_pad100_gc	63104490	44.2	34	94.44	83.77	68.56	55.09	43.81	34.45	9.25
5	SSV6_hg38_pad100_gc	46667872	33.8	25	94.15	77.96	59.22	44.26	32.57	23.62	4.37
6	SSV6_hg38_pad100_gc	42472972	32.3	25	94.18	78.64	58.84	42.92	30.64	21.49	3.39
7	SSV6_hg38_pad100_gc	40638298	30.4	22	93.95	74.26	53.83	38.71	27.61	19.64	3.60
8	SSV6_hg38_pad100_gc	74248466	53.6	40	94.65	85.82	72.92	61.08	50.82	42.05	15.03
9	SSV6_hg38_pad100_gc	47679074	34.8	27	94.13	79.53	61.16	46.06	34.09	24.74	4.42
10	SSV6_hg38_pad100_gc	49239036	35.1	27	94.19	79.38	61.36	46.50	34.60	25.32	4.55
11	SSV6_hg38_pad100_gc	43942312	32.3	23	93.85	74.47	54.95	40.50	29.69	21.76	4.74
12	SSV6_hg38_pad100_gc	40391804	30.5	23	94.19	77.03	56.59	40.50	28.38	19.51	2.77
13	SSV6_hg38_pad100_gc	50027106	36.6	28	94.38	80.53	62.47	47.54	35.68	26.43	5.52

10.5.2 Variants Identified

The 13 patients benefitted from joint calling within a cohort with a total of 1470 individuals from 1354 families. Variants were analysed in three groups: the preterm birth panel, the glycaemic control panel and the inborn errors of metabolism panels. The variants identified in each of these panels is discussed in turn below.

Preterm Birth Panel

Initial searches identified no variants in the preterm birth panel which were classified as pathological, likely pathological or unknown in the ClinVar database. When filtering was altered to a combination of CADD score over 15 and a gnomAD exomes frequency of less than 10%, ten variants were identified in six genes (Table 10-6).

Table 10-6. Variants identified using the preterm birth gene panel after filtering for CADD>15 and gnomAD exome frequency <0.1 ordered by CADD score. rsID - reference single nucleotide polymorphism identity (Ch38); CADD - Combined Annotation–Dependent Depletion Score, Hom/Het – whether homozygous or heterozygous for the listed variant.

Subject ID	gene	Position	Mutation	Hom/ Het	rsID	Type	gnomAD Frequency	CADD Score
2	*ADGRL2*	Chr1:81,990,423	G > C	Het	rs72719419	Missense	$4.1x10^{-3}$	26.6
3	*ADGRL2*	Chr1:81,952,006	A > G	Het		Missense	0	25.5
7	*INPP1*	Chr2:190,371,213	C > G	Het	rs140741805	Missense	$1.6x10^{-4}$	25.5
9	*KCNH7*	Chr2:162,373,501	T > C	Het	rs79262587	Missense	$1.0x10^{-2}$	22.9
7	*NOL10*	Chr2:10,600,863	T > C	Het	rs115424813	Missense	$3.1x10^{-3}$	22
8	*ADGRL2*	Chr1:81,987,338	A > G	Het	rs76995529	Synonymous	$7.7x10^{-3}$	20.4
6, 12	*L3MBTL3*	Chr6:130,139,931	C > T	Het	rs17633592	3 prime UTR variant	$3.5x10^{-2}$	17.72
4	*KCNH7*	Chr2:162,512,573	G > C	Het	rs191243509	Intron	$4.3x10^{-3}$	17.26
12	*RSPO2*	Chr8:108,082,776	T > C	Het	rs112769314	5 prime UTR variant	$1.7x10^{-2}$	16.17
12	*RSPO2*	Chr8:107,901,346	GAAT > G	Hom	rs35321772	Intron	$4.4x10^{-2}$	15.95

These variants were found in a range of genes, each of which is discussed below, along with a discussion of the specific mutations seen in this cohort.

ADGRL2

ADGRL2 codes the adhesion G protein-coupled receptor L2, which is a cell surface membrane receptor implicated in exocytosis (a way of releasing cellular contents into the extracellular environment). The gene was previously called LPHN2. It has a very high loss of function constraint

score, indicating that loss of function mutations are very rare in this gene (compared to all other genes), suggesting an evolutionary pressure to retain its function. It is expressed in many tissues, but especially strongly in the placenta.[247] It has also been shown to be expressed twice as strongly in fetal adrenal glands than in adult adrenal glands.[248] Whilst these distributions of expression raise the possibility of a role in preterm birth, such assertions must be extremely guarded considering the very small number of infants reviewed in this study. The onset of labour, including preterm labour, is controlled by a complex interaction of fetal, maternal and placental hormonal signalling, with the fetal hypothalamic-pituitary-adrenal axis and fetal androgen production (by the adrenal gland) known to be central to the process of labour initiation.[249, 250] Therefore, it is plausible that differences in the control of exocytosis of androgens from the fetal adrenals could moderate the risk of preterm birth.

Each of the variants seen in *ADGRL2* in this group was a heterozygous SNP and no individual had more than one SNP within *APGRL2*. The variant seen in subject 8 is synonymous, meaning that there is no change in the encoded amino acid sequence. This may be expected to render the variant unlikely to cause disruption to biological functioning. However, there are many ways in which synonymous variants can alter function, especially if they interfere with normal mRNA processing or splicing of exons.[251] The variant in subject 8 is predicted to have a high risk of pathogenicity by CADD score. It also has a SpliceAI score of 0.09, indicating a 9% chance of affecting the splicing of the gene. SpliceAI is a deep neural network-trained algorithm used to predict the risk of a variant influencing gene splicing.[252] Therefore, despite this mutation's synonymous sequence status, there are some reasons to believe it may influence production of the encoded protein. The other two identified variants (in subjects 2 and 3) are missense variations within exons of the gene. The variant seen in subject 2 is reasonably common (being present in around 0.4% of individuals in large datasets) and was found in 1% of individuals in the whole group used for joint calling in this study. The variant found in subject 3 is very rare, having not been identified in any of the 140,000 exomes and genomes in the gnomAD project. This unusual level of conservation suggests that mutation at this locus is poorly tolerated in evolutionary terms and is more likely to be deleterious.

INPP1

INNP1 codes for inositol polyphosphate-1-phosphatase. It moderates phosphatidylinositol signalling and is expressed in many organs. Investigation of the gene has identified that it is upregulated in several cancers[253] and may influence response to lithium in psychiatric disorders.[254] It has no obvious biological link to preterm birth. Subject 7 had a heterozygous missense variant in the *INPP1* gene. Various in silico tests indicated that this variant may be deleterious. High CADD and Eigen scores reflected that it is well conserved. The PolyPhen score[255] was also high, predicting impact of the amino acid change on the resultant protein's function.

KCNH7

KCNH7 codes for a voltage-gated potassium channel. It has been identified as enriched in a subpopulation with bipolar disorder within a genetically isolated population.[256] It is expressed in neurological tissues. It has a high degree of missense constraint, indicating that missense variants are less common in this gene than would be expected. Subject 4 had a variant which was rare in reference populations but was within an intron with a very low SpliceAI score. Subject 3 was heterozygous for a missense variant which is common in reference populations (around 1%).

NOL10

NOL10 codes for a nucleolar protein which is involved in RNA binding to ribosomes. It is ubiquitously expressed. It has no biological link to preterm birth. Subject 7 had a missense variant in NOL10 which was uncommon in reference populations (around 0.3%).

L3MBTL3

L3MBTL3 codes for histone methyl-lysine-binding protein, which is probably involved in myelopoiesis[257] but is ubiquitously expressed. Subjects 6 and 12 shared a variant within the 3-prime untranslated region which may influence post-transcription gene expression and which was common in gnomAD genomes (0.5%). However, it should be noted that these subjects had three and four reads respectively at the locus of interest, each with only one read of the minor allele, rendering these results very uncertain.

RSPO2

RSPO2 codes for an r-spondin which is involved in the Wnt/β-catenin signalling, a process key to embryonic development. Other mutations in the RSPO2 gene have been identified as causing severe limb malformations in human fetuses[258] and mouse knockout models have developmental abnormalities which are fatal during fetal life.[259] Subject 12 had two variants in the *RSPO2* gene. The first was within an intron and was poorly characterised by exome sequencing (as expected). The other variant was heterozygous and within the 5-prime untranslated region which was common in gnomAD genomes (1.7%).

Glycaemic Control Panel

Four variants were identified in genes listed in the glycaemic control gene panel (section 9.3.4.2) were labelled as not known to be benign by ClinVar and with a CADD score of greater than 15 (Table 10-7).

Table 10-7 Variants identified using the glycaemic control gene panel after filtering for CADD>15 and ClinVar not benign, ordered by CADD score. rsID - reference single nucleotide polymorphism identity (Ch38); CADD - Combined Annotation–Dependent Depletion Score. Hom/Het – whether homozygous or heterozygous for the listed variant.

Subject ID	Gene	Position	Mutation	Hom/ Het	rsID	Type	gnomAD Frequency	CADD Score
2	*PCBD1*	Chr18:22,171,983	G > C	Het		Missense	$1.2x10^{-5}$	26.4
1	*LIPC*	Chr15:58,563,549	C > T	Het	rs113298164	Missense	$2.9x10^{-3}$	23.4
8	*GATA6*	Chr10:70,884,021	A > G	Het		Missense	$6.7x10^{-5}$	21.2
9	*BLK*	Chr8:11,548,067	G > A	Hom	rs55758736	Missense	$1.2x10^{-2}$	18.47

Each of these variants is discussed below.

PCBD1

PCBD1 codes for a protein involves in the breakdown and processing of amino acids. Homozygous mutations in the gene are associated with unusual diabetes phenotypes characterised by onset in adolescence mediated by insulin resistance.[260] Heterozygote carriers in families with recessive disease had an elevated risk of late-onset diabetes. Subject 2 had a missense variant within the *PCBD1* gene. It was an extremely rare variant in gnomAD and most in silico models predicted this variant to be deleterious, although Polyphen did not predict a change in protein function based on the change in amino acid sequence. A single case report using Sanger sequencing in the early 1990s identified an individual with compound heterozygous mutations of the *PCDB1* gene causing a mild phenotype of hyperphenylalaninemia.[261] One of these mutations was the exact variant seen in subject 2 of this study (c.244T>C causing p.Cys82Arg). This provides additional evidence that the SNP in question disrupts the functioning of the resulting protein and that subject 2 may be at risk of diabetes later in life. Subject 2 was not extremely preterm (born after 29 completed weeks of gestation) and was normoglycaemic during her neonatal stay (except one transient episode of mild hypoglycaemia at around two weeks old) and did not receive insulin treatment. She was managed with parenteral nutrition (PN) for a total of 14 days, comprising a period of one week after birth and a further week at around three weeks of life when feeds were paused due to abdominal distension (but without diagnostic features of necrotising enterocolitis).

LIPC

LIPC codes for hepatic lipase, an enzyme which regulates plasma lipids. A variant in the promoter region of this gene is a known risk factor for development of type two diabetes.[262] Subject 1 had a heterozygous missense variant in this gene, which was predicted to disrupt protein function by Polyphen and had high in silico risk scores. A GWAS of familial dyslipidaemias identified this specific SNP as a risk factor for familial dyslipidaemia.[263] Subject 1 was born after 31 completed

weeks of gestation, received two weeks of parenteral nutrition, did not require insulin therapy, was normoglycaemic throughout her neonatal stay and had an uncomplicated neonatal course.

GATA6

GATA6 codes a transcriptional regulatory protein. It is haploinsufficient (meaning that two functioning copies of the gene are required for normal functioning of the protein). Mutations in *GATA6* are well-recognised causes of congenital cardiac malformations. Several case reports have identified *GATA6* mutations associated with pancreatic dysgenesis and neonatal diabetes.[264] Subject 8 had a heterozygous missense variant in *GATA6*. This variant scored highly on in silico models of deleteriousness, likely due to its rarity in reference populations (0.007% in gnomAD). However, Polyphen predicted that protein functioning was likely to be unaffected by the variant. Subject 8 had no congenital cardiac disease (confirmed by echocardiography). He was born extremely prematurely at 26 weeks of gestation with severe intrauterine growth restriction being born around the 0.4th percentile for weight. He experienced severe hyperglycaemia in response to normal management with parenteral nutrition, required insulin therapy and creation of a bespoke PN formulation to reduce his glucose intake. His growth was normal.

BLK

The *BLK* gene codes for a tyrosine kinase which is involved in insulin production in response to hyperglycaemia. Some mutations in this gene cause maturity onset diabetes of the young (MODY). Subject 9 had a homozygous missense variant in *BLK*. The homozygous state of this variant is rare, being present in 0.01% of gnomAD exome individuals. This variant segregated with MODY in a family with two affected members.[265]Functional data from the same paper suggested that the mutation caused a decrease in production of the target protein. Subject 9 was born extremely prematurely after 27 completed weeks of gestation. She had an uncomplicated neonatal course, receiving PN for just over two weeks. She was mostly normoglycaemic during her neonatal stay (with a few episodes of mild hypoglycaemia after cessation of PN) and did not require insulin therapy.

Inborn Errors of Metabolism

The inborn errors of metabolism panel was designed to identify incidence of rare genetic diseases. Therefore, searches were limited to those variants labelled as pathogenic or likely pathogenic by ClinVar. Variants in these genes are listed in Table 10-8.

Table 10-8 Variants identified using the inborn errors of metabolism gene panel after filtering for CADD>15 and ClinVar pathogenic or likely pathogenic, ordered by CADD score.
rsID - reference single nucleotide polymorphism identity (Ch38); CADD - Combined Annotation–Dependent Depletion Score. Hom/Het – whether homozygous or heterozygous for the listed variant.

Subject ID	Gene	Position	Mutation	Hom / Het	rsID	Type	gnomAD Frequency	CADD Score
6	*GFM1*	Chr3:158,660,906	T > [*]	Het		Frameshift	$4.0x10^{-6}$	35
2	*PCBD1*	Chr10:70,884,021	A > G	Het		Missense	$1.2x10^{-5}$	26.4
11	*GNMT*	Chr6:42,963,149	C > A	Het		Missense	$4.1x10^{-4}$	23.6
9	*BTD*	Chr3:15,645,186	G > C	Het	rs13078881	Missense	$3.2x10^{-2}$	23.4
3	*FKRP*	Chr19:46,756,276	C > A	Het	rs28937900	Missense	$1.0x10^{-3}$	22.2

*TGCATTGTTTGGCATTGACTGTGC

Each of these variants were heterozygous variants in genes which are known to demonstrate an autosomal recessive inheritance pattern. Therefore, these variants represent carrier status for a range of inborn errors of metabolism but none of the infants are likely to manifest clinical disease.

Pan-Genomic Screening

Pan-genomic screening identified 30 variants in genes marked as pathogenic, likely pathogenic or of uncertain significance by ClinVar (Table 10-9).

Most variants were found in the heterozygous state in genes where disease is known to be inherited in a recessive pattern. Three variants raised issues which may be relevant to findings from universal sequencing. One infant was found to be a carrier for the ΔF508 mutation of the CFTR gene – the most common mutation causing cystic fibrosis in Caucasian populations. A second infant had a mutation in the BCO1 gene. This specific mutation has been previously reported as the likely cause of hypercarotenaemia and vitamin A deficiency in one patient who was heterozygous for the mutation.[266] In silico verification confirmed that the mutant form of the protein was not functional and the paper suggested that the gene was haploinsufficient. A third patient had a mutation in the MYO6 gene. This variant has been identified as a cause of deafness in the heterozygote state.[267] The infant in question has been found to have some high-frequency hearing loss during routine hearing screening. However, the genome quality score for this variant was low at 18 (with a read depth of 13 and an allelic balance of 0.15). Furthermore, this variant was identified in 1% of the call set used for joint calling (32 out of 2934 alleles) but only in 0.05% of subjects in the GnomAD exomes reference population, suggesting that a systematic error in the

sequencing of the local call set may have led to false calls at this locus. The implications of these results are discussed below.

Table 10-9. Variants identified by pan-genomic screening, after filtering for CADD>15 and ClinVar pathogenic or likely pathogenic, ordered by CADD score. rsID - reference single nucleotide polymorphism identity (Ch38); CADD - Combined Annotation–Dependent Depletion Score. Hom/Het – whether homozygous or heterozygous for the listed variant. Genome quality scores <30 highlighted.

Subj.	Gene	Position	Mutation	Hom/ Het	rsID	Type	gnomAD Frequency	CADD Score	Genome Quality Score
8	*C8B*	Chr1:56,940,965	G > A	Het		Stop gained	1.1E-03	41	99
6	*MYO6*	Chr6:75,890,140	G > GA	Het	rs551348450	Frameshift	3.6E-04	35	18
11	*FAM161A*	Chr2:61,839,695	T > A	Het		Stop gained	2.8E-04	35	99
6	*GFM1*	Chr3:158,660,906	T > [*]	Het		Frameshift	4.0E-06	35	99
4	*MRI1*	Chr19:13,768,642	G > A	Het		Missense	1.3E-04	33	99
11	*ANO5*	Chr11:22,276,148	GC > G	Het		Frameshift	8.0E-06	29.7	99
1	*ALOX12B*	Chr17:8,077,109	G > A	Het		Missense	3.2E-05	29	99
11	*CYP24A1*	Chr20:54,158,136	G > A	Het		Missense	6.6E-04	28.9	99
2	*MUTYH*	Chr1:45,332,803	T > C	Het		Missense	1.5E-03	27.7	99
6	*PADI3*	Chr1:17,262,194	T > A	Het	rs142129409	Missense	4.5E-03	27.4	99
3	*CHRND*	Chr2:232,531,350	A > C	Het		Splice site missense	2.7E-04	27.2	99
2	*CEP290*	Chr12:88,118,527	AT > A	Het		Frameshift	1.7E-03	26.7	9
3	*M1AP*	Chr2:74,581,766	C > CA	Het		Frameshift	2.1E-03	26.5	99
2	*PCBD1*	Chr10:70,884,021	A > G	Het		Missense	1.2E-05	26.4	99
7	*SEC63*	Chr6:107,893,550	G > GT	Het		Frameshift	1.6E-04	25.6	4
13	*BARD1*	Chr2:214,781,245	GTT > G	Het		Frameshift	4.3E-06	25.1	99
2	*GNRHR*	Chr4:67,754,019	T > C	Het	rs104893836	Missense	2.8E-03	24	99
1	*BCO1*	Chr16:81,264,677	C > T	Het	rs119478057	Missense	1.4E-03	23.7	99
11	*GNMT*	Chr6:42,963,149	C > A	Het		Missense	4.1E-04	23.6	99
9	*FCGR1A*	Chr1:149,784,224	C > T	Het	rs74315310	Stop gained	4.1E-03	23.4	99
9	*BTD*	Chr3:15,645,186	G > C	Het	rs13078881	Missense	3.2E-02	23.4	99
3	*KIAA0586*	Chr14:58,428,294	G > GA	Het		Frameshift	0.0E+00	23.4	99
4	*PRRT2*	Chr16:29,813,694	GC > G	Het		Frameshift	5.3E-03	23.1	33
3	*FKRP*	Chr19:46,756,276	C > A	Het	rs28937900	Missense	1.0E-03	22.2	99
4	*CFTR*	Chr7:117,559,590	ATCT > A	Het	rs1297060838	Deletion	7.1E-03	22	99
8	*PKHD1*	Chr6:52,065,000	TG > T	Het		Frameshift	2.4E-05	21.9	99
9	*DNAAF4*	Chr15:55,467,043	A > AT	Het		Frameshift	1.4E-04	21	21
6	*PLG*	Chr6:160,706,469	A > G	Het	rs73015965	Missense	3.0E-03	19.73	99
4	*GNRHR2*	Chr1:145,919,695	T > C	Het		Missense	0.0E+00	14.89	99
11	*EMP2*	Chr16:10,547,590	C > T	Het		Missense	1.2E-04	10.15	99

* TGCATTGTTTGGCATTGACTGTGC

10.6 Discussion

Sample Acquisition

Whole blood samples were obtained for 13 preterm infants. Each of the samples was of sufficient volume and quality to allow DNA extraction. The parents of several infants declined consent for their infants to be included in genetic analysis. Of these parents, one expressed concern about the storage, use and data security of such sensitive personal data. The remaining parents who declined involvement were concerned about the volume of blood required, even after discussion of the intention to take only 1ml of blood. This is not unreasonable in such small and vulnerable infants (who may have a total blood volume of only 40ml).

Two approaches which could be used to improve recruitment are apparent. Firstly, genetic material not from blood samples taken from the infant could be used. Buccal swabs have been used to obtain DNA for NGS analysis, including in extremely preterm neonates.[268] Concerns about contamination with maternal DNA have largely been theoretical and have not been encountered in real life situations. Cord blood is also often readily available and suitable for NGS.[269] It is often not possible to take detailed informed consent from parents prior to delivery, and therefore an ethics application for such collection would likely need to allow storage of cord blood prior to consent, with DNA extraction carried out only if consent is subsequently given.

Secondly, smaller blood volumes could be used. The lowest yield of DNA in this study was 11.7μg, when a minimum of 400ng is required for the NGS method used. DNA extraction becomes somewhat more difficult with lower volumes of blood as the DNA pellet becomes difficult to see with the naked eye. However, that problem is not insurmountable. Even with a generous margin for differences in white cell count and extraction success, it is likely that a significant reduction in blood volume would still allow for good DNA yields, as low as 250μl of blood and possibly considerably less. Subsequently to the batch of DNA extractions documented in this thesis, further samples have been taken at 500μl rather than 1ml.

DNA Extraction and Sequencing Coverage

As discussed in the section above, DNA yields were high in every case. Relative absorbances indicated high quality of the extracted DNA. Coverage of the consensus coding sequence calculated by Picard was good at around 95%. The coverage quoted by Novogene was much higher, as the company used the GRCh37 assembly of the human genome whereas the in-house analysis was performed using the updated GRCh38 assembly. Planned improvements in the exome capture kit are likely to improve coverage of the coding region in the future. No contamination, mixing or mislabelling of samples could be detected.

Variants

The ability to make novel discoveries from these data is severely limited by the small number of infants included. The following sections explore the findings from the three gene panels in the context of the few babies included.

Preterm Birth Panel

Preterm birth is a complex process which is definitely influenced by environmental factors such as infection. Properties of the pregnancy, especially multiple pregnancy, make preterm delivery much more likely and maternal problems such as pre-eclampsia also increase the risk of preterm labour. Maternal genetic factors are likely to influence preterm labour, either directly or by moderating the risk of complications causing preterm labour (such as pre-eclampsia). The role of the fetus in initiating preterm labour is less well understood. Fetal factors undoubtedly influence preterm birth, for example growth-restricted fetuses are much more likely to be delivered preterm.[270] Fetal and maternal genetic factors are likely to interact in complex ways to influence placentation, which in turn influences the risk of growth restriction and preterm labour.

Initiation of labour is a complex process but it is likely that fetal stress signals can prompt it.[249, 250] Therefore, genetic variation in the fetal endocrine system (and especially the hypothalamic-pituitary-adrenal axis) may contribute to onset of (preterm) labour. *ADGRL2* variants may influence adrenal response and were found in three infants in this sample and in a GWAS of preterm delivery [231] GWAS studies identify variation in a large number of SNPs across the genome, highlighting genomic regions which are associated with presence or absence of a phenotype (typically a disease). The common variants identified in a GWAS are not likely to cause the disease in question but point towards a region of interest. Therefore, the identification of specific mutations within the *ADGRL2* gene in this study is consistent with the signals from previous GWAS studies. However this result must be interpreted with extreme caution, especially as one of the variants was intronic (albeit near a splice site) and another was very common. The approach of selecting a gene panel based on the results of a previous GWAS means that there is a risk of confirmation bias in these findings. Nevertheless, the identification in this study of mutations in *ADGRL2*, taken together with the GWAS data, suggest that aberrations in this gene may be useful targets for scrutiny in larger genomic studies in preterm infants. The remaining variants identified in the preterm birth panel do not have convincing biological links to likely mechanisms of premature onset of labour.

Glycaemic Control Panel

A handful of variants were identified which could plausibly interfere with glucose handling based on previous reports. However, the infants in question mostly did not display aberrant glucose handling. The one infant who did have significant glucose intolerance was at very high risk for glucose handling problems based on conventional risk factors (including extreme prematurity, sepsis and severe intrauterine growth restriction). This patient highlights the difficulty of unpicking the influence of genetic risk factors from environmental influences known to impact glucose tolerance. The other variants identified also raise questions about how such genomic information should be treated if it were to become more commonly available. Protocols would need to be formed to define whether there should be periodic reassessment of genomic information as understanding increased, or in response to changes in the phenotype or disease state of the individual in question.

Inborn Errors of Metabolism

Not unexpectedly, none of the infants in the study were diagnosed with an inborn error of metabolism based on the selected panel. Several infants were identified as likely heterozygous carriers of recessively inherited diseases. The Newborn Genomes programme has not yet defined how such data would be managed.[225] It seems likely that such information would not be routinely communicated to parents as the child would not be expected to develop clinical disease. However, the Data Protection Act (2018)[271] specifies that data pertaining to a person's health must be accessible to him or her. Ethical approaches to information such as carrier status of serious diseases are actively being developed.[272] Current guidelines recommend analysing only genetic regions of interest to the clinical question and not divulging genetic information affecting reproductive choices until the child reaches adulthood. It is unclear whether such approaches are compliant with the Data Protection Act, in the case when a whole genome or exome has been sequenced and is therefore known. Routine sequencing of newborns will generate a vast amount of information which has the potential to inform understanding of disease carrier status or risk of future ill-health. The findings of likely carrier status, even in a restricted panel of genes in a small number of infants, highlights the importance of a robust framework to manage subject access requests and interest in accessing one's own genomic data if routine newborn sequencing were implemented.

Pan-Genomic Screening

As discussed in the paragraph above, it is not yet clear how the genomic results from universal screening would be analysed and filtered. Even if a focused panel is applied, more extensive genomic data may be available to subjects of the screening. The findings of the pan-genomic

screening for variants known to cause disease illustrate a number of issues which could arise as a result of these data being available to individuals and families. One patient was identified as a carrier of the ΔF508 mutation of the CFTR gene, which is known to cause cystic fibrosis (CF). This variant will have no clinical impact for the subject, but it may influence future reproductive decisions. Partners of CF carriers are eligible for testing. If such results are discovered during genomic screening, there would need to be clear pathways for deciding if and when to inform the subject (when he or she reaches a level of maturity to understand the information) or his or her family.

Another infant had a genetic mutation which was thought to be the cause of hypercarotenaemia in a previous case report. It is unclear whether this result is significant, especially without correlation with any clinical signs or symptoms in the child. The incidence of the mutation in the reference populations was around 0.2%. Whilst the true incidence of familial hypercarotenaemia is unknown, it is likely to be much rarer than 0.2%, suggesting that the mutation may not have full penetrance or may rely on a gene-environment interaction (for example, high intake of beta-carotene) to show a clinical effect. The interpretation and communication of these data is complex, and an infrastructure to deal with such uncertain findings would be required if patients were likely to be able to obtain such data.

Thirdly, one infant with hearing impairment had a genetic mutation related to deafness. However, the quality of this call was very questionable as discussed in the results section. Hearing impairment is common in infants born extremely preterm (prevalence of 1-4%, around ten times that in the term-born population).[273] Considering the poor quality of the read at this locus, it is likely that this base call is unrelated to the child's hearing impairment.

The pan-genomic element of this analysis demonstrates the risks and benefits of universal screening. Whilst no diagnoses of monogenic disorders were made, there were several findings of uncertain significance in only 13 infants. It remains unclear to what extent genomic information would be available to families, but it is not clear that such information could be legally withheld. If genomic information were to become available in this way, there would be a significant need for families (and subjects themselves as they matured) to access genetic interpretation and counselling. Current UK health sector systems are not configured to deal with such a large workload, meaning that significant investment and training may be required to meet that demand. Whilst genomic testing may provide some benefits in terms of diagnosing rare monogenic disorders, careful consideration of how the other genomic information is processed, stored and made available to subjects and their families is required to minimise accompanying risks to the child (who cannot give consent to any testing).

10.7 Conclusions

The genomic element of this project has demonstrated local capacity to sample and extract DNA from blood samples of preterm infants, and to use an external contractor to sequence the resulting DNA samples. Local data analytical skills allowed joint calling of the sequences. Considering the small number of infants, it was not possible to identify variants likely to influence metabolic processing or growth. ADGRL2 has previously been identified as a possible target for genomic influences on preterm growth, and my data provide some very limited support for that hypothesis. Genes for glycaemic control could not be identified. Processing of exomes for inborn errors of metabolism highlighted some challenges which could be expected if universal genomic screening of newborns were implemented.

Where variants have been identified which may influence glucose handling, the SPND database has been used to identify whether these infants suffered derangement of glucose homeostasis. The limited number of infants included in genomic analysis precluded more formal integration of genomic results with analysis of nutritional intake or growth.

Chapter 11 Metabolomic Profiling

Metabolic processing of nutrients is immature at birth in preterm infants. This chapter aimed to develop a workflow to analyse the urinary metabolomic profile of preterm infants and to explore metabolomic signals relating to maturation. Ninety-three urine samples from 15 infants were analysed and this chapter sets out the methods, results and analyses of the NMR spectra generated.

Sample acquisition was led by me and was carried out by me and by research nurses on the neonatal unit. Urine samples were processed for analysis by me. Running samples on NMR apparatus was carried out by Prof Jonathan Swann and analysis of the spectra was performed by me with his advice.

11.1 Background

Central Concepts of Metabolomics

Metabolomics is the analysis of small molecules in biological mixtures.[274] Whilst conventional analytical approaches to the biochemical makeup of biological fluids focus on targeted quantification of one molecule (or a handful of molecules), metabolomic profiling aims to describe all small molecules present in a sample. This characterisation can use several methods and apply them to different fluids to gain insights into the metabolic processes occurring within the body. The following subsections explore the methods available, the fluid samples which can be analysed and the current state of metabolomic research in preterm infants.

Previous Metabolomic Investigations into the Growth and Nutrition of Preterm Infants

Several groups have used metabolomic methods to assess matters related to the growth and nutrition of preterm infants. Most such studies have been based on a specific intervention. For example, a group in Newcastle demonstrated some changes in the stool metabolome associated with introduction of a probiotic product.[275]

Necrotising enterocolitis (NEC) has been a frequent target for studies investigating the uses of metabolomic profiling in preterms. Several groups have identified systematic differences between urine metabolomics in cases of NEC and control infants.[276, 277] The Newcastle group found

differences in the serum metabolome.[278] Each of these studies found significant differences in the metabolomic profile of those infants diagnosed with NEC, although an "early warning" signal of impending NEC has not been discovered.

Other studies have focused on describing the normal trajectory of the metabolomic profile over time. A Norwegian group performed secondary analysis of some urine samples gathered during an interventional trial of different nutritional strategies.[99] They did not find a significant difference between the intervention and control groups. This study identified profound differences between the urinary metabolomic profile of infants during their first week of life and during subsequent weeks. The age of a preterm infants is a more complex concept than that for an infant born at term. Each infant possesses a postmenstrual age (which is essentially a post-conception age) and a postnatal age (i.e. the amount of time which has elapsed since birth). Tanner illustrates this point with examples: glucose-6-phosphotase activity in the liver is prompted by the need for glycogenolysis at birth (regardless of postmenstrual age), whereas the shift from production of fetal to adult haemoglobin operates strictly based on postmenstrual (or postconceptional) age.[116]

The Norwegian study identified an overlapping group of metabolites which changed in response to postmenstrual and postnatal ages. In addition, they identified a metabolomic signature which was associated with intrauterine growth restriction. For example, during the first week of life, glycine level was found to be significantly higher in the urine of infants with IUGR than those with normal growth in utero.

Two studies have used mass spectrometry of serum samples to examine metabolite changes associated with differing patterns of ex utero growth.[279, 280] Younge and co-workers identified a maturation signature using a random forests modelling approach.[279] They could then demonstrate that infants exhibiting slower ex utero growth had a slower rise in this metabolic maturity score. Dudzik and co-workers trained a modified partial least squares model to differentiate between discharge samples from infants with extrauterine growth restriction (defined as weight below the 10th percentile at discharge) compared to appropriately grown infants.

Recent studies have examined metabolomic changes over time, alongside interrogation of the influences of nutritional intake on the metabolome.[281, 282] A Mexican group identified differences between the urinary metabolome (assessed by proton NMR) of infants being fed enterally and parenterally using a partial least squares method.[282] Unsupervised principal component analysis (PCA) provided partial differentiation between the groups. Nilsson and co-workers[281] identified differences in urinary proton NMR spectra (and certain metabolites) associated with gestational age (in the first week of life) and with postnatal age. They also found differences in the metabolome in response to different nutrition type (i.e. enteral vs parenteral nutrition).

In summary, there have been a number of studies into the metabolomics of preterm infants during recent years. Whilst, in general terms, each study has found differences in the metabolomic profile dependent on postmenstrual or postnatal age, it is unclear whether individual metabolomic signatures can be validated across different studies.

Metabolomic Methods

Metabolite profiling is usually performed by nuclear magnetic resonance spectroscopy (NMR) or by mass spectrometry (MS). Each of these methods comes with advantages and drawbacks and both are described below.

Nuclear Magnetic Resonance Spectroscopy

NMR analysis of metabolites was developed in the 1960s utilising advances in electromagnet technology and the mathematics of the Fourier transform.[274] It takes advantage of the properties of nuclei which have an odd number of protons and/or an odd number of neutrons. As these nuclei move (or spin) they generate a magnetic field with opposing north and south ends. This magnetic moment (i.e. a direction of magnetism across the nucleus) is known as spin.

During NMR spectroscopy, a strong magnetic field is applied to the sample. This causes the protons to align their magnetic fields with the magnet's field. These protons can either align in the same direction as the magnetic field (the more common situation, in which case they are in a lower energy state) or against it (in a high energy state). A stronger magnetic field will generate a larger difference between these two energy states (ΔE). Once protons have been aligned to the magnetic field, radio frequency radiation is pulsed into the sample. This radiation is absorbed only if the energy of the radio waves is equal to ΔE. The energy of a radio wave is proportionate to its frequency, meaning that only one frequency of radio wave will be absorbed at a given ΔE. When energy is absorbed, some low energy protons will be raised to the high energy state (a "spin flip"), until there is an equal mixture of high- and low-energy protons. When the radio frequency energy is stopped, some high-energy protons will return to the low energy state to restore the previous equilibrium. This "relaxation" produces a radio wave signal which can be measured. This radio signal can then be processed by Fourier transformation to produce a spectrum of radio frequencies detected for the sample.

Nuclei surrounded by a large number of electrons will be more "shielded" from the magnetic field than those in low-electron environments. This shielding will cause the nucleus to resonate at a lower frequency. Therefore, protons in different electronic environments will resonate at

different frequencies. Figure 11-1 illustrates this phenomenon taking caffeine (commonly used as a therapy for preterm infants) as an example.

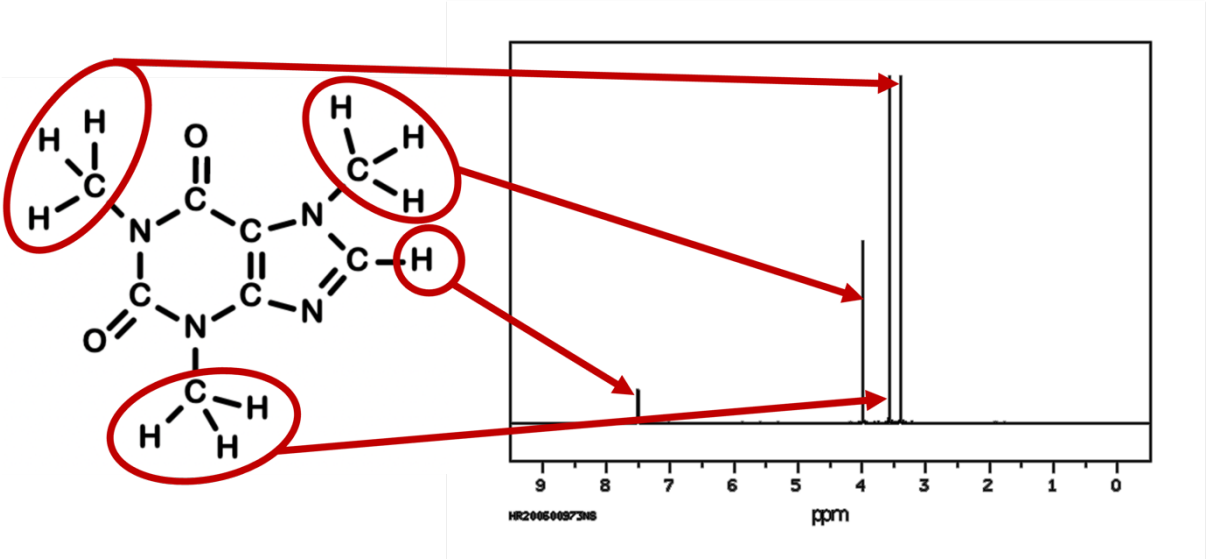


Figure 11-1. Example 1H-NMR spectrum for caffeine, demonstrating the different peaks caused by hydrogen atoms in different electronic environments. Spectrum from chemicalbook.com and molecular diagram produced using chemspider.com.

The spectrum resulting from a mixture of chemicals (such as is found in biological samples), will contain the overlapping features of all the chemicals included in analysis. Depending on the question being asked, the spectrum can be taken in its raw form to discover differences in spectra between different groups (e.g. treatment vs control groups) or the relative abundance of different chemicals can be derived by matching the peaks to reference values for the chemical in question.

Mass Spectrometry

Mass spectrometry (MS) is an alternative approach to NMR for metabolomic analysis. It is considered briefly here as it was not used in this research project. In MS, samples must be converted to gas phase ions prior to analysis. Once ionised, they are fed through the mass spectrometry equipment which calculates their mass-to-charge ratio, the main output of MS analysis. By comparison with reference values, the abundance of different chemicals in the sample can be derived.

Advantages and Disadvantages of NMR and MS

NMR provides metabolomic information quickly (a few minutes per sample) and at a low cost per sample, whereas mass spectrometry is more expensive (after initial equipment purchase) and takes longer for each sample.[283] NMR is also considered to be more reproducible than MS. MS is much more sensitive and is therefore especially useful when seeking to detect chemicals found in very low concentrations. MS also provides more targeted analysis, with NMR profiling often

hindered by the overlapping spectral peaks, even when very sensitive equipment is used. This study used NMR spectroscopy due to its ease of use, its low cost and so that comparison could be made with other studies in preterm infants (see below).

Sample Type

Any liquid can be analysed using NMR techniques. Blood serum or plasma are obvious targets for metabolomic analysis as they are likely to closely reflect the biochemical state of the body as a whole. However, around 500μl of fluid is required for analysis, meaning that at least 1ml of whole blood would be required to acquire sufficient serum or plasma. This is a non-trivial volume of blood for a preterm infant (when the smallest preterms may have a total blood volume of 40ml). Much of the promise of metabolomic profiling is in longitudinally tracking changes in the metabolome over time. Therefore, repeated samples are required from individuals, further increasing the total volumes of blood required and reducing its suitability for assessing preterm infants.

Urine sampling provides a promising alternative to blood-based approaches. As a natural waste product, there is no harm to an infant caused by collecting his or her urine. Studies have shown that urinary metabolites can provide important information about the metabolic status of an infant and has been shown to segregate infants by intrauterine growth status.[99] Collecting urine samples from preterm infants is not entirely straightforward, although there are methods which can be used to achieve good sample collection (see below). Urine was selected as the fluid to be analysed in this project due to its ease of collection and its use in previous research into preterm neonates.

11.2 Hypothesis

I hypothesised that:

- Urine samples could be obtained non-invasively from preterm infants.
- Urinary metabolites (or NMR spectral properties) could be identified which would correlate with the growth of these infants, as recorded in the SPND.

11.3 Aim

This metabolomics element of this doctoral research project aimed to:

1. Build capacity for metabolomic analysis of preterm infants in Southampton;
2. To identify whether developmental trends could be identified in the urinary metabolomic profile of infants; and

3. To begin to assess whether different growth or nutrition patterns contributed to metabolomic differences.

11.4 Methods

11.4.1 Clinical Sampling

Patients were recruited as a subset of infants whose information was held on the Southampton Preterm Nutrition Database. Urine samples were collected as soon as possible after recruitment into the study and then fortnightly. Urine samples were collected in accordance with the ethical approvals for the GAP study (14/SC/1275). The protocol for taking urine samples was drawn up by me, alongside clinical aids to support clinical staff in collecting urine samples.

Urine collection can be difficult in infants as micturition is unpredictable and as urine is usually captured by a nappy, with modern nappies being designed to irretrievably absorb urine into a pad or gel matrix. Various methods are available for urine collection. Urine bags can be stuck to term born infants to collect urine samples but their use is not tested in preterm infants and using adhesives on the underdeveloped skin of preterm infants carries risk and is often ineffective.

In this study, urine was collected by placing balls of cotton wool in the nappy. The nappies of the infants were periodically checked according to normal clinical care practices. When the cotton wool balls were seen to be soaked with urine, they were removed. Soaked cotton wool balls were placed in the barrel of a syringe from which the plunger had been removed. The plunger was then used to squeeze the urine sample out of the cotton wool, through a syringe filter and into a 2ml centrifuge tube. Syringe filters were provided sterile and with a pore size of 0.2μm (Fisherbrand Sterile PES Syringe Filter, product 15206869). This filtration step was included to remove solids from the urine or from the cotton wool and to render the sample free from bacterial contamination for storage. Samples were then labelled with the study number and the date of sampling, were stored at -20°C on the neonatal unit and were periodically transferred to a -80°C freezer. Urine sample acquisition was usually carried out by clinical nurses. Urine sample processing was carried out by me and by research nurses employed by UHSFT.

Urine samples were taken as soon as possible after informed written consent was obtained and then at roughly weekly intervals. When a urine sample was contaminated by stool or was less than 500μl in volume, it was discarded and a new sample was sought.

Weight Gain

A measure of weight gain velocity at the time of urine sampling was required to test for associations between weight gain and the metabolomic spectra.

Daily weight values for each infant were interpolated using cubic b-splines with a knot at every weight value (see Methods chapter). This provided very minimal smoothing of the data, aiming to capture growth velocity over short time periods. Weight gain status at the time of urine sampling was summarised by taking the change in weight z-score (compared to reanalysed UK1990 data[115]) from one week before urine sampling to the day of urine sampling. An example of how this metric was generated is given as Figure 11-2.

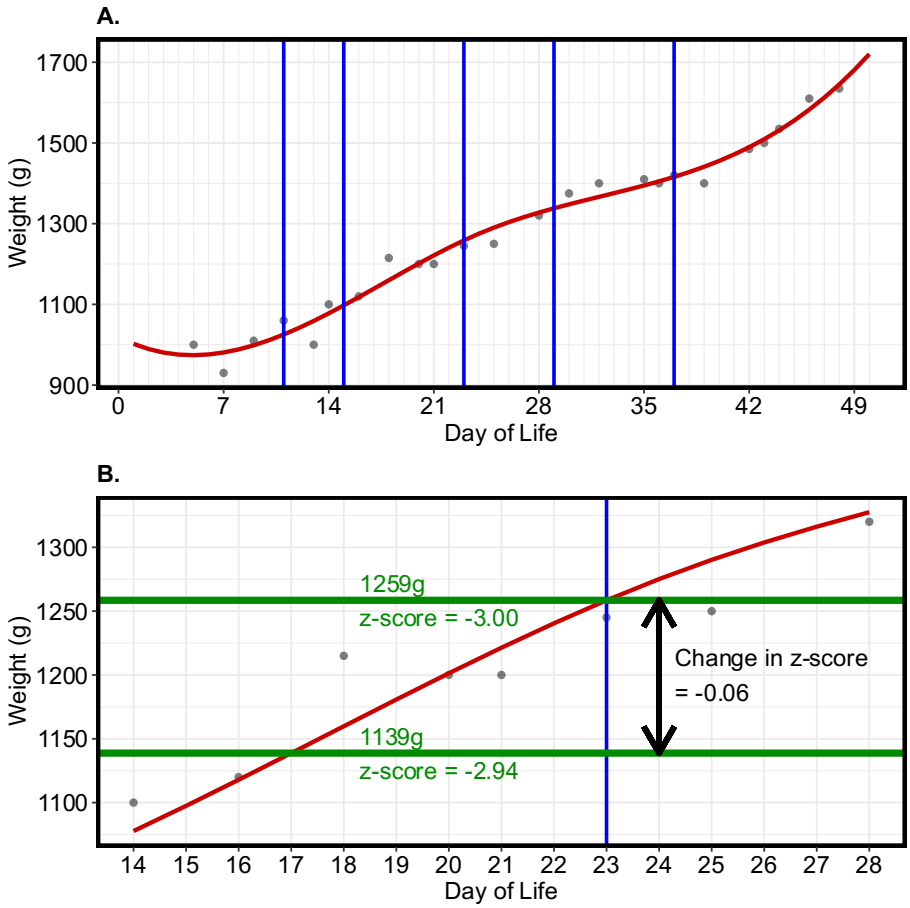


Figure 11-2. A. Weight chart for an example infant with weight values as grey dots, interpolated weight trend by cubic b-splines as a red line and timings of urine sampling indicated by blue vertical lines. B. Example of calculation of weight gain status for the urine sample taken on day 23 for this infant, with green lines indicating the weight values on the day of sampling and one week before sampling and the change in weight z-score demonstrated by black arrow and text.

11.4.2 Sample Processing for NMR

Buffer

A buffer solution is required to control the pH environment and to provide reference NMR peaks. The NMR buffer was prepared by Prof Jonathan Swann. It consisted of 50ml deuterium oxide (D_2O) with: KH_2PO_4 (10.2g, acting as a pH buffer), NaN_3 (6.5mg, to prevent bacterial growth and to maintain comparability to other NMR samples) and deuterated TSP (trimethylsilylpropanoic acid, 50mg). the deuterated TSP provides a reference peak on the resulting NMR spectra which allows each spectrum to be aligned with reference values for interpretation.

Sample Preparation

Sample preparation was carried out by me under the guidance of Prof Swann, with loading into NMR tubes split between me and Prof Swann to expedite the process.

Urine samples were centrifuged at 10,000rpm for 10 minutes so that any solid impurities would form a pellet at the bottom of the Eppendorf tube. A second set of Eppendorf tubes were loaded with 60μl of the NMR buffer solution described above and then 540μl of the urine supernatant formed by centrifugation was added to the buffer. Each sample was vortexed to mix. 575μl of each sample was pipetted into an NMR sample tube (Bruker SampleJet NMR tubes, L 103.5 mm, O.D. 5.0 mm) which were loaded into racks for NMR analysis. Sample placement was randomized prior to loading into sample tubes.

NMR Analysis

After sample preparation, NMR analysis was performed by Prof Swann. Analysis was performed using a Bruker Avance Neo 700 MHz narrow bore four channel NMR spectrometer (Bruker Corporation, Billerica, MA, USA) at the University of Southampton's Multi-Purpose NMR Laboratory with a proton frequency of 700.13MHz. Detection of signals and Fourier transform to generate spectra was handled by the Topspin program (Bruker Corporation). The Topspin software produced a file structure containing spectrum data (as 1r files) along with metadata describing the experimental parameters.

11.4.3 Data Analysis

Processing of Spectra

Analysis of raw spectrum data was performed by me using the AlpsNMR package[284] for R.[111] Spectra were imported and regions with peaks for water and TSP were removed, their positions

having been confirmed by visual inspection of raw spectra. The inbuilt features of the AlpsNMR package were used to identify and exclude outliers by principal component analysis, with confirmation by visual inspection of raw spectra. Peaks were detected using established protocols for identifying peaks by size and morphology.[285] These peak values were then used to align the individual spectra using the cluster-based peak alignment (CluPA) technique[286] which was automated within AlpsNMR.

Probabilistic quotient normalization (PQN) was used to adjust for differences in the overall concentration of the urine samples.[287] In summary, this process uses all the spectral data for each infant to estimate a coefficient for each infant which will bring the overall concentration into concordance with a reference sample. After this step, spectral values essentially reflect the relative abundance of a molecule in the mixture (rather than the absolute concentration).

Statistical Testing

After alignment and normalisation, spectral data were formed into a matrix with each row representing an individual sample and each column assigned to a chemical shift value.

Principal Component Analysis

Spectral data were very high dimensional (with each sample being represented by over 29,000 chemical shift values). Dimensionality was reduced using PCA (see chapter 3 for a description of PCA analysis). A PCA model was created from the aligned and normalised spectral data using the *prcomp* function of the base package of R (with scaling and centring performed on-the-fly using options within the *prcomp* function). Linear regression was then used to assess whether the principal components discovered in the data correlated with postmenstrual or postnatal age of the infants at the time of sampling.

Partial Least Squares Regression

Partial least squares regression (PLS) was used to identify regions of the NMR spectra which were associated with postmenstrual age, postnatal age and weight gain velocity. PLS is an extension of multiple linear regression which aims to deal with highly dimensional and collinear data.[288] It reduces dimensions to a series of components which are defined by their ability to differentiate between predefined outcome states. It therefore reduces dimensionality in a similar way to PCA, but its components are optimised to describe variance in the predefined response variable rather than the variance in the independent variables (as is the case for PCA). PLS regression was carried out using the postmenstrual age, postnatal age and weight gain of infants at the time of sampling as the dependent variables and a matrix of the spectrum values as the independent variables. The

plsr function of the pls package[289] for R was used (with scaling and centring performed within the function as for PCA analysis). PLS models have a risk of over-fitting and cross-validation is required to prevent this. K-fold cross-validation was used. In k-fold cross-validation, the data is separated into subgroups and sequential models are formed, using each subgroup as a test set and the remainder of the dataset as a training set. The results of each of these models is then averaged, providing a final model. Data derived from cross-validation provides measures of the number of components which will best predict the outcome variable without overfitting which would render the results less generalisable. The root mean square error of prediction (RMSEP) provides a measure of the uncertainty of a model when applied to new predictions, with the lowest value representing the most predictive model which is not overfitted.

The PLS model contains information about which variables in the underlying data most strongly contribute to each of the components and these data are known as loadings. The loading data for PLS models were interrogated to identify regions within the spectra which were most strongly associated with postnatal age or weight gain.

Metabolite Concentration Determination

Concentrations of individual metabolites were estimated by me using manual alignment of reference spectral signatures using Chenomx NMR Suite (v9.02, Chenomx Inc., Edmonton, Canada). Once peaks were aligned with the reference spectrum, the concentration was recorded as calculated by the software program by integration. The concentration of target metabolites was adjusted to the overall concentration of the urine sample using the PQN quotients generated in the normalisation step described above. Therefore, the values used for metabolite concentration were relative concentrations rather than absolute concentrations in the urine samples. Relationships between the relative concentration of the metabolite and dependent variables including postnatal age were assessed using simple linear regression.

11.5 Results

Urine Sampling

Urine sampling was ongoing at the time of writing this thesis. A subset of urine samples was selected for interim analysis by me. This subset comprised 93 samples from a total of 15 infants, selecting those for whom genomic information was available and then from whom the most samples had been collected and who had been discharged from the study at the time of analysis. Figure 11-3 illustrates the time points of each of the samples taken in terms of the postnatal age

(A) and the postmenstrual age (B) of the infant. Only three samples were taken during the first week of life.

Each of these samples was of an appropriate volume for analysis and was analysed by ^{1}H-NMR as described above.

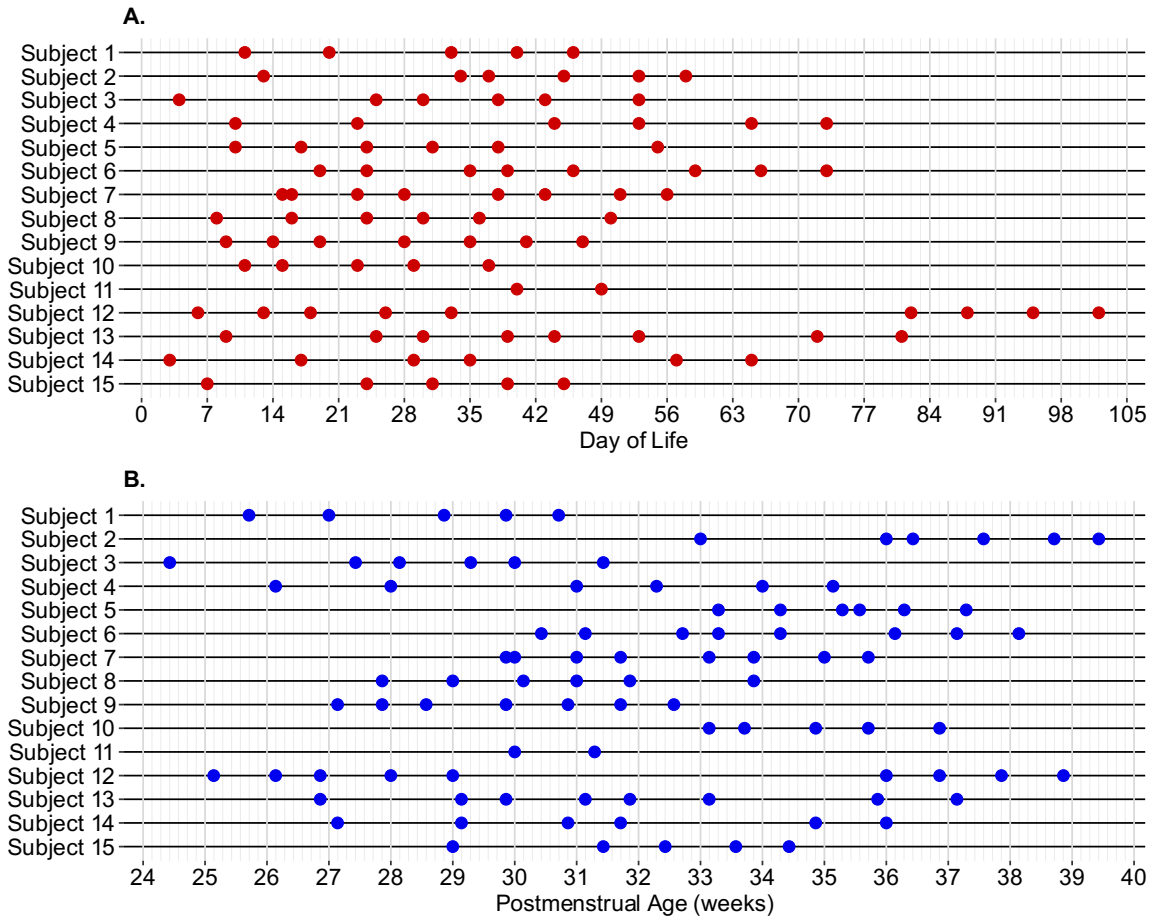


Figure 11-3. Timing of urine samples (coloured dots) for metabolomic analysis by A. postnatal age and B. postmenstrual age.

Processing of Spectra

Visual examination of spectra confirmed that the peaks for water (chemical shift 4.75 to 4.94) and TSP (-0.05 to 1) were in the expected positions and these regions were excluded from further analysis.

PCA outlier analysis using the AlpsNMR package identified one sample as a significant outlier (sample 91 as depicted in Figure 11-4-A). Visual examination of the raw spectra confirmed that this spectrum had a large peak which was not a plausible NMR signal (Figure 11-4-B) and so this sample was removed from further analysis.

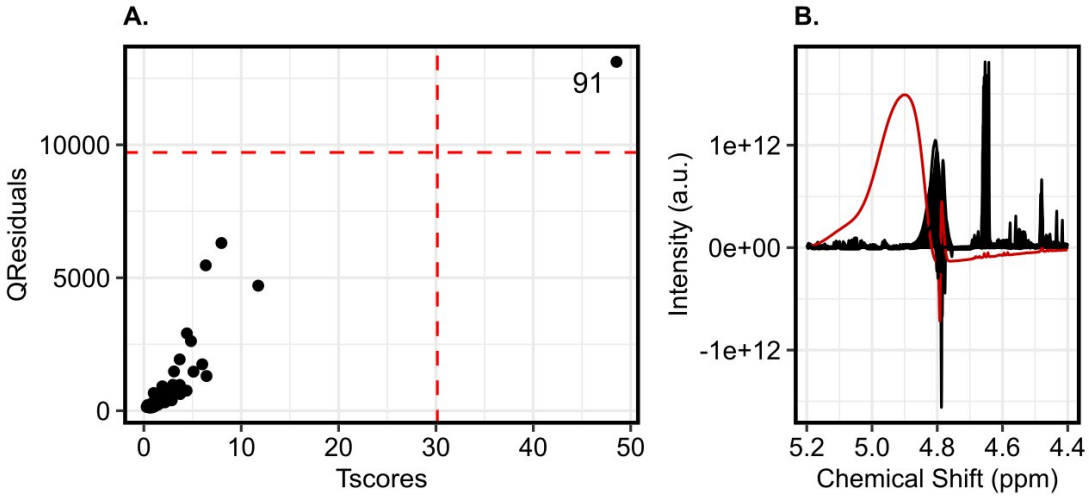


Figure 11-4 Outlier detection using A. PCA analysis within the AlpsNMR package, identifying sample 91 as an outlier; and B. by examination of all spectra (black lines) with a large abnormal peak in sample 91 (red line). PCA – principal component analysis.

The peak identification protocol defined plausible peaks based on analysis the most representative spectrum (sample 21) (Figure 11-5-A). Spectra were visibly much more closely aligned to one another after the alignment step, confirming the success of the automated alignment by the CluPA technique (Figure 11-5-B). PQN successfully unified overall concentration across the samples.

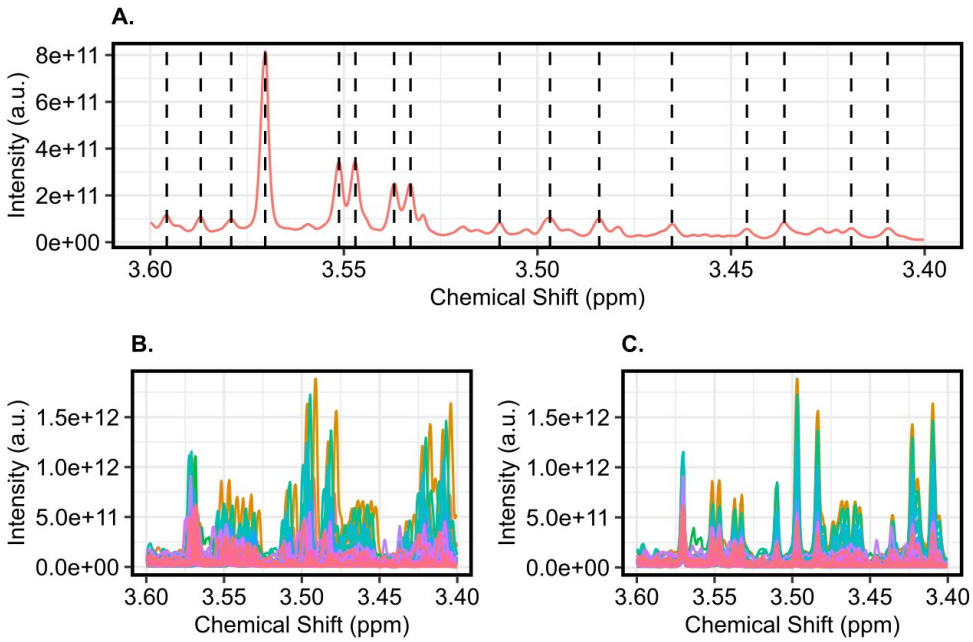


Figure 11-5. A. Peaks identified within a sample chemical shift range. Raw spectra of all infants B. prior to alignment and C. after alignment by CluPA.

Principal Component Analysis

A PCA model was produced using all the spectral intensity values. The first principal component explained 27% of the variance in spectral data, and the first five principal components explained a cumulative total of 47% of the variance (Figure 11-6). This confirms that PCA managed to explain a

considerably proportion of the variance between samples whilst very significantly reducing the dimensionality of the data.

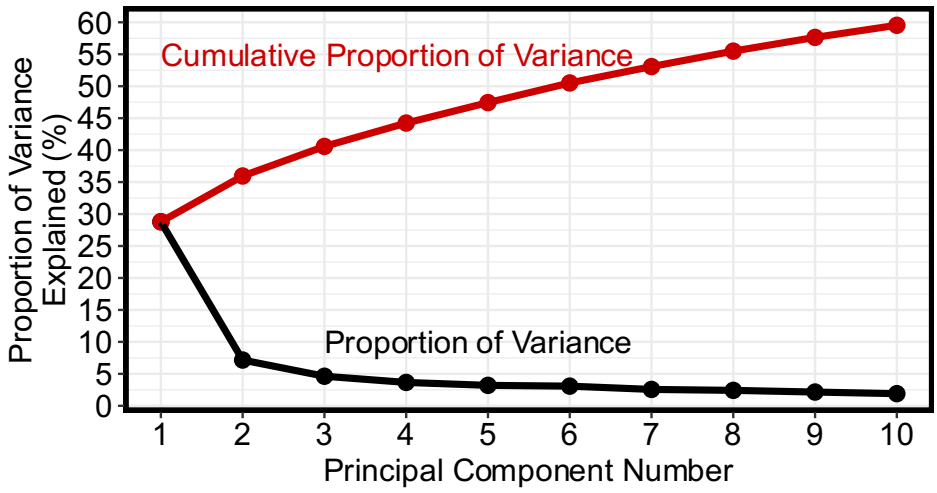


Figure 11-6. Scree plot of PCA analysis of NMR spectra for 92 infants, with the proportion of variance explained by each component in black and the cumulative total of the explained variance in red.

11.5.1 Correlation Analysis of PCA Components with Postnatal Age and Postmenstrual Age

Maturation

The first three principal components significantly correlated with postnatal age at time of urine sampling. Initial regression models using these components identified sample 15 as an extreme outlier. After removal of sample 15, a multiple linear regression model was built using the values for the first three principal components. This model had an adjusted R^2 value of 0.29 indicating that the model explained 29% of the variance in postnatal age. Root mean square error (RMSE) was 17.6. Principal component 2 (PC2) was by far the most closely correlated with postnatal age, explaining 22% of the variance in a simple linear regression model using only PC2 ($p=2 \times 10^{-6}$).

The first two principal components were also significantly correlated to postmenstrual age (PMA). However, these correlations were much weaker than for postnatal age, with a multiple linear regression model for postmenstrual age using the first two principal components explaining only 9% of the variance in PMA. When postnatal age was included as a predictor in this model, no principal components were statistically significantly associated with PMA, indicating that the association of the principal components with PMA was confounded by postnatal age. Including subject identity as a random effect in a mixed effects model did not change the associations.

Having identified a correlation between the first three principal components and postnatal age, a random forest model was created to examine whether this technique would better predict postnatal age from these principal components. For this analysis, samples were split into a

training set (consisting of 60 samples) and a testing set (31 samples). A random forest model was created using the training set and taking the first three principal components as independent variables and the postnatal age as the dependent variable, with 10-fold cross-validation. When applied to the testing set, this model performed less well than multiple linear regression, with an R^2 value of 0.18 and an RMSE of 20 days.

Weight Gain

The second principal component was also correlated with weight gain during the week prior to urine sampling (p=0.02 after adjustment for postnatal age). Including subject identity as a random effect in a mixed effects model did not change the association. No other principal components showed a significant correlation with weight gain.

11.5.2 PLS Regression

Maturation

PLS regression was performed using postnatal age as the dependent variable with 10-fold cross-validation. Figure 11-7-A shows that the lowest value for RMSEP was at two components, with models using more components overfitted to the underlying data. For models taking postmenstrual age as the dependent variable, models were overfitted when more than one component was used (Figure 11-7-B).

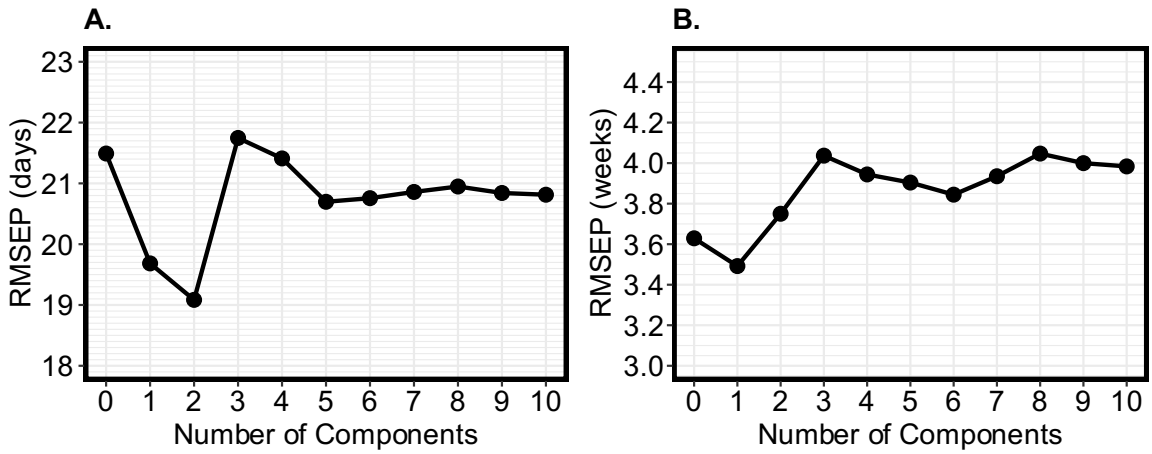


Figure 11-7. Root mean square error of prediction (RMSEP) for PLS regression for spectrum values as a predictor of A. postnatal age and B. postmenstrual age.

The PLS model using two components explained 47% of the variance in postnatal age, outperforming methods using regression unsupervised PCA components. This association between the values of PLS components and postnatal age suggests that there are spectral regions which change as preterm infants get older. The metabolites associated with spectral regions were identified and analysed as set out below. PLS models using postmenstrual age explained 15% of

the variance in postmenstrual age in the group, suggesting that NMR spectra were less predictive of postmenstrual age.

Weight Gain

PLS regression was performed to explore correlations between NMR spectra and weight gain status at the time of urine sampling. Two components were identified as optimal, using RMSEP as described above. A model using these components explained 56% of the variance in weight gain, allowing the loadings from these PLS analyses to identify spectral regions (and associated metabolites) which were associated with faster weight gain.

11.5.3 Metabolite Concentrations

Maturation

Examination of PLS models indicated several areas of the spectra which were strongly related to postnatal age. Manual examination of reference spectra within Chenomx NMR Suite identified myo-inositol as a contributor to peaks at several of the most significant spectral locations. Therefore, the concentration of myo-inositol was manually estimated by alignment within Chenomx NMR Suite. Figure 11-8 gives an example of alignment of the reference spectrum for myo-inositol with a spectrum from one of the samples. Values were normalised using the PQN quotients.

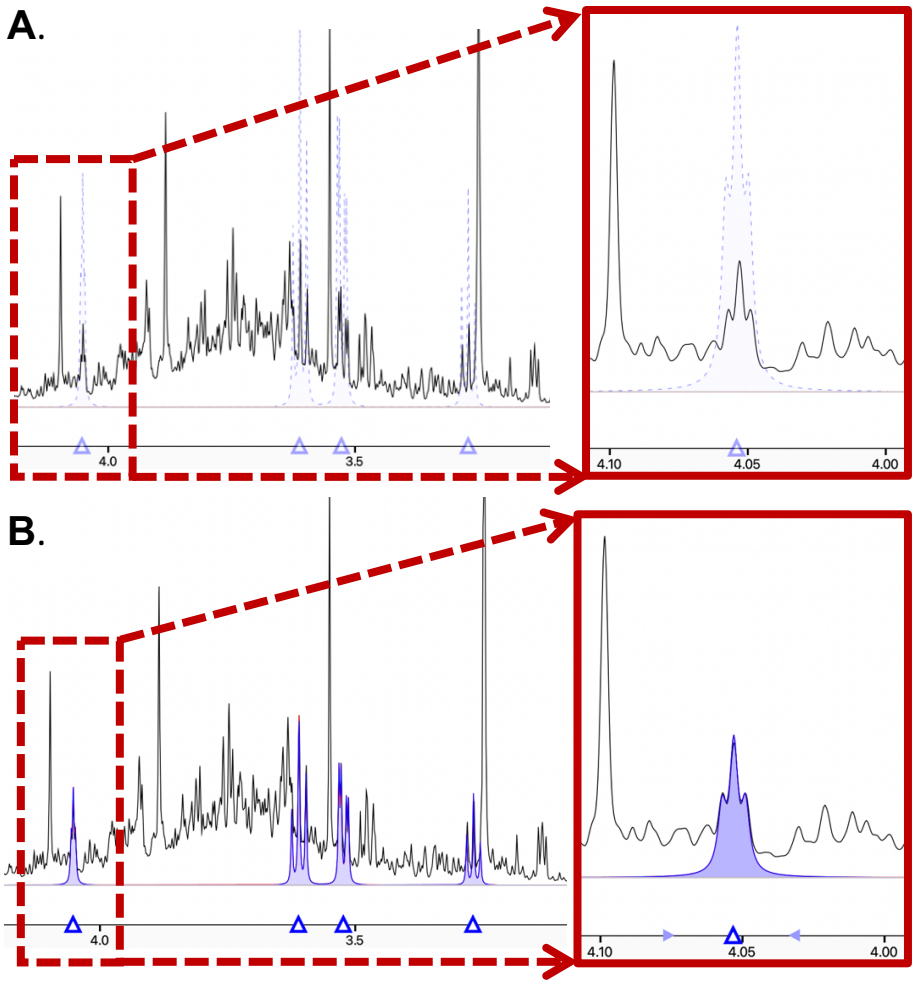


Figure 11-8. Example proton-NMR spectrum from a urine sample A. prior to myo-inositol alignment with the reference spectrum for myo-inositol shown with dotted blue line (with part of the spectrum magnified) and B. after alignment of the myo-inositol reference spectrum.

Simple linear regression confirmed a negative association between relative myo-inositol concentration and postnatal age ($p=0.003$). As postnatal age increased, dietary inositol intake increased ($p<0.001$) whilst urine myo-inositol level decreased, suggesting that the negative association between postnatal age and urinary myo-inositol was not confounded by dietary intake. Urine myo-inositol was also negatively associated with postmenstrual age, but this effect was not present after adjustment for postnatal age.

Weight Gain

Examination of PLS models for weight gain suggested that chemical shifts around 8.5 and 2.91 may have been important. The 8.5 region reflects aromatic compounds which were not well characterised by the NMR spectra. The 2.91 region contained a clear peak, which may have been associated with *N,N*-dimethylglycine. Therefore, the normalised *N,N*-dimethylglycine concentration was estimated using the same process as for myo-inositol in section 10.4.6.1 above.

There was a significant positive correlation between normalised urinary *N,N*-dimethylglycine and weight gain during the week preceding urine sampling (p=0.009). This relationship was not confounded by postnatal age, postmenstrual age or by inclusion of subject identity as a random effect.

11.6 Discussion

The metabolomic section of this doctoral project established ethical approval for urine sampling and analysis, along with developing an effective sampling and analytical pathway in Southampton. Despite the relatively small number of infants included in analysis, there were sufficient data to generate hypotheses for metabolites associated with maturation and with weight gain status.

Of 93 samples included in initial analysis, 92 yielded usable proton-NMR spectra. Very few urine samples were taken during the first week of life, significantly limiting this study's ability to describe the changes in the urinary metabolome of preterm infants during postnatal maturation. This paucity of data during the first week was caused by the timing of taking consent from parents. During the first week of their child's neonatal stay, parents receive a very large amount of information and are often in a state of profound distress; there is an understandable urge to resist overwhelming them with research information during this time. Furthermore, many interventional trials require parental consent within the first few days of life, and so consent for those studies is prioritised by research staff.

Urine sampling is non-invasive and does not require disturbance of the vulnerable infant in excess of that required by normal neonatal care. Therefore, it may be acceptable for urine samples to be gathered and stored (but not analysed) prior to parental consent, which would be deferred. There has been increasing interest in using deferred consent in neonatal studies.[290, 291] This approach could potentially significantly improve the density of metabolomic data during the first week of life if ethical approval were granted.

Metabolomic Changes During Maturation

Principal component analysis of spectra identified principal components which summarised much of the variance in spectra whilst reducing the dimensionality of the data. The first three principal components correlated with the postnatal age, indicating that the unsupervised PCA analysis identified some factors which changed during the postnatal maturation of preterm infants. Other groups have used similar techniques to define a "metabolic maturity age" with relation to postmenstrual age.[279, 292] The postnatal age PCA and PLS models developed during this project could provide a means to examine the differing rates of metabolomic maturation amongst infants exposed to different interventions, including nutritional approaches. The number of infants

included to date precludes further detailed analysis. Models using postmenstrual age were less successful.

A possible negative correlation was identified linking postnatal age to relative myo-inositol concentration in the urine, suggesting myo-inositol as a marker of the postnatal maturation of preterm infants. This finding is consistent with the identification of a similar relationship by Moltu and co-workers.[99]

Myo-inositol has been intensively studied in preterm infants, with its greater abundance in the serum of preterm infants compared to term-born infants having been discovered several decades ago.[293] That study also identified that a rapid drop in serum myo-inositol was associated with more severe respiratory distress syndrome. Several clinical trials were performed to assess whether inositol supplementation could improve respiratory outcomes, but meta-analysis did not show an effect.[294] Therefore, the finding in this study of declining urinary myo-inositol during the postnatal life of preterm infants is not a new discovery. However, it is reassuring that an agnostic process using Southampton data identified a metabolite which is known to change in this way over time, including in a modern study using similar techniques.

Metabolomic Associations with Weight Gain

PCA and PLS models identified spectral signatures associated with faster weight gain. These findings can be compared to those of Dudzik and co-workers.[280] That group determined plasma metabolomic signatures using mass spectrometry near discharge and examined them for associations with low body weight at the point of discharge. They used orthogonal partial least squares discriminant analysis (a similar process to PLS) and identified a metabolomic signature which discriminated between smaller and larger infants. They also explored individual metabolites as markers of low weight at discharge, identifying a significantly lower concentration of serum *N,N*-dimethylglycine in infants with lower weight at discharge. That finding is in good agreement with the identification in this study of a positive association between urinary *N,N*-dimethylglycine and weight gain velocity. Dimethylglycine is involved in choline metabolism. Its role in the growth of preterm human infants is unknown. One animal study demonstrated increased weight gain in low birthweight piglets who were supplemented with *N,N*-dimethylglycine.[295]

11.7 Conclusions

Collection of urine samples, NMR analysis and data processing pathways for metabolomic data have been established. Several metabolites have been identified as candidate markers of maturation and growth velocity in preterm infants.

The findings of the metabolomics element of this research project were limited by the small number of infants included and by the paucity of sampling during the first week of life. Interpretation of NMR spectra to extract reliable metabolite concentrations for many metabolites in parallel is a very complex process requiring significant training and expertise. Therefore, examination of individual metabolites was very restricted.

Using the Southampton Preterm Nutrition Database (SPND), it was possibly to quickly generate a complex metric of weight gain velocity timed to coincide with urine sampling. This allowed for exploration of associations between metabolomic spectra and growth velocity. The SPND also tracked detailed daily nutrient intakes, including for inositol, allowing the correlation between myo-inositol and postnatal age to be adjusted for dietary inositol intake. This approach used untargeted assessment of urinary NMR spectra to identify metabolites of interest, followed by targeted analysis of the dietary intake of inositol. Future approaches may leverage the SPND to assess associations between nutrient intakes, growth and metabolomic profiles, without requiring identification of a target nutrient. Urine samples will continue to be collected and integration of these data with the SPND will permit further insights into relationships between the metabolome, nutrition and growth of preterm infants.

Chapter 12 Discussion

12.1 Overview

In this thesis I have:

- Established the Southampton Preterm Nutrition Database as a comprehensive clinical database containing information on the demographic, clinical, nutritional, genomic and metabolomic status of over 600 preterm infants.
- Described the local and national growth of preterm infants in detail, presenting two alternative means of generating growth charts relevant to preterm infants.
- Identified macronutrient intakes associated with faster growth of preterm infants.
- Established the current understanding of changes in total body water of preterm infants as they grow, and taken some measurements of total body water in growing preterm infants.
- Established pathways to collect and analyse genomic and metabolic information from growing preterm infants.
- Identified candidate metabolites as markers of metabolic maturation and growth velocity in preterm infants.

12.2 Research Findings

This thesis provides information required to underpin a comprehensive toolkit to monitor and manage the growth of preterm infants during their initial hospital stay. The Southampton Preterm Nutrition Database provides a platform for further investigation of the influences on preterm growth and genomic and metabolomic methods have been developed which can be integrated with these clinical data.

Meaningful growth charts are key to monitoring the growth of preterm infants. Current growth charts reflect fetal growth patterns and it is unclear to what extent postnatal growth of preterm infants should follow this pattern. In this thesis, I have presented two alternative means of generating growth charts relevant to preterm infants. The first option was created using the birth measurements and subsequent measurements of preterm infants cared for in Southampton and managed according to current nutritional guidance. These charts have the benefit of reflecting achievable growth in infants who are mostly well-nourished (by current international consensus guidance) but they imply that growth should be similar at different postnatal ages, i.e. that a newborn infant at 31 weeks of gestation should grow in the same way as a four-week-old infant

born at a gestation of 27 weeks. This is counter-intuitive; the SITAR lines I generate for this thesis confirm that infants usually experience an initial period of weight loss or very slow weight gain, followed by a phase of accelerating weight gain. This phenomenon is likely to be at least partly physiological and a result of water loss.[296] The second set of charts presented in Chapter 6 and online at bit.ly/preterm-plotter take this early weight pattern into account, by presenting an individualised growth chart for the gestation and birthweight centiles of the infant in question. These charts were largely inspired by those created by Landau-Crangle and co-workers.[34] Whereas Landau-Crangle leveraged physiological discoveries and targeted a trajectory which would converge with the infant's birth centile a at 42 weeks PMA with her charts, the charts I present are based on the observed growth trajectories of infants with favourable medium-term neurodevelopmental outcomes. My charts do not make assumptions about the ideal weight gain pattern in early life and do not rely on assumptions about early weight loss which are themselves reliant on data gathered from a population of infants who may not have had optimal fluid management at the time of birth. They can be implemented as part of a nutrition toolkit to track growth and to alert clinicians to growth which is faltering below that which is seen in infants who have a positive outcome.

I used detailed data from infants in Southampton to discover nutritional intake patterns associated with faster weight gain and head growth. Whilst collinearity prevented me from assessing the impact of micronutrients, machine learning techniques identified some key observations about macronutrient intake. These data suggest that clinicians should aim for a protein intake of 3.5g/kg/day, energy intake of 125kcal/kg/day and fat intake of 4g/kg/day, with intakes in excess of these values failing to promote faster growth. These findings are in good agreement with current ESPGHAN guidelines. Investigations into the link between protein intake and serum urea in Chapter 7 did not identify a workable way that this information could be used to guide clinical care. Examination of growth metrics in response to feeding approach suggested that a brisk transition from parenteral to enteral feeding promotes better growth than receiving a mixture of the two for a significant period of time. These data provide a foundation for clinicians making feeding decisions about infants in their care.

I used English national data to identify some demographic and perinatal factors which influence weight gain. These data confirm that the most preterm infants and those born small for gestational age are at risk of poorer growth. They also show that girls grow more slowly than boys (even when expressed in z-score terms, correcting for the different average sizes of the sexes) and that a need for resuscitation and ventilation at birth are associated with slower growth (likely reflecting general physiological instability). Early use of parenteral nutrition and birth in a unit providing intensive care were associated with faster weight gain, providing some support for recent efforts to centralise births of preterm infants and to initiate parenteral nutrition as soon as

possible. The other key modifiable element of care is the nutritional intake provided to preterm infants.

Real-time tracking of body composition in preterm infants remains dishearteningly out of reach. The systematic review in Chapter 8 confirms that preterm infants have a significantly higher total body water percentage at birth than infants born at full term, but the postnatal pattern of body composition change in preterm infants remains largely undefined. Efforts in this PhD to describe that pattern more closely were impaired by the COVID-19 pandemic and by equipment failures. In addition, results which were available from the pilot phase of that study were difficult to interpret, often providing implausible values. These issues highlight the difficulty of measuring body composition in this patient group. They also stymied plans to assess the correlation between limb circumference measurements and body composition. The growth charts presented in Chapter 6 demonstrate that it is feasible to track limb circumference growth. However, it is difficult to assess how meaningful these data are in the absence of evidence that limb circumference growth is related to body composition.

The genomic and metabolomic elements of this thesis were primarily carried out to demonstrate the feasibility of collecting such information and to build capacity for ongoing work in the department. The acquisition of these data provides a model for storing multi-omic data alongside clinical information. Examination of urinary metabolites identified dimethylglycine as a candidate marker for improved growth, suggesting that the choline pathway may merit further investigation. This system would also provide a platform for recording other high-density data, with stool microbiome information likely to be useful as a marker of gut health.

The establishment of the Southampton Preterm Nutrition Database will be a lasting legacy of this doctoral project. It provides validated, structured information on a wide range of clinical and nutritional factors. It is flexible enough to integrate additional fields if these are required to support further research. At the time of submission, work on the database continues and is focused on further automation of the process, especially as it relates to nutritional intake. This process will use data from the detailed clinical information system in use in Southampton. The structure of the database would also provide a means by which even more information could be gathered from the clinical information system, including minute-to-minute measurements of vital observations and a comprehensive record of all medications used. In its current form, the database is limited to infants receiving care in Southampton. The implementation of clinical information systems in other units provides an opportunity to broaden the inclusion of infants from elsewhere, providing more generalisable insights into the factors influencing growth during neonatal care.

12.3 Limitations

This thesis uses observational data to seek insights into the factors influencing growth. Observational approaches are prone to confounding and to reverse causation, and the work included in this thesis is not an exception. For example, it is plausible that there is a cohort of "healthy" preterm infants who are destined to grow well. These infants are likely to receive and tolerate nutritional intakes close to those recommended by international bodies. There is a danger that the growth of these infants is identified as having been caused by their nutritional care, whereas in reality their general lack of severe comorbidities leads to good growth and high nutrient intakes (i.e. the association between nutrition and growth is confounded). Where possible, I have aimed to adjust for this effect by including markers of overall morbidity in models using regression or machine learning techniques to assess associations, but residual confounding cannot be excluded. The complex interaction of growth and clinical decision-making around nutrition is less amenable to mathematical adjustment. There remains a possibility that prior growth of the infant influences nutritional decision-making and that subsequent growth appears to reflect the nutritional care provided but actually reflects an ongoing pattern of growth.

The data in this work is also mostly limited to short-term outcomes, especially growth parameters. Growth is not, in itself, an important outcome for neonatal care, although its association with later neurodevelopment suggests that it may be a useful short-term surrogate marker. A recently developed list of core outcomes for neonatal research did not include growth, but did include the broad concepts of gross motor function, general cognitive ability and quality of life.[297] Future work will need to focus on these outcomes as they relate to growth and nutrition during the neonatal period.

12.4 Implications for Practice

Figure 12-1 illustrates the way that the findings from this thesis could be implemented to support the parallel management of growth in a preterm infant and observational data gathering for research.

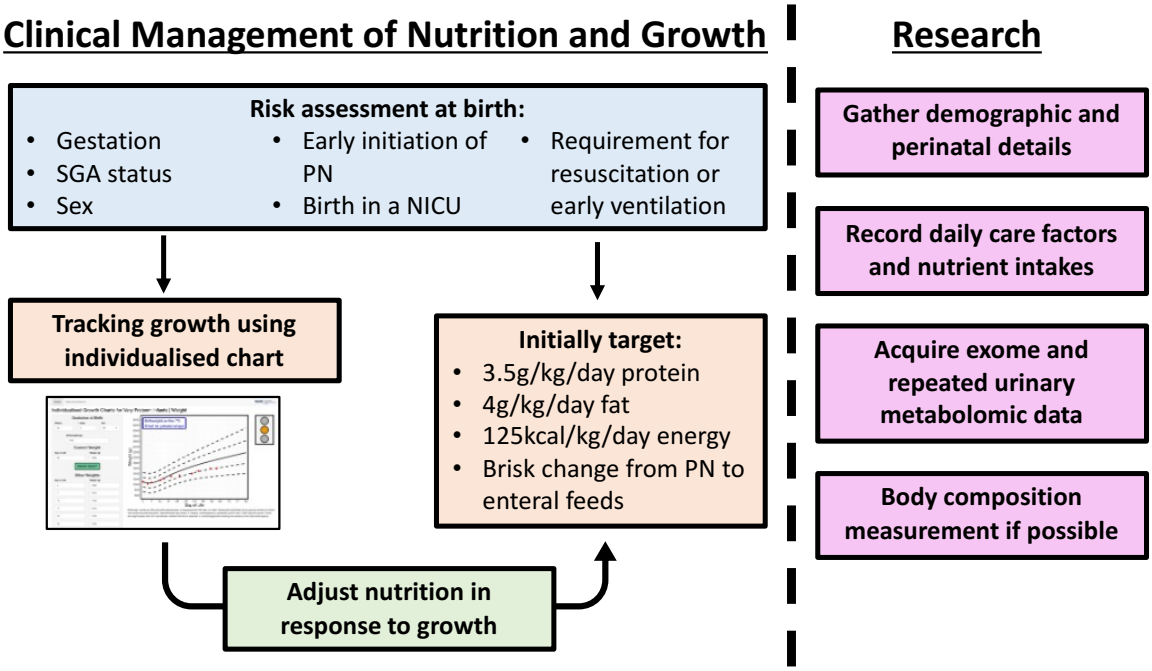


Figure 12-1. Proposed use of findings and methods from this thesis in the parallel clinical care and research into growth of preterm infants. PN – parenteral nutrition, NICU – (Level 3) neonatal intensive care unit

More work is required to prepare these outputs for clinical implementation. As set out in Chapter 6, individualised growth charts are likely to need to take account of a wider range of outcome measures to improve confidence that they do not represent a pattern of growth which would have adverse metabolic consequences. There remains some risk that the listed macronutrient targets reflect the provision of guideline-compliant nutrition to infants who are thriving already. Nevertheless, the work of this thesis provides a structure for a toolkit in which infants can be assessed for their risk of poor growth, start a course of nutrition which is expected to be sufficient and then have their growth tracked against a meaningful standard.

Currently, the SPND is updated periodically to retrospectively integrate data from its sources. Therefore, its use is limited to observational research applications. With the advent of comprehensive clinical information systems for neonatal use, there is a potential for the sorts of information gathered in the SPND to be made available to the clinician in real time to guide clinical care. The Metavision system in use in Southampton can already report the intake of macronutrients for the previous day to the user. Integration of more complex data could allow the system to highlight infants at the highest risk of growth failure, to provide an overview of the nutrition given over a longer time period and to identify suboptimal growth patterns more clearly.

12.5 Future Work

As well as the concrete outcomes detailed above, the work of this doctoral thesis has prepared me for my future academic career. It has provided me with the knowledge and skills to extend this work into the future.

I aim to continue research along four workstreams:

1. Developing the Southampton Preterm Nutrition Database (SPND).
2. Working towards clinical markers of body composition.
3. Expanding multi-omics data collection and analysis.
4. Developing a toolkit to support the growth of preterm infants.

The SPND currently provides a robust platform for gathering and analysing information on the growth of preterm infants in Southampton. With the introduction of a comprehensive clinical information system, there is scope for both additional automation of data gathering and of an expansion of the data collected. In the first instance, I am working with the clinical information system team in Southampton to automatically gather nutrient intake data from the system, reducing the workload of manually inputting feed and fluid data directly. In the future, the minute-to-minute observations of each infant could also be captured. It would be useful to be able to expand the use of the SPND to include infants from other neonatal units. As other units take on clinical information systems, I will investigate how their data could be included.

I have collected deuterium dilution samples from a further cohort of infants but the analysis of these samples has been delayed by our collaborator's equipment failure. This arm of the study was also delayed by the COVID-19 pandemic in that research nursing time was withdrawn and laboratory staff were unable to work. When these data are available, there may be sufficient information to assess whether limb circumference measurements are capable of tracking body composition in real time.

Collection of blood and urine for genomic and metabolomic analysis continues in Southampton. The analysis of these samples will provide further opportunities to investigate the links between the genome, the metabolome and the growth and maturation of preterm infants. The metabolomics department of the University of Southampton is expanding and there will be opportunities to leverage the expertise of new members of that team to discover new insights. Characterisation of the gut microbiome is missing from the current dataset and I will need to build collaborations to include these data in future projects.

Finally, the information gathered during this thesis and in future projects needs to be shaped and presented so that it can impact the growth and health of preterm infants in the future. Currently, research findings feed into guidelines which are applied broadly across infants with different

phenotypes and different nutritional needs. The individualised growth charts presented in this thesis provide an example of how the phenotype of an infant can be used to generate goals which are appropriate to that individual. Over the next decade, there will be opportunities to integrate research findings, guidelines and detailed phenotyping to inform the personalised management of the nutrition of each individual infant. This doctoral project will place me in a position to contribute to developing those personalised tools to optimise the nutrition, growth and development of these vulnerable infants.

References

1. Young A, Andrews ET, Ashton JJ, Pearson F, Beattie RM, Johnson MJ. Generating longitudinal growth charts from preterm infants fed to current recommendations. Arch Dis Child Fetal Neonatal Ed. 2020;105(6):646-51.
2. Young A, Brown LK, Ennis S, Beattie RM, Johnson MJ. Total body water in full-term and preterm newborns: systematic review and meta-analysis. Arch Dis Child Fetal Neonatal Ed. 2021;106(5):542-8.
3. Young A, Beattie RM, Johnson MJ. Optimising growth in very preterm infants: reviewing the evidence. Arch Dis Child Fetal Neonatal Ed. 2022.
4. Young A, Cole TJ, Cheng G, Ennis S, Beattie RM, Johnson MJ. Changes in the growth of very preterm infants in England 2006-2018. Arch Dis Child Fetal Neonatal Ed. 2022.
5. Marlow N, Wolke D, Bracewell MA, Samara M, Group EPS. Neurologic and developmental disability at six years of age after extremely preterm birth. N Engl J Med. 2005;352(1):9-19.
6. Linsell L, Johnson S, Wolke D, O'Reilly H, Morris JK, Kurinczuk JJ, et al. Cognitive trajectories from infancy to early adulthood following birth before 26 weeks of gestation: a prospective, population-based cohort study. Arch Dis Child. 2018;103(4):363-70.
7. Johnson S, Kochhar P, Hennessy E, Marlow N, Wolke D, Hollis C. Antecedents of Attention-Deficit/Hyperactivity Disorder Symptoms in Children Born Extremely Preterm. J Dev Behav Pediatr. 2016;37(4):285-97.
8. Johnson S, Hollis C, Kochhar P, Hennessy E, Wolke D, Marlow N. Psychiatric disorders in extremely preterm children: longitudinal finding at age 11 years in the EPICure study. J Am Acad Child Adolesc Psychiatry. 2010;49(5):453-63 e1.
9. O'Reilly H, Ni Y, Johnson S, Wolke D, Marlow N. Extremely preterm birth and autistic traits in young adulthood: the EPICure study. Mol Autism. 2021;12(1):30.
10. Marlow N, Ni Y, Lancaster R, Suonpera E, Bernardi M, Fahy A, et al. No change in neurodevelopment at 11 years after extremely preterm birth. Arch Dis Child Fetal Neonatal Ed. 2021;106(4):418-24.
11. Pierrat V, Marchand-Martin L, Arnaud C, Kaminski M, Resche-Rigon M, Lebeaux C, et al. Neurodevelopmental outcome at 2 years for preterm children born at 22 to 34 weeks' gestation in France in 2011: EPIPAGE-2 cohort study. BMJ. 2017;358:j3448.
12. Belfort MB, Rifas-Shiman SL, Sullivan T, Collins CT, McPhee AJ, Ryan P, et al. Infant growth before and after term: effects on neurodevelopment in preterm infants. Pediatrics. 2011;128(4):e899-906.
13. Ehrenkranz RA, Dusick AM, Vohr BR, Wright LL, Wrage LA, Poole WK. Growth in the neonatal intensive care unit influences neurodevelopmental and growth outcomes of extremely low birth weight infants. Pediatrics. 2006;117(4):1253-61.
14. Cordova EG, Cherkerzian S, Bell K, Joung KE, Collins CT, Makrides M, et al. Association of Poor Postnatal Growth with Neurodevelopmental Impairment in Infancy and Childhood: Comparing the Fetus and the Healthy Preterm Infant References. J Pediatr. 2020;225:37-43 e5.
15. Stephens BE, Walden RV, Gargus RA, Tucker R, McKinley L, Mance M, et al. First-week protein and energy intakes are associated with 18-month developmental outcomes in extremely low birth weight infants. Pediatrics. 2009;123(5):1337-43.
16. Parkinson JR, Hyde MJ, Gale C, Santhakumaran S, Modi N. Preterm birth and the metabolic syndrome in adult life: a systematic review and meta-analysis. Pediatrics. 2013;131(4):e1240-63.
17. Bayman E, Drake AJ, Piyasena C. Prematurity and programming of cardiovascular disease risk: a future challenge for public health? Arch Dis Child Fetal Neonatal Ed. 2014;99(6):F510-4.

18. Vaag A. Low birth weight and early weight gain in the metabolic syndrome: consequences for infant nutrition. Int J Gynaecol Obstet. 2009;104 Suppl 1:S32-4.
19. Ekelund U, Ong KK, Linne Y, Neovius M, Brage S, Dunger DB, et al. Association of weight gain in infancy and early childhood with metabolic risk in young adults. J Clin Endocrinol Metab. 2007;92(1):98-103.
20. Ni Y, Beckmann J, Hurst JR, Morris JK, Marlow N. Size at birth, growth trajectory in early life, and cardiovascular and metabolic risks in early adulthood: EPICure study. Arch Dis Child Fetal Neonatal Ed. 2021;106(2):149-55.
21. Embleton ND, Korada M, Wood CL, Pearce MS, Swamy R, Cheetham TD. Catch-up growth and metabolic outcomes in adolescents born preterm. Arch Dis Child. 2016;101(11):1026-31.
22. Menon G, Davidson AL, Drake AJ, Embleton ND. Is preterm nutrition a trade-off between head and heart? Arch Dis Child Fetal Neonatal Ed. 2019;104(3):F232-F4.
23. Agostoni C, Buonocore G, Carnielli VP, De Curtis M, Darmaun D, Decsi T, et al. Enteral nutrient supply for preterm infants: commentary from the European Society of Paediatric Gastroenterology, Hepatology and Nutrition Committee on Nutrition. J Pediatr Gastroenterol Nutr. 2010;50(1):85-91.
24. American Academy of Pediatrics Committee on Nutrition. Nutritional Needs of the Preterm Infant. 2019 [cited 8th February 2021]. In: Pediatric Nutrition [Internet]. Itasca, IL: American Academy of Pediatrics. 8th edition., [cited 8th February 2021].
25. Cole TJ, Statnikov Y, Santhakumaran S, Pan H, Modi N, Neonatal Data Analysis U, et al. Birth weight and longitudinal growth in infants born below 32 weeks' gestation: a UK population study. Arch Dis Child Fetal Neonatal Ed. 2014;99(1):F34-40.
26. Andrews ET, Ashton JJ, Pearson F, Beattie RM, Johnson MJ. Early postnatal growth failure in preterm infants is not inevitable. Arch Dis Child Fetal Neonatal Ed. 2019;104(3):F235-F41.
27. RCPCH. UK-WHO growth charts - neonatal and infant close monitoring (NICM) London, UK: Royal College of Paediatrics and Child Health; 2019. Available from: https://www.rcpch.ac.uk/resources/uk-who-growth-charts-neonatal-infant-close-monitoring-nicm.
28. Cole TJ. The LMS method for constructing normalized growth standards. Eur J Clin Nutr. 1990;44(1):45-60.
29. Fenton TR. A new growth chart for preterm babies: Babson and Benda's chart updated with recent data and a new format. BMC Pediatr. 2003;3:13.
30. Fenton TR, Kim JH. A systematic review and meta-analysis to revise the Fenton growth chart for preterm infants. BMC Pediatr. 2013;13:59.
31. Villar J, Altman DG, Purwar M, Noble JA, Knight HE, Ruyan P, et al. The objectives, design and implementation of the INTERGROWTH-21st Project. Bjog. 2013;120 Suppl 2:9-26, v.
32. Villar J, Giuliani F, Fenton TR, Ohuma EO, Ismail LC, Kennedy SH, et al. INTERGROWTH-21st very preterm size at birth reference charts. Lancet. 2016;387(10021):844-5.
33. Villar J, Giuliani F, Bhutta ZA, Bertino E, Ohuma EO, Ismail LC, et al. Postnatal growth standards for preterm infants: the Preterm Postnatal Follow-up Study of the INTERGROWTH-21(st) Project. Lancet Glob Health. 2015;3(11):e681-91.
34. Landau-Crangle E, Rochow N, Fenton TR, Liu K, Ali A, So HY, et al. Individualized Postnatal Growth Trajectories for Preterm Infants. JPEN J Parenter Enteral Nutr. 2018;42(6):1084-92.
35. Johnson MJ, Wootton SA, Leaf AA, Jackson AA. Preterm birth and body composition at term equivalent age: a systematic review and meta-analysis. Pediatrics. 2012;130(3):e640-9.
36. Ramel SE, Gray HL, Christiansen E, Boys C, Georgieff MK, Demerath EW. Greater Early Gains in Fat-Free Mass, but Not Fat Mass, Are Associated with Improved Neurodevelopment at 1 Year Corrected Age for Prematurity in Very Low Birth Weight Preterm Infants. J Pediatr. 2016;173:108-15.
37. Uthaya S, Thomas EL, Hamilton G, Dore CJ, Bell J, Modi N. Altered adiposity after extremely preterm birth. Pediatr Res. 2005;57(2):211-5.

38. Dyke JP, Garfinkel AC, Groves AM, Kovanlikaya A. High-resolution rapid neonatal whole-body composition using 3.0 Tesla chemical shift magnetic resonance imaging. Pediatric Research. 2018;83(3):638-44.
39. Olhager E, Flinke E, Hannerstad U, Forsum E. Studies on human body composition during the first 4 months of life using magnetic resonance imaging and isotope dilution. Pediatric Research. 2003;54(6):906-12.
40. Roggero P, Gianni ML, Orsi A, Piemontese P, Amato O, Moioli C, et al. Neonatal period: body composition changes in breast-fed full-term newborns. Neonatology. 2010;97(2):139-43.
41. Venkataraman PS, Ahluwalia BW. Total bone mineral content and body composition by x-ray densitometry in newborns. Pediatrics. 1992;90(5):767-70.
42. Godang K, Qvigstad E, Voldner N, Isaksen GA, Froslie KF, Notthellen J, et al. Assessing body composition in healthy newborn infants: reliability of dual-energy x-ray absorptiometry. J Clin Densitom. 2010;13(2):151-60.
43. Forsum E, Olhager E, Tornqvist C. An Evaluation of the Pea Pod System for Assessing Body Composition of Moderately Premature Infants. Nutrients. 2016;8(4):238.
44. Brans YW, Shannon DL, Hunter MA. Maternal diabetes and neonatal macrosomia. III. Neonatal body water estimates. Early Human Development. 1983;8(3):307-16.
45. Cassady G. Effect of cesarean section on neonatal body water spaces. N Engl J Med. 1971;285(16):887-91.
46. Cheek DB, Maddison TG, Malinek M, Coldbegk JH. Further observations on the corrected bromide space of the neonate and investigation of water and electrolyte status in infants born of diabetic mothers. Pediatrics. 1961;28(6):861.
47. Christian JR, Jackson H, Karazeris MP, Larson M, Spafford NR, Talson PJ. Total body water and exchangeable sodium in normal full-term newborn infants. AMA J Dis Child. 1956;92(4):325-9.
48. Thornton CJ, Shannon DL, Hunter MA, Ramamurthy RS, Brans YW. Body water estimates in neonatal polycythemia. Journal of Pediatrics. 1983;102(1):113-7.
49. International Atomic Energy Agency. Assessment of Body Composition and Total Energy Expenditure in Humans Using Stable Isotope Techniques. Vienna; 2009.
50. Tint MT, Ward LC, Soh SE, Aris IM, Chinnadurai A, Saw SM, et al. Estimation of fat-free mass in Asian neonates using bioelectrical impedance analysis. British Journal of Nutrition. 2016;115(6):1033-42.
51. Grijalva-Eternod CS, Wells JC, Girma T, Kaestel P, Admassu B, Friis H, et al. Midupper arm circumference and weight-for-length z scores have different associations with body composition: evidence from a cohort of Ethiopian infants. Am J Clin Nutr. 2015;102(3):593-9.
52. Ashton JJ, Johnson MJ, Pond J, Crowley P, Dimitrov BD, Pearson F, et al. Assessing the growth of preterm infants using detailed anthropometry. Acta Paediatr. 2017;106(6):889-96.
53. Koletzko B, Poindexter B, Uauy R. Recommended nutrient intake levels for stable, fully enterally fed very low birth weight infants. World Rev Nutr Diet. 2014;110:297-9.
54. National Institute for Health and Care Excellence. Neonatal parenteral nutrition. London, UK; 2020 26 Feb 2020. Contract No.: NG154.
55. Mihatsch WA, Braegger C, Bronsky J, Cai W, Campoy C, Carnielli V, et al. ESPGHAN/ESPEN/ESPR/CSPEN guidelines on pediatric parenteral nutrition. Clin Nutr. 2018;37(6 Pt B):2303-5.
56. Baxter Healthcare. Summary of Product Characteristics - Numeta G13%E Preterm emulsion for infusion. electronic medicines compendium. Surrey, UK: Datapharm; 2020.
57. Ziegler EE. Nutrient requirements of premature infants. Nestle Nutr Workshop Ser Pediatr Program. 2007;59:161-72; discussion 72-6.
58. Ziegler EE, O'Donnell AM, Nelson SE, Fomon SJ. Body composition of the reference fetus. Growth. 1976;40(4):329-41.
59. Uthaya S, Liu X, Babalis D, Dore CJ, Warwick J, Bell J, et al. Nutritional Evaluation and Optimisation in Neonates: a randomized, double-blind controlled trial of amino acid regimen

and intravenous lipid composition in preterm parenteral nutrition. Am J Clin Nutr. 2016;103(6):1443-52.

60. Embleton ND, Van Den Akker CH. Early parenteral amino acid intakes in preterm babies: does NEON light the way? Arch Dis Child Fetal Neonatal Ed. 2018;103(2):F92-F4.
61. Morgan C, McGowan P, Herwitker S, Hart AE, Turner MA. Postnatal head growth in preterm infants: a randomized controlled parenteral nutrition study. Pediatrics. 2014;133(1):e120-8.
62. Strommen K, Haag A, Moltu SJ, Veierod MB, Blakstad EW, Nakstad B, et al. Enhanced nutrient supply to very low birth weight infants is associated with higher blood amino acid concentrations and improved growth. Clin Nutr ESPEN. 2017;18:16-22.
63. Roelants JA, Vlaardingerbroek H, van den Akker CHP, de Jonge RCJ, van Goudoever JB, Vermeulen MJ. Two-Year Follow-up of a Randomized Controlled Nutrition Intervention Trial in Very Low-Birth-Weight Infants. JPEN J Parenter Enteral Nutr. 2018;42(1):122-31.
64. Mathur NB, Saini A, Mishra TK. Assessment of Adequacy of Supplementation of Vitamin D in Very Low Birth Weight Preterm Neonates: A Randomized Controlled Trial. J Trop Pediatr. 2016;62(6):429-35.
65. Mactier H, McCulloch DL, Hamilton R, Galloway P, Bradnam MS, Young D, et al. Vitamin A supplementation improves retinal function in infants at risk of retinopathy of prematurity. J Pediatr. 2012;160(6):954-9 e1.
66. Bernhard W, Bockmann K, Maas C, Mathes M, Hovelmann J, Shunova A, et al. Combined choline and DHA supplementation: a randomized controlled trial. Eur J Nutr. 2020;59(2):729-39.
67. Moon K, Rao SC, Schulzke SM, Patole SK, Simmer K. Longchain polyunsaturated fatty acid supplementation in preterm infants. Cochrane Database Syst Rev. 2016;12:CD000375.
68. Staub E, Evers K, Askie LM. Enteral zinc supplementation for prevention of morbidity and mortality in preterm neonates. Cochrane Database Syst Rev. 2021;3:CD012797.
69. Vazquez-Gomis R, Bosch-Gimenez V, Juste-Ruiz M, Vazquez-Gomis C, Izquierdo-Fos I, Pastor-Rosado J. Zinc concentration in preterm newborns at term age, a prospective observational study. BMJ Paediatr Open. 2019;3(1):e000527.
70. Beardsall K. Hyperglycaemia in the Newborn Infant. Physiology Verses Pathology. Front Pediatr. 2021;9:641306.
71. van der Lugt NM, Smits-Wintjens VE, van Zwieten PH, Walther FJ. Short and long term outcome of neonatal hyperglycemia in very preterm infants: a retrospective follow-up study. BMC Pediatr. 2010;10:52.
72. Beardsall K, Thomson L, Guy C, Iglesias-Platas I, van Weissenbruch MM, Bond S, et al. Real-time continuous glucose monitoring in preterm infants (REACT): an international, open-label, randomised controlled trial. Lancet Child Adolesc Health. 2021.
73. Giretti I, D'Ascenzo R, Correani A, Antognoli L, Monachesi C, Biagetti C, et al. Hypertriglyceridemia and lipid tolerance in preterm infants with a birth weight of less than 1250 g on routine parenteral nutrition. Clin Nutr. 2021.
74. Alburaki W, Yusuf K, Dobry J, Sheinfeld R, Alshaikh B. High Early Parenteral Lipid in Very Preterm Infants: A Randomized-Controlled Trial. J Pediatr. 2021;228:16-23 e1.
75. Wang HJ, Hsieh YT, Liu LY, Huang CF, Lin SC, Tsao PN, et al. Use of sodium glycerophosphate in neonatal parenteral nutrition solutions to increase calcium and phosphate compatibility for preterm infants. Pediatr Neonatol. 2020;61(3):331-7.
76. Vermont Oxford Network. Vermont Oxford Network Annual Report 2019. Burlington, VT, USA; 2019 2020.
77. Ramel SE, Brown LD, Georgieff MK. The Impact of Neonatal Illness on Nutritional Requirements-One Size Does Not Fit All. Curr Pediatr Rep. 2014;2(4):248-54.
78. Casaer MP, Mesotten D, Hermans G, Wouters PJ, Schetz M, Meyfroidt G, et al. Early versus late parenteral nutrition in critically ill adults. N Engl J Med. 2011;365(6):506-17.

79. Fivez T, Kerklaan D, Mesotten D, Verbruggen S, Wouters PJ, Vanhorebeek I, et al. Early versus Late Parenteral Nutrition in Critically Ill Children. N Engl J Med. 2016;374(12):1111-22.
80. Excellence NIfHaC. Neonatal parenteral nutrition. London, UK; 2020.
81. Kuschel CA, Harding JE. Multicomponent fortified human milk for promoting growth in preterm infants. Cochrane Database Syst Rev. 2004(1):CD000343.
82. Embleton ND, King C, Jarvis C, Mactier H, Pearson F, Menon G. Effectiveness of human milk-based fortifiers for preventing necrotizing enterocolitis in preterm infants: case not proven. Breastfeed Med. 2013;8(4):421.
83. Grace E, Hilditch C, Gomersall J, Collins CT, Rumbold A, Keir AK. Safety and efficacy of human milk-based fortifier in enterally fed preterm and/or low birthweight infants: a systematic review and meta-analysis. Arch Dis Child Fetal Neonatal Ed. 2021;106(2):137-42.
84. Rochow N, Fusch G, Choi A, Chessell L, Elliott L, McDonald K, et al. Target fortification of breast milk with fat, protein, and carbohydrates for preterm infants. J Pediatr. 2013;163(4):1001-7.
85. Arslanoglu S, Moro GE, Ziegler EE. Adjustable fortification of human milk fed to preterm infants: does it make a difference? J Perinatol. 2006;26(10):614-21.
86. Brion LP, Rosenfeld CR, Heyne R, Brown LS, Lair CS, Petrosyan E, et al. Optimizing individual nutrition in preterm very low birth weight infants: double-blinded randomized controlled trial. J Perinatol. 2020;40(4):655-65.
87. Fabrizio V, Trzaski JM, Brownell EA, Esposito P, Lainwala S, Lussier MM, et al. Individualized versus standard diet fortification for growth and development in preterm infants receiving human milk. Cochrane Database Syst Rev. 2020;11:CD013465.
88. Quigley M, Embleton ND, McGuire W. Formula versus donor breast milk for feeding preterm or low birth weight infants. Cochrane Database of Systematic Reviews. 2019(7).
89. Polberger SK, Axelsson IE, Raiha NC. Urinary and serum urea as indicators of protein metabolism in very low birthweight infants fed varying human milk protein intakes. Acta Paediatr Scand. 1990;79(8-9):737-42.
90. Embleton ND, Cooke RJ. Protein Requirements in Preterm Infants: Effect of Different Levels of Protein Intake on Growth and Body Composition. Pediatric Research. 2005;58(5):855-60.
91. Giretti I, Correani A, Antognoli L, Monachesi C, Marchionni P, Biagetti C, et al. Blood urea in preterm infants on routine parenteral nutrition: A multiple linear regression analysis. Clin Nutr. 2021;40(1):153-6.
92. Horikoshi M, Yaghootkar H, Mook-Kanamori DO, Sovio U, Taal HR, Hennig BJ, et al. New loci associated with birth weight identify genetic links between intrauterine growth and adult height and metabolism. Nat Genet. 2013;45(1):76-82.
93. Freathy RM, Mook-Kanamori DO, Sovio U, Prokopenko I, Timpson NJ, Berry DJ, et al. Variants in ADCY5 and near CCNL1 are associated with fetal growth and birth weight. Nat Genet. 2010;42(5):430-5.
94. Horikoshi M, Beaumont RN, Day FR, Warrington NM, Kooijman MN, Fernandez-Tajes J, et al. Genome-wide associations for birth weight and correlations with adult disease. Nature. 2016;538(7624):248-52.
95. Ryckman KK, Feenstra B, Shaffer JR, Bream EN, Geller F, Feingold E, et al. Replication of a genome-wide association study of birth weight in preterm neonates. J Pediatr. 2012;160(1):19-24 e4.
96. Hamvas A, Feng R, Bi Y, Wang F, Bhattacharya S, Mereness J, et al. Exome sequencing identifies gene variants and networks associated with extreme respiratory outcomes following preterm birth. BMC Genet. 2018;19(1):94.
97. Salis ER, Reith DM, Wheeler BJ, Broadbent RS, Medlicott NJ. Insulin resistance, glucagon-like peptide-1 and factors influencing glucose homeostasis in neonates. Archives of Disease in Childhood - Fetal and Neonatal Edition. 2017;102(2):F162-F6.
98. Blanco CL, Liang H, Joya-Galeana J, DeFronzo RA, McCurnin D, Musi N. The ontogeny of insulin signaling in the preterm baboon model. Endocrinology. 2010;151(5):1990-7.

99. Moltu SJ, Sachse D, Blakstad EW, Strommen K, Nakstad B, Almaas AN, et al. Urinary metabolite profiles in premature infants show early postnatal metabolic adaptation and maturation. Nutrients. 2014;6(5):1913-30.

100. Karlberg J, Albertsson-Wikland K. Growth in full-term small-for-gestational-age infants: from birth to final height. Pediatr Res. 1995;38(5):733-9.

101. Gonzalez-Luis GE, van Westering-Kroon E, Villamor-Martinez E, Huizing MJ, Kilani MA, Kramer BW, et al. Tobacco Smoking During Pregnancy Is Associated With Increased Risk of Moderate/Severe Bronchopulmonary Dysplasia: A Systematic Review and Meta-Analysis. Front Pediatr. 2020;8:160.

102. Mehra R, Shebl FM, Cunningham SD, Magriples U, Barrette E, Herrera C, et al. Area-level deprivation and preterm birth: results from a national, commercially-insured population. BMC Public Health. 2019;19(1):236.

103. Bonet M, Smith LK, Pilkington H, Draper ES, Zeitlin J. Neighbourhood deprivation and very preterm birth in an English and French cohort. BMC Pregnancy Childbirth. 2013;13:97.

104. Nobre RG, de Azevedo DV, de Almeida PC, de Almeida NM, Feitosa FE. Weight-Gain Velocity in Newborn Infants Managed with the Kangaroo Method and Associated Variables. Matern Child Health J. 2017;21(1):128-35.

105. Sladkevicius P, Saltvedt S, Almstrom H, Kublickas M, Grunewald C, Valentin L. Ultrasound dating at 12-14 weeks of gestation. A prospective cross-validation of established dating formulae in in-vitro fertilized pregnancies. Ultrasound Obstet Gynecol. 2005;26(5):504-11.

106. Marks KA, Reichman B, Lusky A, Zmora E, Israel Neonatal N. Fetal growth and postnatal growth failure in very-low-birthweight infants. Acta Paediatr. 2006;95(2):236-42.

107. Zozaya C, Avila-Alvarez A, Garcia-Munoz Rodrigo F, Couce ML, Arruza L, Fernandez-Perez C, et al. The Impact of Postnatal Systemic Steroids on the Growth of Preterm Infants: A Multicenter Cohort Study. Nutrients. 2019;11(11).

108. Williams EE, Gunawardana S, Donaldson NK, Dassios T, Greenough A. Postnatal diuretics, weight gain and home oxygen requirement in extremely preterm infants. J Perinat Med. 2022;50(1):100-7.

109. Codd EF. A Relational Model of Data for Large Shared Data Banks. Communications of the Acm. 1970;13(6):377-&.

110. Wickham H. Tidy Data. Journal of Statistical Software. 2014;59(10):1-23.

111. R Core Team. R: A Language and Environment for Statistical Computing. 3.5.3 ed. Vienna, Austria: R Foundation for Statistical Computing; 2019.

112. Clevermed Ltd. BadgerNet. Edinburgh, UK: Clevermed Ltd.; 2021.

113. Battersby C, Statnikov Y, Santhakumaran S, Gray D, Modi N, Costeloe K, et al. The United Kingdom National Neonatal Research Database: A validation study. PLoS One. 2018;13(8):e0201815.

114. Office for National Statistics. Postcode to Output Area to Lower Layer Super Output Area to Middle Layer Super Output Area to Local Authority District (November 2019) Lookup in the UK. London, UK2019.

115. Cole TJ, Williams AF, Wright CM, Group RGCE. Revised birth centiles for weight, length and head circumference in the UK-WHO growth charts. Ann Hum Biol. 2011;38(1):7-11.

116. Tanner JM. Foetus into man: physical growth from conception to maturity. 2nd ed: Castlemead; 1989.

117. Cole TJ, Donaldson MD, Ben-Shlomo Y. SITAR--a useful instrument for growth curve analysis. Int J Epidemiol. 2010;39(6):1558-66.

118. Beath KJ. Infant growth modelling using a shape invariant model with random effects. Stat Med. 2007;26(12):2547-64.

119. de Onis M, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. Bull World Health Organ. 2007;85(9):660-7.

120. Box GEP, Cox DR. An Analysis of Transformations Revisited, Rebutted. Journal of the American Statistical Association. 1982;77(377):209-10.

121. Group WHOMGRS. WHO Child Growth Standards based on length/height, weight and age. Acta Paediatr Suppl. 2006;450:76-85.

122. Grajeda LM, Ivanescu A, Saito M, Crainiceanu C, Jaganath D, Gilman RH, et al. Modelling subject-specific childhood growth using linear mixed-effect models with cubic regression splines. Emerg Themes Epidemiol. 2016;13:1.

123. Tilling K, Macdonald-Wallis C, Lawlor DA, Hughes RA, Howe LD. Modelling childhood growth using fractional polynomials and linear splines. Ann Nutr Metab. 2014;65(2-3):129-38.

124. Das R, Banerjee M, Nan B, Zheng H. Fast estimation of regression parameters in a broken-stick model for longitudinal data. J Am Stat Assoc. 2016;111(515):1132-43.

125. Ooka T, Johno H, Nakamoto K, Yoda Y, Yokomichi H, Yamagata Z. Random forest approach for determining risk prediction and predictive factors of type 2 diabetes: large-scale health check-up data in Japan. BMJ Nutr Prev Health. 2021;4(1):140-8.

126. Lugo-Martinez J, Xu S, Levesque J, Gallagher D, Parker LA, Neu J, et al. Integrating longitudinal clinical and microbiome data to predict growth faltering in preterm infants. J Biomed Inform. 2022:104031.

127. Breiman L, Cutler A, Liaw A, Wiener M. Breiman and Cutler's Random Forests for Classification and Regression. 4.7-1.1 ed2022.

128. Kamath U, Liu J. Explainable artificial intelligence : an introduction to interpretable machine learning. Cham: Springer; 2021.

129. Komisarczyk K, Kozminski P, Maksymiuk S, Biecek P. Fast SHAP values computation for tree ensemble models. 0.1.1 ed2022.

130. Ringnér M. What is principal component analysis? Nature Biotechnology. 2008;26(3):303-4.

131. Rochow N, Raja P, Liu K, Fenton T, Landau-Crangle E, Gottler S, et al. Physiological adjustment to postnatal growth trajectories in healthy preterm infants. Pediatr Res. 2016;79(6):870-9.

132. Greenbury SF, Angelini ED, Ougham K, Battersby C, Gale C, Uthaya S, et al. Birthweight and patterns of postnatal weight gain in very and extremely preterm babies in England and Wales, 2008-19: a cohort study. Lancet Child Adolesc Health. 2021;5(10):719-28.

133. Williams EE, Dassios T, Mann M, Greenough A. The effect of postnatal corticosteroids on growth parameters in infants with bronchopulmonary dysplasia. J Perinat Med. 2021;49(9):1141-4.

134. Tijsseling D, Ter Wolbeek M, Derks JB, de Vries WB, Heijnen CJ, van Bel F, et al. Neonatal corticosteroid therapy affects growth patterns in early infancy. PLoS One. 2018;13(2):e0192162.

135. Buchiboyina AK, Yip CSA, Kohan R, Nathan EA, Shrestha D, Davis J, et al. Effect of cumulative dexamethasone dose in preterm infants on neurodevelopmental and growth outcomes: a Western Australia experience. Arch Dis Child Fetal Neonatal Ed. 2021;106(1):69-75.

136. Howe LD, Matijasevich A, Tilling K, Brion MJ, Leary SD, Smith GD, et al. Maternal smoking during pregnancy and offspring trajectories of height and adiposity: comparing maternal and paternal associations. Int J Epidemiol. 2012;41(3):722-32.

137. Fenton TR, Griffin IJ, Hoyos A, Groh-Wargo S, Anderson D, Ehrenkranz RA, et al. Accuracy of preterm infant weight gain velocity calculations vary depending on method used and infant age at time of measurement. Pediatric Research. 2019;85(5):650-4.

138. NHS Digital. Statistics on Women's Smoking Status at Time of Delivery: England. 2022.

139. Kalien K. The impact of maternal smoking during pregnancy on delivery outcome. European Journal of Public Health. 2001;11(3):329-33.

140. Nabet C, Ancel P-Y, Burguet A, Kaminski M. Smoking during pregnancy and preterm birth according to obstetric history: French national perinatal surveys. Paediatric and Perinatal Epidemiology. 2005;19(2):88-96.

141. Jardine J, Walker K, Gurol-Urganci I, Webster K, Muller P, Hawdon J, et al. Adverse pregnancy outcomes attributable to socioeconomic and ethnic inequalities in England: a national cohort study. Lancet. 2021;398(10314):1905-12.

142. Ministry of Housing Communities & Local Government. English indices of deprivation 2019. 2019.

143. Embleton ND. Optimal protein and energy intakes in preterm infants. Early Hum Dev. 2007;83(12):831-7.

144. Embleton ND, van den Akker CHP. Protein intakes to optimize outcomes for preterm infants. Semin Perinatol. 2019;43(7):151154.

145. Strobel NA, Adams C, McAullay DR, Edmond KM. Mother's Own Milk Compared With Formula Milk for Feeding Preterm or Low Birth Weight Infants: Systematic Review and Meta-analysis. Pediatrics. 2022;150(Supplement 1):e2022057092D.

146. National Neonatal Research Database [Internet]. Imperial College London. 2021. Available from: https://web.www.healthdatagateway.org/dataset/67020745-9def-4c6e-b5ac-bb273bd0a20e.

147. Gale C, Morris I, Neonatal Data Analysis Unit Steering B. The UK National Neonatal Research Database: using neonatal data for research, quality improvement and more. Arch Dis Child Educ Pract Ed. 2016;101(4):216-8.

148. Young A, Andrews ET, Ashton JJ, Pearson F, Beattie RM, Johnson MJ. Generating longitudinal growth charts from preterm infants fed to current recommendations. Arch Dis Child Fetal Neonatal Ed. 2020.

149. . !!! INVALID CITATION !!! 114.

150. Jiang J. Applied medical statistics. Hoboken, NJ: John Wiley & Sons, Inc.,; 2022.

151. British Association of Perinatal Medicine. Service Standards for Hospitals Providing Neonatal Care (3rd Ed.). BAPM; 2010.

152. East of England ODN. Enteral Feeding of Preterm Infants on the Neonatal Unit. 2020.

153. Grainger C, NHS Greater Galsgow and Clyde. Enteral feeding of preterm infants. 2020.

154. . !!! INVALID CITATION !!! 112.

155. Marlow N, Bennett C, Draper ES, Hennessy EM, Morgan AS, Costeloe KL. Perinatal outcomes for extremely preterm babies in relation to place of birth in England: the EPICure 2 study. Arch Dis Child Fetal Neonatal Ed. 2014;99(3):F181-8.

156. Andrews ET, Beattie RM, Johnson MJ. Measuring body composition in the preterm infant: Evidence base and practicalities. Clin Nutr. 2019.

157. Young A, Andrews ET, Ashton JJ, Pearson CF, Beattie RM, Johnson MJ, editors. Using Early Postnatal Growth in a Well-Grown Preterm Cohort to Generate Weight Centile Charts. Neonatal Society Summer Meeting; 2019 29th June 2019; Cambridge, UK.

158. Chang W, Cheng J, Allaire JJ, Xie Y, McPherson J. shiny: Web Application Framework for R. 2019.

159. Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape. Journal of the Royal Statistical Society: Series C (Applied Statistics). 2005;54(3):507-54.

160. Cole TJ, Green PJ. Smoothing reference centile curves: the LMS method and penalized likelihood. Stat Med. 1992;11(10):1305-19.

161. Cole TJ, Freeman JV, Preece MA. British 1990 growth reference centiles for weight, height, body mass index and head circumference fitted by maximum penalized likelihood. Stat Med. 1998;17(4):407-29.

162. Shaffer SG, Bradt SK, Meade VM, Hall RT. Extracellular fluid volume changes in very low birth weight infants during first 2 postnatal months. J Pediatr. 1987;111(1):124-8.

163. Ginnell L, Boardman JP, Reynolds RM, Fletcher-Watson S. Parent priorities for research and communication concerning childhood outcomes following preterm birth. Wellcome Open Res. 2021;6:151.

164. Molony CL, Hiscock R, Kaufman J, Keenan E, Hastie R, Brownfoot FC. Growth trajectory of preterm small-for-gestational-age neonates. The Journal of Maternal-Fetal & Neonatal Medicine. 2022;35(25):8400-6.

165. Embleton NE, Pang N, Cooke RJ. Postnatal malnutrition and growth retardation: an inevitable consequence of current recommendations in preterm infants? Pediatrics. 2001;107(2):270-3.

166. Joosten K, Embleton N, Yan W, Senterre T, nutrition EEECwgopp. ESPGHAN/ESPEN/ESPR/CSPEN guidelines on pediatric parenteral nutrition: Energy. Clin Nutr. 2018;37(6 Pt B):2309-14.

167. Kadioglu Simsek G, Alyamac Dizdar E, Arayici S, Canpolat FE, Sari FN, Uras N, et al. Comparison of the Effect of Three Different Fortification Methods on Growth of Very Low Birth Weight Infants. Breastfeed Med. 2019;14(1):63-8.

168. Brion LP, Fleischman AR, McCarton C, Schwartz GJ. A simple estimate of glomerular filtration rate in low birth weight infants during the first year of life: noninvasive assessment of body composition and growth. J Pediatr. 1986;109(4):698-707.

169. Muhari-Stark E, Burckart GJ. Glomerular Filtration Rate Estimation Formulas for Pediatric and Neonatal Use. J Pediatr Pharmacol Ther. 2018;23(6):424-31.

170. Bates D, Maechler M, Bolker B, Walker S, Christensen RHB, Singmann H, et al. Linear Mixed-Effects Models using 'Eigen' and S4. 1.1-27.1 ed2021.

171. Gallini F, Maggio L, Romagnoli C, Marrocco G, Tortorolo G. Progression of renal function in preterm neonates with gestational age < or = 32 weeks. Pediatr Nephrol. 2000;15(1-2):119-24.

172. Obi Y, Kalantar-Zadeh K, Streja E, Daugirdas JT. Prediction equation for calculating residual kidney urea clearance using urine collections for different hemodialysis treatment frequencies and interdialytic intervals. Nephrol Dial Transplant. 2018;33(3):530-9.

173. Hoffer LJ. How much protein do parenteral amino acid mixtures provide? The American Journal of Clinical Nutrition. 2011;94(6):1396-8.

174. D.R. F. Lippincott Illustrated Reviews: Biochemistry. 7th ed. Philadelphia, PA: Wolters Kluwer; 2017.

175. Weiner ID, Mitch WE, Sands JM. Urea and Ammonia Metabolism and the Control of Renal Nitrogen Excretion. Clin J Am Soc Nephrol. 2015;10(8):1444-58.

176. Van den Akker CH, Van Goudoever JB. Recent advances in our understanding of protein and amino acid metabolism in the human fetus. Curr Opin Clin Nutr Metab Care. 2010;13(1):75-80.

177. van Goudoever JB, H. V, C.H. vdA, F. dG, S.R.D. vdS. Amino Acids and Proteins. In: B. K, B. P, R. U, editors. Nutritional Care of Preterm Infants: Scientific Basis and Practical Guidelines. 110. Basel: Karger; 2014. p. 49-63.

178. Leenders E, de Waard M, van Goudoever JB. Low- versus High-Dose and Early versus Late Parenteral Amino-Acid Administration in Very-Low-Birth-Weight Infants: A Systematic Review and Meta-Analysis. Neonatology. 2018;113(3):187-205.

179. Roggero P, Gianni ML, Morlacchi L, Piemontese P, Liotto N, Taroni F, et al. Blood urea nitrogen concentrations in low-birth-weight preterm infants during parenteral and enteral nutrition. J Pediatr Gastroenterol Nutr. 2010;51(2):213-5.

180. Mathes M, Maas C, Bleeker C, Vek J, Bernhard W, Peter A, et al. Effect of increased enteral protein intake on plasma and urinary urea concentrations in preterm infants born at < 32 weeks gestation and < 1500 g birth weight enrolled in a randomized controlled trial - a secondary analysis. BMC Pediatr. 2018;18(1):154.

181. Young A, Brown LK, Ennis S, Beattie RM, Johnson MJ. Total body water in full-term and preterm newborns: systematic review and meta-analysis. Arch Dis Child Fetal Neonatal Ed. 2021.

182. Bonsante F, Iacobelli S, Chantegret C, Martin D, Gouyon JB. The effect of parenteral nitrogen and energy intake on electrolyte balance in the preterm infant. Eur J Clin Nutr. 2011;65(10):1088-93.

183. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(7):e1000097.

184. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev. 2016;5(1):210.

185. Moola S, Munn Z, Tufanaru C, Aromataris E, Sears K, Sfetcu R, et al. Chapter 7: Systematic reviews of etiology and risk. 2017. In: Joanna Briggs Institute Reviewer's Manual [Internet]. The Joanna Briggs Institute. Available from: https://reviewersmanual.joannabriggs.org/

186. Balduzzi S, Rucker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. Evid Based Ment Health. 2019;22(4):153-60.

187. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC Med Res Methodol. 2014;14:135.

188. Andersen JB. The effect of diuretics in late pregnancy on the newborn infant. Acta Paediatrica Scandinavica. 1970;59(6):659-63.

189. Cheek DB, Wishart J, MacLennan AH, Haslam R. Cell hydration in the normally grown, the premature and the low weight for gestational age infant. Early Human Development. 1984;10(1-2):75-84.

190. Cheek DB, Wishart J, MacLennan AH, Haslam R, Fitzgerald A. Hydration in the first 24 h of postnatal life in normal infants born vaginally or by ceasarean section. Early Human Development. 1982;7(4):323-30.

191. Clapp WM, Butterfield LJ, O'Brien D. Body Water Compartments in the Premature Infant with Special Reference to the Effects of the Respiratory Distress Syndrome and of Maternal Diabetes and Toxemia. Pediatrics. 1962;29(6):883-9.

192. Cochran WJ, Klish WJ, Wong WW, Klein PD. Total body electrical conductivity used to determine body composition in infants. Pediatric Research. 1986;20(6):561-4.

193. Llanos A, Segovia N, Gallegos R, Torres J, Mitnik M. DEUTERIUM-OXIDE SPACE AS AN INDEX OF TOTAL-BODY WATER IN FETAL GROWTH RETARDED NEWBORN-INFANTS. Journal of Pediatrics. 1976;89(2):323-4.

194. Offringa PJ, Boersma ER, Brunsting JR, Meeuwsen WP, Velvis H. Weight loss in full-term negroid infants: relationship to body water compartments at birth? Early Human Development. 1990;21(2):73-81.

195. Olhager E, Forsum E. Total energy expenditure, body composition and weight gain in moderately preterm and full-term infants at term postconceptional age. Acta Paediatr. 2003;92(11):1327-34.

196. Osler M. Body water of newborn infants of diabetic mothers. Acta Endocrinol (Copenh). 1960;34:261-76.

197. Singhi SC, Ganguli NK, Bhakoo ON, Dhall K, Sood V, Kaur A. Body water distribution in newborn infants appropriate for gestational age. Indian Journal of Medical Research. 1995;101:193-200.

198. Baarsma R, Hof P, Zijlstra WG, Zock JP, Okken A. Measurement of total bodywater volume with deuterated water in newborn infants. Biology of the Neonate. 1992;62(2-3):108-12.

199. Bauer K, Bovermann G, Roithmaier A, Gotz M, Proiss A, Versmold HT. Body composition, nutrition, and fluid balance during the first two weeks of life in preterm neonates weighing less than 1500 grams. Journal of Pediatrics. 1991;118(4 Pt 1):615-20.

200. Bhatia J, Rassin DK. Growth and total body water in premature infants fed "in-utero" or "ex-utero". Acta Paediatrica Scandinavica. 1988;77(3):326-31.

201. Hartnoll G, Betremieux P, Modi N. Body water content of extremely preterm infants at birth. Archives of Disease in Childhood Fetal & Neonatal Edition. 2000;83(1):F56-9.

202. Heimler R, Doumas BT, Jendrzejczak BM, Nemeth PB, Hoffman RG, Nelin LD. Relationship between nutrition, weight change, and fluid compartments in preterm infants during the first week of life. Journal of Pediatrics. 1993;122(1):110-4.

203. Raghavan CV, Super DM, Chatburn RL, Savin SM, Fanaroff AA, Kalhan SC. Estimation of total body water in very-low-birth-weight infants by using anthropometry with and without bioelectrical impedance and H2[(18)O]. American Journal of Clinical Nutrition. 1998;68(3):668-74.

204. Singhi S, Sood V, Bhakoo ON, Ganguly NK, Kaur A. Composition of postnatal weight loss & subsequent weight gain in preterm infants. Indian Journal of Medical Research. 1995;101:157-62.

205. Tang W, Ridout D, Modi N. Influence of respiratory distress syndrome on body composition after preterm birth. Arch Dis Child Fetal Neonatal Ed. 1997;77(1):F28-31.

206. Dimitriou G, Kavvadia V, Marcou M, Greenough A. Antenatal steroids and fluid balance in very low birthweight infants. Arch Dis Child Fetal Neonatal Ed. 2005;90(6):F509-13.

207. Stewart LA, Parmar MK. Meta-analysis of the literature or of individual patient data: is there a difference? Lancet. 1993;341(8842):418-22.

208. American Academy of Pediatrics. Committee on Nutrition., Barness LA. Pediatric nutrition handbook. 6th ed. Elk Grove Village, IL: American Academy of Pediatrics; 2009. xlix, 1470 p. p.

209. Lingwood BE. Bioelectrical impedance analysis for assessment of fluid status and body composition in neonates - The good, the bad and the unknown. European Journal of Clinical Nutrition. 2013;67(SUPPL. 1):S28-S33.

210. International Atomic Energy Agency. Introduction to Body Composition Assessment Using the Deuterium Dilution Technique with Analysis of Urine Samples by Isotope Ratio Mass Spectrometry. Vienna: International Atomic Energy Agency; 2010.

211. Coplen TB. Letter to the Editor: New IUPAC guidelines for the reporting of stable hydrogen, carbon, and oxygen isotope-ratio data. J Res Natl Inst Stand Technol. 1995;100(3):285.

212. Vaz-Drago R, Custodio N, Carmo-Fonseca M. Deep intronic mutations and human disease. Hum Genet. 2017;136(9):1093-111.

213. Illumina Inc. Planning your NGS budget [Website]. San Diego, CA: Illumina Inc.; 2022. Available from: https://emea.illumina.com/science/technology/next-generation-sequencing/beginners/ngs-cost.html.

214. Novogene Ltd. Next Generation Sequencing (NGS) 2022. Available from: https://en.novogene.com/resources/blog/ngs-beginners-guide/.

215. Hu T, Chitnis N, Monos D, Dinh A. Next-generation sequencing technologies: An overview. Hum Immunol. 2021;82(11):801-11.

216. Agilent Technologies Inc. SureSelect Human All Exon V6 Datasheet. 2015.

217. Pfeiffer F, Grober C, Blank M, Handler K, Beyer M, Schultze JL, et al. Systematic evaluation of error rates and causes in short samples in next-generation sequencing. Sci Rep. 2018;8(1):10950.

218. Pais LS, Snow H, Weisburd B, Zhang S, Baxter SM, DiTroia S, et al. seqr: A web-based analysis and collaboration tool for rare disease genomics. Hum Mutat. 2022;43(6):698-707.

219. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. Nucleic Acids Res. 2014;42(Database issue):D980-5.

220. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17(5):405-24.

221. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014;46(3):310-5.

222. Investigators GPP, Smedley D, Smith KR, Martin A, Thomas EA, McDonagh EM, et al. 100,000 Genomes Pilot on Rare-Disease Diagnosis in Health Care - Preliminary Report. N Engl J Med. 2021;385(20):1868-80.

223. Williamson SL, Rasanayagam CN, Glover KJ, Baptista J, Naik S, Satodia P, et al. Rapid exome sequencing: revolutionises the management of acutely unwell neonates. Eur J Pediatr. 2021;180(12):3587-91.

224. UK National Screening Committee. Criteria for appraising the viability, effectiveness and appropriateness of a screening programme. In: Care DoHaS, editor. London, UK2015.

225. Genomics England. Newborn Genomes Programme Vision Document. In: National Health Service, editor. London, UK2021.

226. Office for National Statistics. Child mortality (death cohort) tables in England and Wales. In: Office for National Statistics, editor. London, UK2022.

227. Rappoport N, Toung J, Hadley D, Wong RJ, Fujioka K, Reuter J, et al. A genome-wide association study identifies only two ancestry specific variants associated with spontaneous preterm birth. Sci Rep. 2018;8(1):226.

228. Salem H, Yatchenko Y, Anosov M, Rosenfeld T, Altarescu G, Grisaru-Granovsky S, et al. Maternal and neonatal irisin precursor gene FNDC5 polymorphism is associated with preterm birth. Gene. 2018;649:58-62.

229. Zhang G, Feenstra B, Bacelis J, Liu X, Muglia LM, Juodakis J, et al. Genetic Associations with Gestational Duration and Spontaneous Preterm Birth. N Engl J Med. 2017;377(12):1156-67.

230. Strauss JF, 3rd, Romero R, Gomez-Lopez N, Haymond-Thornburg H, Modi BP, Teves ME, et al. Spontaneous preterm birth: advances toward the discovery of genetic predisposition. Am J Obstet Gynecol. 2018;218(3):294-314 e2.

231. Zhang H, Baldwin DA, Bukowski RK, Parry S, Xu Y, Song C, et al. A genome-wide association study of early spontaneous preterm delivery. Genet Epidemiol. 2015;39(3):217-26.

232. NIH National Library of Medicine Variation Viewer v2.1.4 [Internet]. 2019. Available from: https://www.ncbi.nlm.nih.gov/variation/view/.

233. Yang X, Liu G, Yi B. Mdr3 gene mutation in preterm infants with parenteral nutrition-associated cholestasis. Mol Genet Genomic Med. 2022;10(3):e1875.

234. Ramel S, Rao R. Hyperglycemia in Extremely Preterm Infants. Neoreviews. 2020;21(2):e89-e97.

235. Akmal DM, Razek ARAA, Musa N, El-Aziz AGA. Incidence, risk factors and complications of hyperglycemia in very low birth weight infants. Egyptian Pediatric Association Gazette. 2017;65(3):72-9.

236. Mironov PI, Mingazov NN, Valiev RR, Lekmanov AU. Associations of molecular genetic predictors of type 2 diabetes mellitus with hyperglycemia in extremely low birth weight infants. Messenger of ANESTHESIOLOGY AND RESUSCITATION. 2021;18(5):62-8.

237. Gene; PPARG peroxisome proliferator activated receptor gamma [Internet]. National Center for Biotechnology Information. 2022 [cited 27th May 2022]. Available from: https://www.ncbi.nlm.nih.gov/gene/5468.

238. PanelApp; Familial diabetes (Version 1.66) [Internet]. Genomics England. 2022. Available from: https://panelapp.genomicsengland.co.uk/panels/152/.

239. University of California Santa Cruz. LiftOver. UCSC Genome Browser2022.

240. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26(6):841-2.

241. PanelApp; Inborn errors of metabolism (Version 2.257) [Internet]. Genomics England. 2022. Available from: https://panelapp.genomicsengland.co.uk/panels/467/.

242. Li H. bwa - Burrows-Wheeler Alignment Tool. 0.6 ed2013.

243. Broad Institute. Picard Tools. Cambridge, MA2022.

244. Jun G, Flickinger M, Hetrick KN, Romm JM, Doheny KF, Abecasis GR, et al. Detecting and estimating contamination of human DNA samples in sequencing and array-based genotype data. Am J Hum Genet. 2012;91(5):839-48.

245. Pedersen BS, Quinlan AR. Who's Who? Detecting and Resolving Sample Anomalies in Human DNA Sequencing Studies with Peddy. Am J Hum Genet. 2017;100(3):406-13.

246. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016;536(7616):285-91.

247. Ichtchenko K, Bittner MA, Krasnoperov V, Little AR, Chepurny O, Holz RW, et al. A Novel Ubiquitously Expressed α-Latrotoxin Receptor Is a Member of the CIRL Family of G-protein-coupled Receptors*. Journal of Biological Chemistry. 1999;274(9):5491-8.

248. Xing Y, Nakamura Y, Rainey WE. G protein-coupled receptor expression in the adult and fetal adrenal glands. Mol Cell Endocrinol. 2009;300(1-2):43-50.

249. Gibb W, Challis JR. Mechanisms of term and preterm birth. J Obstet Gynaecol Can. 2002;24(11):874-83.

250. Behrman RE, Butler AS, Institute of Medicine (U.S.). Committee on Understanding Premature Birth and Assuring Healthy Outcomes. Preterm birth : causes, consequences, and prevention. Washington, D.C.: National Academies Press; 2007. xvi, 772 p. p.

251. Chen R, Davydov EV, Sirota M, Butte AJ. Non-synonymous and synonymous coding SNPs show similar likelihood and effect size of human disease association. PLoS One. 2010;5(10):e13574.

252. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, Darbandi SF, Knowles D, Li YI, et al. Predicting Splicing from Primary Sequence with Deep Learning. Cell. 2019;176(3):535-48 e24.

253. Li P, Zhang Q, Tang H. INPP1 up-regulation by miR-27a contributes to the growth, migration and invasion of human cervical cancer. J Cell Mol Med. 2019;23(11):7709-16.

254. Michelon L, Meira-Lima I, Cordeiro Q, Miguita K, Breen G, Collier D, et al. Association study of the INPP1, 5HTT, BDNF, AP-2beta and GSK-3beta GENE variants and restrospectively scored response to lithium prophylaxis in bipolar disorder. Neurosci Lett. 2006;403(3):288-93.

255. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7(4):248-9.

256. Strauss KA, Markx S, Georgi B, Paul SM, Jinks RN, Hoshi T, et al. A population-based study of KCNH7 p.Arg394His and bipolar spectrum disorder. Hum Mol Genet. 2014;23(23):6395-406.

257. Arai S, Miyazaki T. Impaired maturation of myeloid progenitors in mice lacking novel Polycomb group protein MBT-1. EMBO J. 2005;24(10):1863-73.

258. Szenker-Ravi E, Altunoglu U, Leushacke M, Bosso-Lefevre C, Khatoo M, Thi Tran H, et al. RSPO2 inhibition of RNF43 and ZNRF3 governs limb development independently of LGR4/5/6. Nature. 2018;557(7706):564-9.

259. Yamada W, Nagao K, Horikoshi K, Fujikura A, Ikeda E, Inagaki Y, et al. Craniofacial malformation in R-spondin2 knockout mice. Biochem Biophys Res Commun. 2009;381(3):453-8.

260. Simaite D, Kofent J, Gong M, Ruschendorf F, Jia S, Arn P, et al. Recessive mutations in PCBD1 cause a new type of early-onset diabetes. Diabetes. 2014;63(10):3557-64.

261. Citron BA, Kaufman S, Milstien S, Naylor EW, Greene CL, Davis MD. Mutation in the 4a-carbinolamine dehydratase gene leads to mild hyperphenylalaninemia with defective cofactor metabolism. Am J Hum Genet. 1993;53(3):768-74.

262. Todorova B, Kubaszek A, Pihlajamaki J, Lindstrom J, Eriksson J, Valle TT, et al. The G-250A promoter polymorphism of the hepatic lipase gene predicts the conversion from impaired glucose tolerance to type 2 diabetes mellitus: the Finnish Diabetes Prevention Study. J Clin Endocrinol Metab. 2004;89(5):2019-23.

263. Ripatti P, Ramo JT, Soderlund S, Surakka I, Matikainen N, Pirinen M, et al. The Contribution of GWAS Loci in Familial Dyslipidemias. PLoS Genet. 2016;12(5):e1006078.

264. Yorifuji T, Kawakita R, Hosokawa Y, Fujimaru R, Yamaguchi E, Tamagawa N. Dominantly inherited diabetes mellitus caused by GATA6 haploinsufficiency: variable intrafamilial presentation. J Med Genet. 2012;49(10):642-3.

265. Borowiec M, Liew CW, Thompson R, Boonyasrisawat W, Hu J, Mlynarski WM, et al. Mutations at the BLK locus linked to maturity onset diabetes of the young and beta-cell dysfunction. Proc Natl Acad Sci U S A. 2009;106(34):14460-5.

266. Lindqvist A, Sharvill J, Sharvill DE, Andersson S. Loss-of-Function Mutation in Carotenoid 15,15'-Monooxygenase Identified in a Patient with Hypercarotenemia and Hypovitaminosis A. The Journal of Nutrition. 2007;137(11):2346-50.

267. National Library of Medicine. NM_004999.4(MYO6):c.2751dup (p.Gln918fs) AND Autosomal dominant nonsyndromic deafness 22. 2020.

268. Said M, Cappiello C, Devaney JM, Podini D, Beres AL, Vukmanovic S, et al. Genomics in premature infants: a non-invasive strategy to obtain high-quality DNA. Sci Rep. 2014;4:4286.

269. Lehmann AS, Haas DM, McCormick CL, Skaar TC, Renbarger JL. Collection of human genomic DNA from neonates: a comparison between umbilical cord blood and buccal swabs. American Journal of Obstetrics and Gynecology. 2011;204(4):362.e1-.e6.

270. Morken NH, Kallen K, Jacobsson B. Fetal growth and onset of delivery: a nationwide population-based study of preterm infants. Am J Obstet Gynecol. 2006;195(1):154-61.

271. The Data Protection Act, (2018).

272. Royal College of Physicians. Consent and confidentiality in genomic medicine. 2019.

273. Pierrat V, Marchand-Martin L, Marret S, Arnaud C, Benhammou V, Cambonie G, et al. Neurodevelopmental outcomes at age 5 among children born preterm: EPIPAGE-2 cohort study. BMJ. 2021;373:n741.

274. Gowda GAN, Raftery D. NMR-based metabolomics : methods and protocols. New York, NY: Humana Press; 2019.

275. Abdulkadir B, Nelson A, Skeath T, Marrs EC, Perry JD, Cummings SP, et al. Routine Use of Probiotics in Preterm Infants: Longitudinal Impact on the Microbiome and Metabolome. Neonatology. 2016;109(4):239-47.

276. Thomaidou A, Chatziioannou AC, Deda O, Benaki D, Gika H, Mikros E, et al. A pilot case-control study of urine metabolomics in preterm neonates with necrotizing enterocolitis. J Chromatogr B Analyt Technol Biomed Life Sci. 2019;1117:10-21.

277. Picaud JC, De Magistris A, Mussap M, Corbu S, Dessi A, Noto A, et al. Urine NMR Metabolomics Profile of Preterm Infants With Necrotizing Enterocolitis Over the First Two Months of Life: A Pilot Longitudinal Case-Control Study. Front Mol Biosci. 2021;8:680159.

278. Stewart CJ, Nelson A, Treumann A, Skeath T, Cummings SP, Embleton ND, et al. Metabolomic and proteomic analysis of serum from preterm infants with necrotising entercolitis and late-onset sepsis. Pediatr Res. 2016;79(3):425-31.

279. Younge NE, Newgard CB, Cotten CM, Goldberg RN, Muehlbauer MJ, Bain JR, et al. Disrupted Maturation of the Microbiota and Metabolome among Extremely Preterm Infants with Postnatal Growth Failure. Sci Rep. 2019;9(1):8167.

280. Dudzik D, Iglesias Platas I, Izquierdo Renau M, Balcells Esponera C, Del Rey Hurtado de Mendoza B, Lerin C, et al. Plasma Metabolome Alterations Associated with Extrauterine Growth Restriction. Nutrients. 2020;12(4).

281. Nilsson AK, Tebani A, Malmodin D, Pedersen A, Hellgren G, Lofqvist C, et al. Longitudinal Serum Metabolomics in Extremely Premature Infants: Relationships With Gestational Age, Nutrition, and Morbidities. Front Neurosci. 2022;16:830884.

282. Esturau-Escofet N, Rodriguez de San Miguel E, Vela-Amieva M, Garcia-Aguilera ME, Hernandez-Espino CC, Macias-Kauffer L, et al. A Longitudinal (1)H NMR-Based Metabolic Profile Analysis of Urine from Hospitalized Premature Newborns Receiving Enteral and Parenteral Nutrition. Metabolites. 2022;12(3).

283. Bjerrum JT. Metabonomics : Methods and Protocols: New York : Humana Press, [2015].

284. Madrid-Gambin F, Oller-Moreno S, Fernandez L, Bartova S, Giner MP, Joyce C, et al. AlpsNMR: an R package for signal processing of fully untargeted NMR-based metabolomics. Bioinformatics. 2020;36(9):2943-5.

285. Du P, Kibbe WA, Lin SM. Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching. Bioinformatics. 2006;22(17):2059-65.

286. Vu TN, Valkenborg D, Smets K, Verwaest KA, Dommisse R, Lemière F, et al. An integrated workflow for robust alignment and simplified quantitative analysis of NMR spectrometry data. BMC Bioinformatics. 2011;12(1):405.

287. Dieterle F, Ross A, Schlotterbeck G, Senn H. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in 1H NMR metabonomics. Anal Chem. 2006;78(13):4281-90.

288. Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems. 2001;58(2):109-30.

289. Mevik B-H, Wehrens R. The pls Package: Principal Component and Partial Least Squares Regression in R. Journal of Statistical Software. 2007;18(2):1 - 23.

290. Sloss S, Dawson JA, McGrory L, Rafferty AR, Davis PG, Owen LS. Observational study of parental opinion of deferred consent for neonatal research. Archives of Disease in Childhood - Fetal and Neonatal Edition. 2021;106(3):258.

291. Imbulana DI, Owen LS, Prentice TM, Davis PG, Manley BJ. Deferred Consent in Neonatal Clinical Research: Why, When, How? Pediatric Drugs. 2021;23(6):565-73.

292. Ernst M, Rogers S, Lausten-Thomsen U, Bjorkbom A, Laursen SS, Courraud J, et al. Gestational age-dependent development of the neonatal metabolome. Pediatr Res. 2021;89(6):1396-404.

293. Hallman M, Saugstad OD, Porreco RP, Epstein BL, Gluck L. Role of myoinositol in regulation of surfactant phospholipids in the newborn. Early Human Development. 1985;10(3):245-54.

294. Howlett A, Ohlsson A, Plakkal N. Inositol in preterm infants at risk for or having respiratory distress syndrome. Cochrane Database Syst Rev. 2019;7:CD000366.

295. Feng C, Bai K, Wang A, Ge X, Zhao Y, Zhang L, et al. Effects of dimethylglycine sodium salt supplementation on growth performance, hepatic antioxidant capacity, and mitochondria-related gene expression in weanling piglets born with low birth weight1. J Anim Sci. 2018;96(9):3791-803.

296. Rochow N, Landau-Crangle E, Raja P, Fusch G, Poeschl J, Heckmann M, et al. Postnatal weight loss of healthy preterm infants and crossing of percentiles on growth charts during postnatal adaptation. Monatsschrift fur Kinderheilkunde. 2016;164 (2 Supplement):S217-S8.

297. Webbe JWH, Duffy JMN, Afonso E, Al-Muzaffar I, Brunton G, Greenough A, et al. Core outcomes in neonatology: development of a core outcome set for neonatal research. Archives of Disease in Childhood - Fetal and Neonatal Edition. 2020;105(4):425.

Appendices

Appendix 1

Published paper: Young A, Andrews ET, Ashton JJ, Pearson F, Beattie RM, Johnson MJ. Generating longitudinal growth charts from preterm infants fed to current recommendations. Arch Dis Child Fetal Neonatal Ed. 2020;105(6):646-51

Archives of
Disease in Childhood

Generating Longitudinal Growth Charts from Preterm Infants Fed to Current Recommendations

Journal:	*Archives of Disease in Childhood*
Manuscript ID	fetalneonatal-2019-318404.R1
Article Type:	Original research
Date Submitted by the Author:	19-Feb-2020
Complete List of Authors:	Young, Aneurin; University Hospital Southampton NHS Foundation Trust, Department of Neonatal Medicine; University of Southampton and University Hospital Southampton NHS Foundation Trust, NIHR Southampton Biomedical Research Centre Andrews, Edward; University Hospital Southampton NHS Foundation Trust, Department of Neonatal Medicine Ashton, James; University Hospital Southampton NHS Foundation Trust, Department of Paediatric Gastroenterology; University of Southampton, Department of Human Genetics and Genomic Medicine Pearson, Freya; University Hospital Southampton NHS Foundation Trust, Department of Neonatal Medicine Beattie, R; University Hospital Southampton NHS Foundation Trust, Department of Paediatric Gastroenterology Johnson, Mark; University Hospital Southampton NHS Foundation Trust, Department of Neonatal Medicine; University of Southampton and University Hospital Southampton NHS Foundation Trust, NIHR Southampton Biomedical Research Centre
Keywords:	Neonatology, Nutrition

This article has been accepted for publication in Archives of Disease in Childhood - Fetal and Neonatal Edition, 2020, following peer review, and the Version of Record can be accessed online at http://doi.org/10.1136archdischild-2019-318404.

GENERATING LONGITUDINAL GROWTH CHARTS FROM PRETERM INFANTS FED TO CURRENT RECOMMENDATIONS

Aneurin Young[a,b], MBBS, Edward Thomas Andrews[a], BM, James John Ashton[c,d], BMBS, Freya Pearson[a], MBBS, R Mark Beattie[c], MBBS, Mark John Johnson[a,b], PhD

Affiliations/Addresses: [a]Department of Neonatal Medicine, Southampton Children's Hospital, University Hospital Southampton NHS Foundation Trust, UK; [b]NIHR Southampton Biomedical Research Centre, University Hospital Southampton NHS Foundation Trust and University of Southampton, UK; [c]Department of Paediatric Gastroenterology, Southampton Children's Hospital, University Hospital Southampton NHS Foundation Trust, UK; [d]Department of Human Genetics and Genomic Medicine, University of Southampton, UK

Corresponding Author: Aneurin Young, Neonatal Unit, University Hospital Southampton NHS Foundation Trust, Coxford Road, Southampton, SO16 5YA, UK [a.young@soton.ac.uk] 02380777222

Authors' Contributions

Dr Young made a substantial contribution to the design of the work and to the analysis and interpretation of data for the work. He drafted the work and created the web application.

Dr Andrews and Dr Ashton made substantial contributions to the conception of the work and to the acquisition of data for the work. They revised the work critically for important intellectual content.

Dr Pearson and Prof Beattie made a substantial contribution to the conception and design of the work. They revised the work critically for important intellectual content.

Dr Johnson made a substantial contribution to the conception and design of the work and to the interpretation of data for the work. He revised the work critically for important intellectual content.

All authors had final approval of the version to be published and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Word Count: 2454

Keywords: Prematurity. Nutrition. Growth. Anthropometry.

Abbreviations:

CGA	-	Corrected gestational age
GA	-	Gestational age
HC	-	Head circumference
IQR	-	Interquartile range
MGRS	-	Multicentre Growth Reference Study
MDT	-	Multidisciplinary team
NHS	-	National Health Service
NICM	-	Newborn Infant Close Monitoring chart
RRI	-	Recommended range of intake
SDS	-	Standard deviation score
UK	-	United Kingdom
WHO	-	World Health Organization

WHAT IS KNOWN ABOUT THIS TOPIC

- Current growth charts for the preterm infant are derived from cross-sectional birthweight data, making their applicability to growing preterm infants uncertain.
- The WHO Multicentre Growth Reference Study set a new standard for the production of growth charts using longitudinally gathered repeated measurements of children.
- Despite work in the INTERGROWTH-21st project, this approach has not been fully implemented for preterm infants.

WHAT THIS STUDY ADDS

- This study presents growth charts derived from serial measurements of weight, length and head circumference.
- These charts reflect the postnatal growth of preterm infants, drawn from a real-world population fed in line with current nutritional guidelines.
- A web application has been developed to allow practitioners to plot the growth of infants under their care on these growth charts.

ABSTRACT

Objective: To use repeated measurements of weight, length and, head circumference to generate growth centile charts reflecting real-world growth of a population of very preterm infants with a well-described nutritional intake close to current recommendations.

Design: Infants born before 30 weeks gestational age (GA) were recruited. Infants received nutrition according to an integrated care pathway, with nutrient intake recorded daily, weight recorded twice-weekly and length and head circumference weekly. The LMS method was used to construct growth centile charts between 24 and 36 weeks corrected GA for each parameter.

Setting: A single tertiary neonatal unit in England.

Patients: 212 infants (124 male) (median GA at birth: 27.3 weeks, median birthweight: 900g).

Results: Median daily energy, protein, carbohydrate, and fat intake were within 3% of published recommendations. The total number of measurements recorded was 5944 (3431 for weight, 1227 for length and 1286 for head circumference). Centile charts were formed for each parameter. Data for male and female infants demonstrated similar patterns of growth and were pooled for LMS analysis. A web application was created and published (bit.ly/sotongrowth) to allow infants to be plotted on these charts with changes in SDS of measurements reported and graphically illustrated.

Conclusions: These charts reflect growth in a real world cohort of preterm infants whose nutrient intakes are close to current recommendations. This work demonstrates the feasibility of forming growth charts from serial measurements of growing preterm infants fed according to current recommendations which will aid clinicians in setting a benchmark for achievable early growth.

INTRODUCTION

Preterm infants are at risk of poor growth and tend to leave hospital lighter and shorter than their term-born counterparts.[1] The causes of growth failure are multifactorial and include intercurrent illnesses, complications of prematurity and the sequelae of an adverse *in utero* environment. However, nutritional intake plays a central role and changes in nutritional practice influence growth.[2] [3] Therefore, monitoring of growth against appropriate standards underpins the nutritional care of the preterm infant.

Growth is currently benchmarked against centile charts formed by taking cross-sectional weight data. Examples include the Fenton growth reference[4] which is used extensively, including in North America, and the Newborn Infant Close Monitoring chart (NICM)[5] used in the UK. In both, cross sectional birthweight data, acting as a proxy for in-utero growth, is used to create the chart. Such an approach assumes that preterm infants grow normally up until the point of delivery,[6] and does not generate a true longitudinal growth chart which follows individuals over time. Several studies have shown that, if preterm infants are followed up longitudinally, as a population they exhibit growth failure when compared to such cross-sectional charts, falling down two or more marked centile lines during the first few weeks of life and never recovering. Whether this is 'normal preterm growth' is uncertain.[7] [8] The pattern of growth described in these studies has led clinicians to expect and accept growth which falls well short of that demonstrated by fetuses in utero (based on growth charts derived from birthweight data).This limits the usefulness of current growth charts as significant downward deviation from centile lines is ignored or tolerated.

Recent work by our group has demonstrated that weight loss in the first two weeks and subsequent decline in centile line from birth during the initial neonatal admission is not inevitable,[9] and can be alleviated by a comprehensive approach to nutritional care including guidelines, close monitoring of growth, and regular multidisciplinary nutritional review.[2] Furthermore, growth-restricted infants

subjected to this nutritional approach show a similar growth pattern to their counterparts whose growth is appropriate for their gestational age.[10]

In recent years, the WHO Multicentre Growth Reference Study (MGRS) has facilitated the generation of longitudinal growth charts, tracking the growth of term infants from six different countries between birth and 16 years of age.[11] Such standards do not exist for very preterm infants. Longitudinal charts for preterm infant growth were created for the INTERGROWTH 21st study.[12] However, it is notable that this component of INTERGROWTH data is derived from 201 infants with a mean gestation at birth of 35.5 weeks and only 28 infants (14%) born below 34 weeks GA. Therefore, data on more significantly preterm infants is mostly extrapolated and may not be representative of how such infants can or should grow, calling into question the applicability of INTERGROWTH in the most preterm populations. Furthermore, whilst there was some use of nutritional guidelines in the INTERGROWTH study, actual nutritional intakes were not reported and so the adequacy of the nutrition supplied to the infants cannot be assessed, may have been insufficient and was likely to be heterogenous across the cohort. It is therefore not clear if the INTERGROWTH infants were able to grow to their full potential.

We aimed to generate growth charts for weight, length and head circumference using data generated from a real-world population of growing preterm infants (born before 30 weeks completed gestation) who were fed to current recommendations and had close nutritional monitoring. We also created a web-based tool which can be used to plot the growth of a preterm infant on these charts.

METHODS

Eligibility Criteria

Infants born before 30 weeks gestational age were recruited from the neonatal unit of University Hospital Southampton between April 2014 and March 2019. There are over 5000 deliveries each year at University Hospital Southampton which is one of three tertiary neonatal units set within the Thames Valley and Wessex Operational Delivery Network and provides tertiary level neonatal, surgical and cardiac services to a population of over two million people in urban, suburban and rural settings.

Gestational age at birth was determined by ultrasonographic dating during the first trimester of pregnancy. Infants were recruited after parents gave informed consent in accordance with NHS Research Ethics Committee approval (Oxford A, ref 14/SC/1275). Infants were recruited within one week of birth. Patients who were transferred to the unit for gastro-intestinal or cardiac surgery, or were diagnosed with genetic syndromes known to impact on growth, were excluded. Follow-up ceased at either discharge from the recruiting neonatal unit, reaching 36 weeks corrected gestational age (CGA), or death.

Anthropometry

Infants underwent twice-weekly measurement of weight and weekly measurement of head circumference and length performed by clinical staff trained according to a standard operating procedure as part of routine clinical care.[13]

Nutritional Care

Infants were subject to a standardised nutritional approach. This approach has been previously described[2] but briefly, parenteral nutrition from the first day of life was accompanied by early trophic feeding. Enteral feeds with maternal breastmilk or preterm formula were increased at a standardised rate and breastmilk fortifier was introduced when enteral feeding reached 100-

130ml/kg/day. Weekly nutrition-focused MDT meetings addressed problems and set nutritional strategies.

Nutritional Intake

Daily nutritional intake data were calculated and stored for each infant prospectively based on the type and volume of enteral and parenteral feeds received each day, using a specially designed spreadsheet containing the nutrient content of all feeds for 33 selected nutrients. Nutrient intake is described per kilogram of body weight, based on the infant's most recent weight on that day.

Statistical Analysis

Centile charts describing the growth of infants were constructed using the LMS method.[14] This method estimates three age-specific cubic spline curves describing a Box-Cox power to remove skewness (the L curve), the median (M curve), and the coefficient of variation (S curve). Centile lines are then formed from these cubic spline curves. LMS functions of the GAMLSS package[15] were used in R[16] within the RStudio environment.[17] Data were assessed for normality using the Shapiro-Wilk test, with normal data summarised using mean and standard deviation, and non-normal data summarised using median and range. Descriptive data and nutritional intake charts were prepared using Stata 15.[18] Birthweight SDS, birth head circumference SDS and definitions of SGA and LGA were derived from the UK Neonatal and Infant Close Monitoring Growth Chart[19] constructed from reanalysed UK 1990 data.[14]

Online Package

The Shiny package[20] within RStudio was used to form a web application, published on the shinyapps.io platform (RStudio, 2017).

RESULTS

212 infants were recruited from 1st October 2014 to 31st March 2019 as part of the Growth Assessment of Preterm Infants (GAP) study (Figure 1). Table 1 shows the characteristics of the population studied. Birth characteristics were not normally distributed and are therefore reported as a median and range.

Gestational age at birth	27·29 (23 to 9·86)
Gestational age at discharge or death	36·07 (26 to 51·43)
Length of follow-up (days)	59·5 (7 to 189)
Sex (male), n (%)	124 (58·5)
Birthweight (kg)	0·90 (0·45 to 1·61)
Birthweight SDS	-0·31 (-3·61 to 2·18)
Birth HC (cm)	24·0 (19·9 to 29·0)
Birth HC SDS	-1·02 (-4·20 to 1·63)
Birth length (cm)	34·0 (26·5 to 40·2)
Small for gestational age (weight <10th centile), n (%)	48 (22·6)
Large for gestational age (weight >90th centile), n (%)	6 (2·8)

Table 1. *Demographic characteristics of preterm infants studied. HC = Head circumference; SDS = Standard deviation score. Data are median (range) unless otherwise stated.*

There were a total of 3431 weight measurements (median 18 per infant), 1227 length measurements (median 6 per infant) and 1286 head circumference measurements (median 7 per infant).

Growth Chart Creation

Centile charts for weight, length and head circumference were generated using the LMS method (Figure 2). Male and female infants were pooled for LMS analysis and growth chart creation. Visual comparison of the LMS 50th centile lines for males and females separately indicated that, although male infants were generally larger than females, patterns of growth did not substantially differ for any growth parameter (Supplementary Figure 1), meaning that changes in SDS (equivalent to downwards or upwards 'crossing' of centile lines) can be used to track the progress of growth regardless of sex. Further support for this approach can be found in the striking similarity of the

curves at all centiles and the routine pooling of infants of different sexes in subgroup analysis of a very large cohort of preterm infants by Cole at al.[8]

Nutritional Intake

Nutritional intake data was analysed from birth to 36 weeks CGA or discharge. The median daily intake of selected key macronutrients across stay for all studied infants is set out in Table 2 and the first month of life is plotted in Figure 3.

Also shown is the current recommended range of intake (RRI) for fully enterally fed preterm infants[21] and the median daily intake as a percentage of the minimum RRI. Median daily protein intake (and protein-energy ratio) was just below the RRI. Median daily total energy and fat intakes were within the RRI and median daily carbohydrate intake was above the RRI.

Nutrient	Median Daily Intake (IQR)	Recommended Intake (RI)*	Percentage of minimum RI achieved (IQR)
Energy (kcal/kg)	120 (98-136)	110-130	109 (89-124)
Carbohydrate (g/kg)	14·5 (12·4-16·5)	11·6-13·2	125 (106-142)
Protein (g/kg)	3·40 (2·78-3·92)	3·5-4·5	97 (80-112)
Fat (g/kg)	5·45 (3·32-6·56)	4·8-6·6	114 (69-137)
Protein:Energy Ratio (g/100 kcal)	2·94 (2·65-3·10)	3·2-4·1	92 (83-97)

Table 2. *Median daily intake values for selected macronutrients from birth to 36 weeks CGA, recommended intake (RI),[21] mean daily intake as a proportion of minimum RI as % (IQR). IQR = interquartile range.*

Energy intake increased rapidly during the first week of life and then exhibited a slower rise during the remainder of the first month (Figure 3.A). Carbohydrate intake rose rapidly at first and then stabilized at a level above the recommended intake, whilst protein intake rose slowly after the first week (ultimately achieving the recommended intake range) and fat intake rose steadily during the first three weeks of life (Figure 3.B-D). This pattern led to an initial fall in protein-energy ratio followed by a gradual increase (Figure 3.E).

Online Package

An online tool illustrating the new growth charts was created and published at bit.ly/sotongrowth as the Southampton Preterm Growth Package. This tool allows users to enter weight, length and head circumference growth data for a patient and plots that data on centile charts. It also reports and graphically illustrates changes in SDS and allows the changes in weight, length and head circumference SDS to be easily compared. It also provides a basic simulation of a growing preterm infant to illustrate the application's features.

DISCUSSION

We have shown that longitudinal anthropometric data from preterm infants can be used to generate novel growth charts for weight, length and head circumference. These data are derived from a population of preterm infants for whom detailed nutritional information is available and who generally received nutritional intakes close to current recommendations. They represent a 'real-world' mixed population of preterm infants, including those with intrauterine growth restriction and those who suffered complications of prematurity, and are likely to reflect the general population of preterm infants admitted to neonatal intensive care units. They received a carefully applied protocol of nutritional care using products and strategies which are widely available in developed countries.[2]

Charts defining optimal growth would ideally be derived from a very large number of infants who were representative of the whole preterm population, received standardised care which met nutritional requirements and were free of comorbidities. Forming such a cohort is challenging. There remains substantial controversy surrounding nutritional standards. This study used the standards set out by Koletzko[21]. These standards are substantially similar to other international guidelines, including those of the European Society for Paediatric Gastroenterology, Hepatology and Nutrition (ESPGHAN)[22], but some differences remain. Uauy and Koletzko define nutritional needs as "the amount and chemical form of a nutrient needed to support normal health, growth and development without disturbing the metabolism of other nutrients".[23] There is insufficient evidence available to make recommendations which fully comply with that definition. In practice, nutritional guidelines are derived from clinical trials of dietary interventions and from observed accretion rates and physiological responses to changes in diet. In the absence of certainty regarding optimal nutritional intake, there is substantial heterogeneity in the approach to nutrition between neonatal units and therefore a lack of a large cohort of infants fed similarly and to an agreed standard.

Despite the use of a nutritional care guideline, there were some deviations from recommended intakes (Table 2 and Figure 2) especially during the first days of life. Specifically, carbohydrate

provision was often above the recommended range and the protein provision slightly below it, resulting in a depressed protein-energy ratio. Excess carbohydrate intake has previously been shown to be associated with increased fat storage.[24] We have previously shown that infants managed in this way mimic in utero weight gain, but increases in head circumference and length did not keep pace.[9] This has led to questioning of the quality of their growth, potential for excess adiposity and the impact of our nutritional approach on body composition.[25] Despite a standardised nutritional approach and careful nutritional care, these charts are derived from infants who may not have received optimal nutrition in every respect. Furthermore, this study is limited by the relatively small number of infants included and by its restriction to a single centre, however this did allow standardized capture of nutritional intake. Calculations of nutrient intake were reliable when parenteral nutrition or commercially available formulas are used, but values for maternal and donor breastmilk are derived from historical reference data and so may not reflect the true nutritional composition of the milk received by each individual infant.

Longitudinal charts in older children have typically been created from cohorts of individuals free from comorbidities. However, significantly preterm infants can be expected to experience some complications of their prematurity. It is not obvious how such a cohort of 'well' preterm infants could be differentiated from a 'comorbid' group for the purpose of setting growth standards. We took the approach of including all infants who had not been specifically referred to our centre for surgical care. The resultant charts represent a mixed cohort of infants fed as their clinical condition and comorbidities allowed, as opposed to optimally fed patients without comorbidities. In common with the INTERGROWTH-21st approach, infants were born and entered the study at different gestational ages, meaning that the resultant graphs form a hybrid of cross-sectional birth data and longitudinal growth data.

Tracking individual growth on a chart constructed to include a range of gestations is further complicated by expected weight loss during the first two weeks. The acceptable magnitude of this

weight loss is not clear, is likely to less pronounced at earlier gestations[8] and may be preventable.[9] Innovative personalised growth trajectories may be developed to address this issue.[26]

Web-based applications are becoming increasingly popular tools for the analysis of healthcare information.[27] The package created for this project demonstrates the potential for simple web-based tools to be used to assess the growth of preterm infants, when combined with growth standards. With the increasing ubiquity of comprehensive clinical information systems in neonatal units, there is scope for such applications to be integrated into those systems and into the routine monitoring and clinical care of the growing infant.

CONCLUSIONS

We have demonstrated the feasibility of creating growth charts using longitudinal measurements of preterm infants with known nutritional intake. Current charts are based on cross-sectional data and clinicians tolerate downward deviation through the centiles. In contrast, charts created using the approach used in the present study with larger cohorts (and extending into later growth) may allow clinicians to feel a greater confidence that growth along a centile line represents an achievable target, improving the clinical usefulness of growth charts in the preterm population. Web apps can be created from such charts and may provide a tool to promote their implementation in clinical care.

ACKNOWLEDGEMENTS
With thanks to the staff of University Hospital Southampton Neonatal Unit and especially to the research nursing team. Thanks also to the families who participated in the study.

COMPETING INTERESTS: All authors have no conflicts of interest to disclose.

FUNDING SOURCE: The study was supported by the National Institute for Health Research Biomedical Research Centre Southampton, UK. JJA is funded by an Action Medical Research training fellowship and by an ESPEN personal fellowship.

DATA SHARING STATEMENT: We will share individual participant data that underlie the results reported in this article, after de-identification (text, tables, figures, and appendices). along with the study protocol and analytic code to researchers who provide a methodologically sound proposal beginning three months and ending five years following article publication. Proposals should be directed to a.young@soton.ac.uk.

REFERENCES

1. Johnson MJ, Wootton SA, Leaf AA, et al. Preterm birth and body composition at term equivalent age: a systematic review and meta-analysis. *Pediatrics* 2012;130(3):e640-9. doi: 10.1542/peds.2011-3379 [published Online First: 2012/08/15]
2. Johnson MJ, Leaf AA, Pearson F, et al. Successfully implementing and embedding guidelines to improve the nutrition and growth of preterm infants in neonatal intensive care: a prospective interventional study. *BMJ Open* 2017;7(12):e017727. doi: 10.1136/bmjopen-2017-017727 [published Online First: 2017/12/09]
3. McKenzie BL, Edmonds L, Thomson R, et al. Nutrition Practices and Predictors of Postnatal Growth in Preterm Infants During Hospitalization: A Longitudinal Study. *J Pediatr Gastroenterol Nutr* 2018;66(2):312-17. doi: 10.1097/MPG.0000000000001747 [published Online First: 2017/09/28]
4. Fenton TR, Kim JH. A systematic review and meta-analysis to revise the Fenton growth chart for preterm infants. *BMC Pediatr* 2013;13:59. doi: 10.1186/1471-2431-13-59 [published Online First: 2013/04/23]
5. RCPCH. UK-WHO growth charts - neonatal and infant close monitoring (NICM) London, UK: Royal College of Paediatrics and Child Health; 2019 [Available from: https://www.rcpch.ac.uk/resources/uk-who-growth-charts-neonatal-infant-close-monitoring-nicm.
6. Clark RH, Olsen IE. Do We Need Another Set of Growth Charts for Premature Infants? *Pediatrics* 2016;138(6) doi: 10.1542/peds.2016-3128 [published Online First: 2016/12/13]
7. Ehrenkranz RA. Extrauterine growth restriction: is it preventable? *J Pediatr (Rio J)* 2014;90(1):1-3. doi: 10.1016/j.jped.2013.10.003 [published Online First: 2013/10/26]
8. Cole TJ, Statnikov Y, Santhakumaran S, et al. Birth weight and longitudinal growth in infants born below 32 weeks' gestation: a UK population study. *Arch Dis Child Fetal Neonatal Ed* 2014;99(1):F34-40. doi: 10.1136/archdischild-2012-303536 [published Online First: 2013/08/13]
9. Andrews ET, Ashton JJ, Pearson F, et al. Early postnatal growth failure in preterm infants is not inevitable. *Arch Dis Child Fetal Neonatal Ed* 2019;104(3):F235-F41. doi: 10.1136/archdischild-2018-315082 [published Online First: 2018/08/24]
10. Young A, Andrews ET, Ashton JJ, et al. 'Catch-up' growth of infants with IUGR does not significantly contribute to the whole-cohort weight gain pattern. *Arch Dis Child Fetal Neonatal Ed* 2019 doi: 10.1136/archdischild-2019-317566 [published Online First: 2019/08/01]
11. de Onis M, Garza C, Victora CG. The WHO Multicentre Growth Reference Study: strategy for developing a new international growth reference. *Forum Nutr* 2003;56:238-40. [published Online First: 2005/04/06]
12. Villar J, Giuliani F, Bhutta ZA, et al. Postnatal growth standards for preterm infants: the Preterm Postnatal Follow-up Study of the INTERGROWTH-21(st) Project. *Lancet Glob Health* 2015;3(11):e681-91. doi: 10.1016/S2214-109X(15)00163-1 [published Online First: 2015/10/18]
13. Ashton JJ, Johnson MJ, Pond J, et al. Assessing the growth of preterm infants using detailed anthropometry. *Acta Paediatr* 2017;106(6):889-96. doi: 10.1111/apa.13804 [published Online First: 2017/02/28]
14. Cole TJ, Freeman JV, Preece MA. British 1990 growth reference centiles for weight, height, body mass index and head circumference fitted by maximum penalized likelihood. *Stat Med* 1998;17(4):407-29. [published Online First: 1998/03/13]
15. Rigby RA, Stasinopoulos DM. Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 2005;54(3):507-54. doi: 10.1111/j.1467-9876.2005.00510.x
16. R: A Language and Environment for Statistical Computing [program]. 3.5.3 version. Vienna, Austria: R Foundation for Statistical Computing, 2019.

17. RStudio: Integrated Development Environment for R [program]. Boston, MA: RStudio, Inc., 2018.
18. Stata Statistical Software: Release 15 [program]. College Station, TX: StataCorp LLC., 2017.
19. Department of Health. Neonatal and Infant Close Monitoring Growth Charts. London, UK: Department of Health, 2009.
20. shiny: Web Application Framework for R [program], 2019.
21. Koletzko B, Poindexter B, Uauy R. Recommended nutrient intake levels for stable, fully enterally fed very low birth weight infants. *World Rev Nutr Diet* 2014;110:297-9. doi: 10.1159/000360195 [published Online First: 2014/04/23]
22. Agostoni C, Buonocore G, Carnielli VP, et al. Enteral nutrient supply for preterm infants: commentary from the European Society of Paediatric Gastroenterology, Hepatology and Nutrition Committee on Nutrition. *J Pediatr Gastroenterol Nutr* 2010;50(1):85-91. doi: 10.1097/MPG.0b013e3181adaee0 [published Online First: 2009/11/03]
23. Uauy R, Koletzko B. Defining the Nutritional Needs of Preterm Infants. In: Poindexter B, Uauy R, eds. Nutritional Care of Preterm Infants : Scientific Basis and Practical Guidelines. Basel: Karger 2014:9-15.
24. Putet G, Senterre J, Rigo J, et al. Energy balance and composition of body weight. *Biol Neonate* 1987;52 Suppl 1:17-24. doi: 10.1159/000242736 [published Online First: 1987/01/01]
25. Menon G, Davidson AL, Drake AJ, et al. Is preterm nutrition a trade-off between head and heart? *Arch Dis Child Fetal Neonatal Ed* 2019;104(3):F232-F34. doi: 10.1136/archdischild-2018-315672 [published Online First: 2018/10/14]
26. Rochow N, Landau-Crangle E, So HY, et al. Z-score differences based on cross-sectional growth charts do not reflect the growth rate of very low birth weight infants. *PLoS One* 2019;14(5):e0216048. doi: 10.1371/journal.pone.0216048 [published Online First: 2019/05/08]
27. Lazakidou AA. Web-based applications in healthcare and biomedicine. New York: Springer 2010.

LEGENDS

Figure 1. Recruitment flow chart.

Figure 2. Centile charts for: A) weight, B) length and C) head circumference from 24 to 36 weeks corrected gestational age. Centile lines are marked on the right. Points are individual measurements.

Figure 3. Median (±IQR) daily intake for all infants from first full day of life to 28th day of life of: A) energy, B) carbohydrate, C) protein, D) fat, and E) protein-energy ratio, with recommended intake range indicated by red box.

Supplementary Figure 1. 50th centile lines for male (dashed line) and female (solid line) infants generated using LMS for A) weight, B) length and C) head circumference.

Figure 1

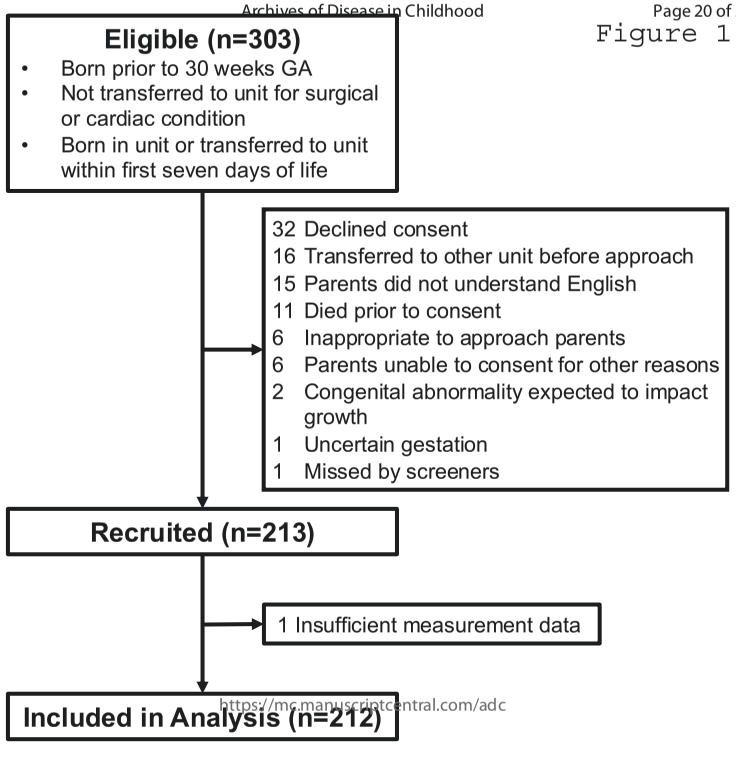


Figure 2

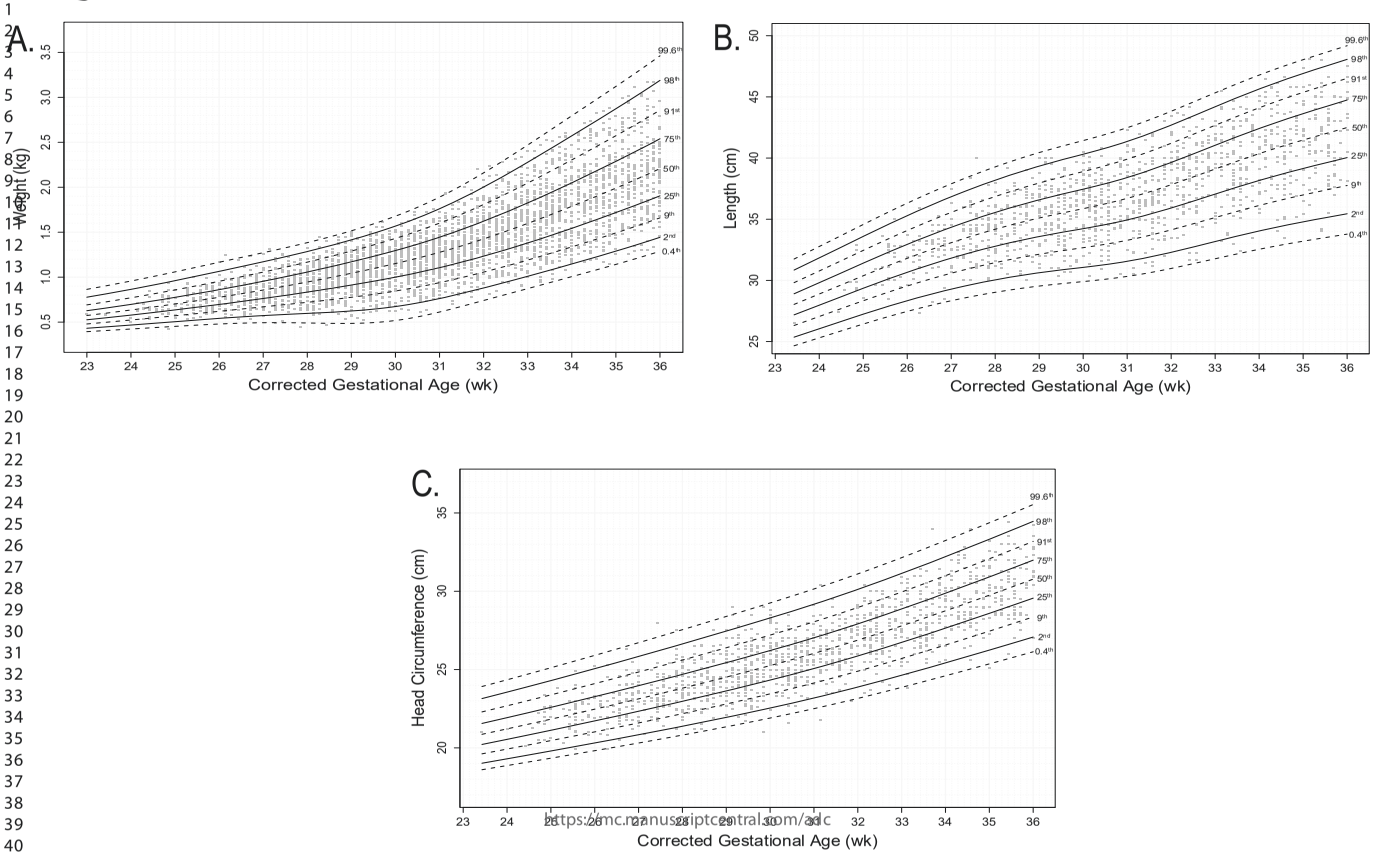
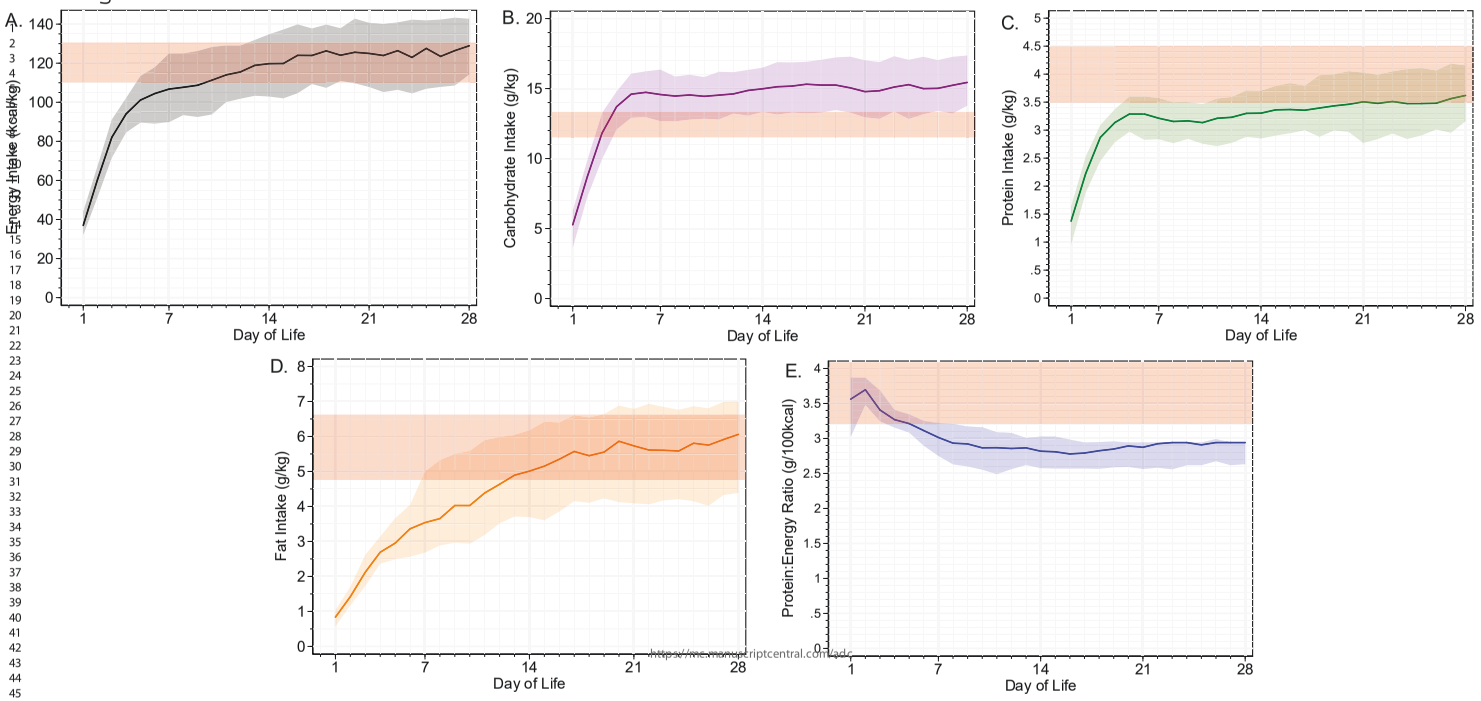
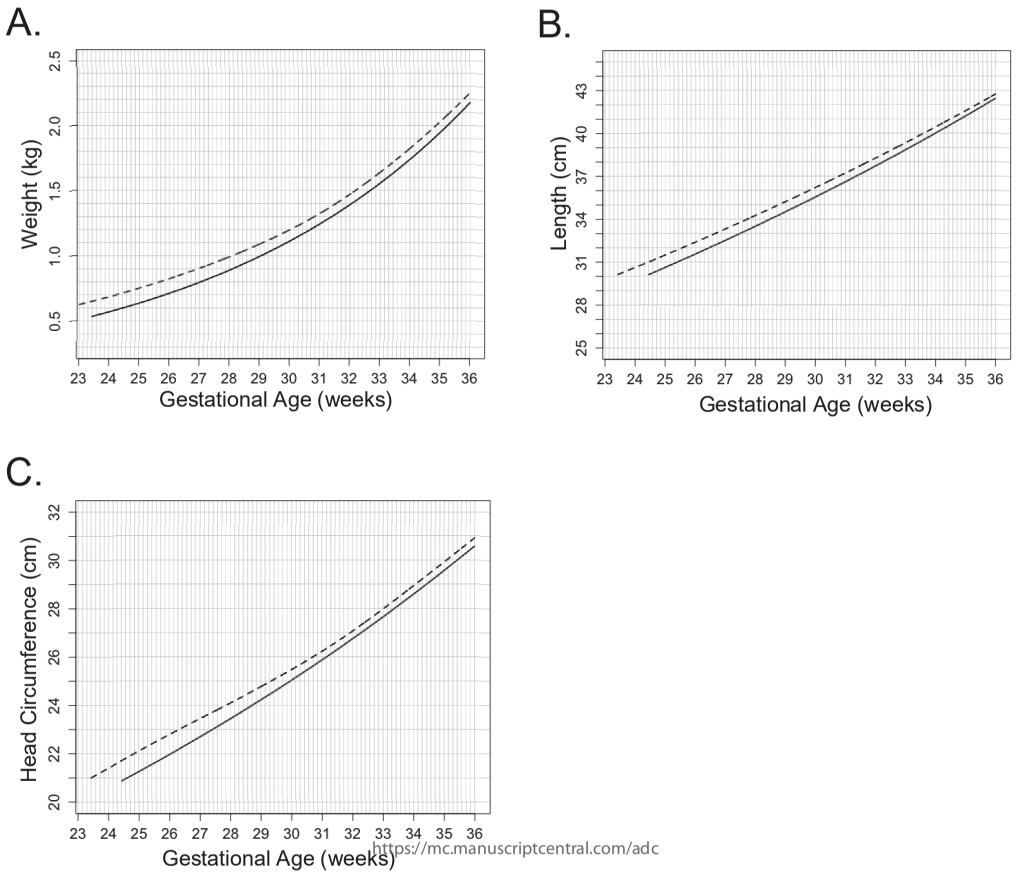


Figure 3



Supplementary Figure 1



Appendix 2

Published paper: Young A, Brown LK, Ennis S, Beattie RM, Johnson MJ. Total body water in full-term and preterm newborns: systematic review and meta-analysis. Arch Dis Child Fetal Neonatal Ed. 2021;106(5):542-8

TOTAL BODY WATER IN FULL TERM AND PRETERM NEWBORNS: SYSTEMATIC REVIEW AND META-ANALYSIS

Aneurin Young[a,b] (ORCID: 0000-0003-3549-3813, Twitter: @AneurinYoung), Lisa K Brown[a] (0000-0002-5982-121), Sarah Ennis[c] (0000-0003-2648-0869), R Mark Beattie[d] (0000-0003-4721-0577, @RMBeattie50), Mark J Johnson[a,b] (0000-0003-1829-9912, @MarkJohnson800).

[a]Department of Neonatal Medicine, Southampton Children's Hospital, University Hospital Southampton NHS Foundation Trust, UK; [b]NIHR Southampton Biomedical Research Centre, University Hospital Southampton NHS Foundation Trust and University of Southampton, UK; [c]Department of Human Genetics and Genomic Medicine, University of Southampton, UK, [d]Department of Paediatric Gastroenterology, Southampton Children's Hospital, University Hospital Southampton NHS Foundation Trust, UK;

Corresponding Author: Aneurin Young, Department of Neonatal Medicine, University Hospital Southampton NHS Foundation Trust, Coxford Road, Southampton, SO16 5YA, UK [a.young@soton.ac.uk] 02380777222

Word Count: 2498

This article has been accepted for publication in Archives of Disease in Childhood - Fetal and Neonatal Edition, 2021, following peer review, and the Version of Record can be accessed online at doi.org/10.1136/archdischild-2020-321112

WHAT IS ALREADY KNOWN ON THIS TOPIC?

- International guidance recommends that the body composition of growing preterm infants should mimic that of the fetus *in utero*.
- Total body water (TBW) is influenced by fluid status, electrolyte balance and body composition.
- Changes in TBW over time in relation to prematurity are not well understood.

WHAT THIS STUDY ADDS

- The typical proportion of total body water in a newborn term infant is around 74%.
- Preterm infants have a greater proportion of total body water at birth, up to 90% in the most premature neonates at 26 weeks, and this decreases with gestational age to 75% at 36 weeks.
- These findings may be of use to the clinician attempting to track changes in fluid status, electrolyte balance and body composition after birth in preterm infants.

ABSTRACT

Background: Total body water (TBW) is one component of fat free mass and changes in TBW are influenced by fluid shifts (especially during transition to postnatal life), electrolyte balance and nutritional status. Normal values for term-born neonates and preterm infants at birth have not been defined in large cohorts, limiting investigation into its monitoring and use in clinical practice.

Objective: To systematically review the evidence base for percentage of TBW in term-born infants, quantify the effect of prematurity on TBW at birth, and describe normal progression of TBW over time in preterm infants.

Methods: Systematic review of Medline, Web of Science Core Collection and EBSCO-CINAHL (January 1946 to January 2020). Included articles used dilutional methods to assess TBW.

Results: Searches identified 2349 articles of which 22 included data suitable for analysis. Mean TBW in term-born newborns was 73.76% (95% CI 72.47% to 75.06%, 15 studies, 433 infants). Meta-regression showed that TBW was higher in preterm infants (up to 90% at 26 weeks gestation, dropping to 74% at 36 weeks corrected gestation) and was negatively correlated with gestation at birth, falling 1.44% per week (95% CI 0.63% to 2.24%, 9 studies, 189 infants). Analysis of TBW over time during the *ex-utero* growth of preterm infants was not possible due to a paucity of data.

Conclusion: This review defines the normal total body water percentage in term born infants and confirms and quantifies previous findings that preterm infants have a higher TBW percentage.

PROSPERO Registration: CRD42019111436

INTRODUCTION

Current European[1] and North American[2] guidance for the nutritional care of preterm infants recommends that *ex utero* growth should match that of the fetus in utero, both in terms of weight gain and body composition. Mimicking *in utero* body composition has proven difficult to achieve in practice, with preterm infants leaving hospital lighter and shorter than their term-born counterparts.[3] Importantly, preterm infants also demonstrate abnormalities of their body composition at term equivalent age (TEA).[3] Specifically, they have a higher percentage of total body fat, mostly due to a relative failure to accrete lean mass.

Postnatal changes in total body water are influenced by a complex and interacting set of factors. During early adaptation there is a weight loss which is mainly mediated by loss of body water (the "preterm contraction of extracellular spaces" or PreCES).[4,5] However, this is also frequently a period of cumulative nutritional deficit. Therefore, TBW changes reflect a composite of early loss of body water, subsequent abnormalities of fluid and electrolyte balance and alterations in body composition. A fuller understanding of normal values and patterns of TBW would inform investigation and clinical management of fluid status, electrolyte management and nutritional approaches.

TBW contributes to the fat free mass (FFM) component of body composition, and measurement of TBW can be used to help derive estimates of FFM (especially after early fluid loss and fluid balance disturbances), using assumptions relating to the water content of lean and fat-containing tissues.[6] Therefore, a greater understanding of TBW patterns may be used as one avenue to explore and monitor elements of the derangements of body composition seen in preterm infants.

It is increasingly recognised that changes in weight are inadequate to understand these abnormalities of body composition, even when taken together with length. Observational data have demonstrated an association between greater early gains in fat-free mass (FFM) (but not fat mass) with improved neurodevelopmental markers at twelve months corrected age. These findings were robust to adjustment for known clinical confounding factors.[7] Preterm infants are also prone to metabolic syndrome in later life.[8,9] The mechanisms underlying this remain uncertain,[10] although differences in adiposity,

particularly the high fat and low lean mass pattern seen in preterm infants, have been associated with cardiometabolic risk.[11]

Our group has previously reviewed and summarised available methods for assessing body composition in infants.[6] In short, several methods which are useful in older children and adults are impractical in the preterm infant receiving intensive care. These include MRI, air displacement plethysmography and DXA scanning. However, assessment of TBW using dilution methods is possible in this group, and involves the administration of an exogenous substance which equilibrates with the TBW pool. Levels of this substance can then be measured, and the total body water volume calculated. Early dilution studies used antipyrine, but this was later superseded due to its slight protein binding, its pharmacological effects and its metabolism.[12,13] Stable isotopes of water (deuterium oxide, 18-oxygen water and tritiated water) produced values in good agreement with antipyrine experiments[12] and lack the problems associated with antipyrine, although analysis is considerably more complex and expensive.

This systematic review aims to assess:

1. The normal percentage TBW of the term-born infant at birth
2. The percentage TBW of preterm infants at birth, and
3. The change in TBW as preterm infants grow and mature.

METHODS

This systematic review was prospectively registered with PROSPERO (CRD42019111436, https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=111436). It is reported in line with the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) guidelines.[14]

Eligibility Criteria

The initial search strategy identified studies using a range of methods to assess total body water percentage. Results arising from different methods could not be combined in meta-analysis and so retained studies were limited to those using dilution techniques (deuterium oxide (^2H_2O), doubly labelled water ($^2H_2{}^{18}O$) or antipyrine dilution).

Studies were included if they used dilutional methods to measure total body water in term-born or preterm infants within two weeks of birth. In addition, selected studies were required to report TBW percentage and sufficient information to define gestational age and corrected gestational age of infants at the time of TBW assessment. Studies were excluded if they concerned infants with congenital abnormalities (e.g. congenital cardiac disease) or infants during the postoperative period. Animal studies, case reports, studies not published in English and review articles without primary data were also excluded.

Search Strategy

Structured searches were made of Medline (Ovid MEDLINE and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions), Web of Science Core Collection and EBSCO-CINAHL (January 1946 to January 2020). Ovid MEDLINE search terms are included as Supplementary File 1. Bibliographies of selected papers were reviewed to find additional papers.

Study Selection

Titles and abstracts were downloaded to EndNote X9 reference management software and duplicates removed. Retained titles and abstracts were uploaded to the Rayyan QCRI web application.[15] Titles and

abstracts were independently screened and selected by two reviewers (AY and LB), with disagreements resolved by consensus. The full text of selected articles was retrieved and reviewed for inclusion and risk of bias, and data were extracted to a custom data extraction form and compiled into a database.

Risk of Bias Assessment

Risk of bias assessment was carried out according to the Joanna Biggs Institute checklist for Analytical Cross-Sectional Studies[16].

Data Analysis

The meta package (version 4.11-0)[17] for R (version 3.6.1) was used for all analyses. Where data were reported as median and range, mean and standard deviation were estimated for comparison and meta-analysis using methods described by *Wan et al*.[18] Random effects meta-analysis was selected for due to anticipated high inter-study heterogeneity (I^2 >75%) and variations in the methodological approaches. Where weight for gestational age groups were listed categorically, author definition of SGA/AGA/LGA status was accepted. Where individual patient data were provided, percentile values were calculated from UK-WHO data[19] (with infants described only as 'term' treated as being born at 40 weeks gestation). Subgroup differences were tested with the Q test for heterogeneity. Random effects meta-regression of the influence of gestational age on TBW percentage at birth of preterm infants was selected due to different methodologies and ranges of gestational ages studied. Narrative synthesis and graphical comparison were used where meta-analysis or meta-regression was not possible due to the heterogeneity of the data.

RESULTS

Study Identification

Searches identified 2349 articles after deduplication. Following screening of titles and abstracts, 128 full text articles were retrieved and assessed for eligibility. From this group, 22 studies met inclusion criteria (Fig. 1). Fifteen papers concerned full term infants[20-34] (Table 1A) and 9 concerned preterm infants[24,33,35-41] (Table 1.B), with two providing information on both term-born and preterm infants[24,33].

A. INFANTS BORN AT FULL TERM

Study	*n*	Method of TBW Assessment	TBW Percentage (mean ±SD)	Birthweight (g) (mean ±SD)	Age at Assessment	Comments
J. B. Andersen (1970)[18]	16	Deuterium oxide dilution	73.93 ± 4.02	3551 ± 622	2-4 hours after birth (9 hours for one infant)	Some mothers received diuretics during labour
Y. W. Brans (1983)[19]	109	Antipyrine space	71.72 ± 8.74	3600 ± 576	0.5 to 12 hours of life	Some infants of diabetic mothers and some macrosomic infants
G. Cassady (1971)[19,20]	26	Antipyrine space	72.95 ± 6.26	3215*	Within 19 hours of birth	Compared normal delivery and Caesarian section
D. B. Cheek, et al. (1961)[21]	11	Antipyrine space	73.18 ± 2.49	Not available	Day 3 of life	Compared diabetic and non-diabetic mothers
D. B. Cheek, et al. (1984)[22]	13	Deuterium oxide dilution	75.70 ± 3.20	2938 ± 563	12 hours of life	Compared preterm, term AGA and term SGA
D. B. Cheek, et al. (1982)[23]	56	Deuterium oxide dilution	75.79 ± 3.48	3931 ± 449	6-24 hours of life	Compared normal delivery and Caesarian section
J. R. Christian, et al. (1956)[24]	17	Antipyrine space	66.21 ± 5.00	3290 ± 330	6-42 hours of life	
W. M. Clapp, et al. (1962)[25]	12	Deuterium oxide dilution	77.40 ± 3.10	range: 2590-4985*	1-6 days of life	Compared term and preterm infants. Found lower total body water in term infants of diabetic mothers.
W. J. Cochran, et al. (1986)[26]	4	Oxygen-18 dilution	79.35 ± 6.38	2335 ± 381	2-3 days of life	Compared total body electrical conductivity with oxygen-18 technique
A. Llanos, et al. (1976)[27]	36	Deuterium oxide dilution	78.50 ± 4.62	2728*	First 30 hours of life	Included AGA and SGA infants
P. J. Offringa, et al. (1990)[28]	13	Deuterium oxide dilution	75.10 ± 5.00	3456 ± 591	First day of life	
E. Olhager, et al. (2003)[29]	9	Doubly labelled water	68.10 ± 4.10	3895 ± 565	4-11 days of life	Compared term and preterm infants
M. Osler (1960)[30]	38	Deuterium oxide dilution	73.37 ± 5.53	3476 ± 523	First day of life	Compared diabetic and non-diabetic mothers
S. C. Singhi, et al. (1995)[31]	55	Tritiated water dilution	73.70 ± 2.60	2836 ± 231	Within 6 hours of birth	Compared term and preterm infants
C. J. Thornton, et al. (1983)[32]	18	Antipyrine space	73.09 ± 2.47	3096 ± 560	Within 12 hours of birth	Compared normocythaemic and polycythaemic infants (no difference)

B. PRETERM INFANTS

Study	*n*	Mean GA (±SD)	Method of TBW Assessment	TBW Percentage (mean ±SD)	Birthweight (g) (mean ±SD)	Age at Assessment	Comments
Baarsma et al. (1992)[33] †	8	31.63 ± 3.66	Deuterium oxide dilution	83.24 ± 4.74	1683 ± 584	2-7 days (one infant at 15 days)	
Bauer et al. (1991)[34] †	8	28.00 ± 1.39††	Deuterium oxide dilution	79.20 ± 11.40	1057 ± 213	First 12 hours of life	
Bhatia et al. (1988)[35]	17	34.60 ± 1.90	Deuterium oxide dilution	74.23 ± 4.32	1990 ± 82	First 7 days of life	

Cheek et al. (1984)[22] †	5	32.50 ± 1.74††	Deuterium oxide dilution	81.90 ± 2.80	2012 ± 397	12 hours of life	Compared preterm, term AGA and term SGA
Hartnoll et al. (2000)[36]	42	27.52 ± 1.01††	Oxygen-18 dilution	83.39 ± 7.75††	1011 ± 238††	Within 18 hours of birth	Compared AGA and SGA infants
Heimler et al. (1993)[37] †	14	30.70 ± 2.40	Deuterium oxide dilution	86.60 ± 6.10	1473 ± 342	Day 1 of life	
Raghavan et al. (1988)[38]	18	27.22 ± 2.24	Oxygen-18 dilution	90.65 ± 2.82	838 ± 161	1-4 days	Validation study for bioelectrical impedance
Singhi et al. (1995)a[39] †	23	33.71 ± 1.00††	Tritiated water dilution	77.34 ± 16.80	1902 ± 242	Within 6 hours of birth	Tracked changes in TBW
Singhi et al (1995)b[31]	44	34.00 ± 1.57††	Tritiated water dilution	77.70 ± 2.60	1978 ± 412	Within 6 hours of birth	Compared term and preterm infants

Table 1. Characteristics of studies included in meta-analysis of total body water percentage in term-born neonates (**A**) and in meta-regression of total body water percentage in preterm infants (**B**). *Insufficient data to calculate or estimate mean and standard deviation. † Study measuring total body water percentage multiple times for each infant. ††Mean and standard deviation estimated from median and range using method described by *Wan et al.*[18]

1. Total Body Water Percentage of Newborn Term Infants

Fifteen studies assessed TBW in a total of 433 newborn term infants (Fig. 2A). The estimated total body water percentage was 73.8% (95% confidence interval 72.47-75.06%, $I^2 = 90\%$). The antipyrine method yielded significantly lower estimates of total body water than the deuterium oxide method (random effects Q test for subgroup differences: $p<0.01$). There were also significant differences between the total body water estimates for SGA infants (78.3%), AGA infants (73.9%) and LGA infants (68.7%) ($p<0.01$) (Fig. 2B-C).

2. Total Body Water Percentage of Preterm Infants at Birth

Mixed effects meta-regression of nine studies measuring total body water in 179 preterm newborns identified a progressive decline in percentage TBW from 85-90% between 26 and 28 weeks completed gestational age, to around 75% at 36 weeks gestational age. The estimated decline in percentage TBW was 1.44% per week (95% confidence interval 0.63-2.24%, $p<0.001$) (regression equation where $y = TBW$ and $x = gestational\ age\ in\ weeks$: $y = 127 - 1.45x$) (Fig. 3).

3. Changes in Total Body Water as Preterm Infants Grow

Five studies measured TBW multiple times for each preterm infant (marked with † in Table 1B).[24,35,36,39,41] These studies assessed 56 infants at heterogenous gestational ages and the repeated measurements were taken on different days of life (Fig. 4A). None of the studies gave detailed information on nutritional intake. *Baarsma et al*.[35] reported two measurements each for two infants, only one of whom had a measurement taken shortly after birth. The remaining four studies did not report individual patient data but summarised the gestation at birth and the day of life when measurements were taken. *Bauer et al*.[36] and *Singhi et al*.[41] focused on early body composition changes as they relate to initial weight loss in preterm infants. *Cheek et al*.[24] and *Heimler et al*.[39] measured body composition at weekly intervals. *Cheek et al*.,[24] *Heimler et al*.[39] and *Singhi et al*.[41] all identified that TBW percentage fell over time, with the one patient for whom early assessment was made in *Baarsma et al*.[35] also

exhibited a falling total body water percentage. In contrast, *Bauer et al*.[36] found an initial drop in total body water followed by a rise almost back to the first measurement.

Data from these five papers were too heterogenous to perform formal meta-analysis. Data were adjusted to plot the change in total body water against the day of life. A linear regression model, weighted for the number of infants assessed in each study, confirmed that percentage TBW fell over time (1.31% per week; 95% CI 0.53-2.09%; $p<0.01$) (Fig. 4B).

DISCUSSION

This systematic review and meta-analysis identifies the normal percentage TBW for term-born infants. Meta-regression of studies examining percentage TBW in newborn preterm infants demonstrates a negative correlation between gestation at birth and percentage TBW. Studies with repeated measurements of preterm infants were too heterogenous for formal meta-analysis but most studies identified a fall in total body water over time.

Studies included in this systematic review were mostly performed at least twenty years ago with many being published forty to sixty years ago. This is likely to limit the generalisability of these findings to modern cohorts of preterm infants. Advances in neonatal care have led to a rapidly changing phenotype of preterm infants during this period. Despite uncertainty surrounding the impact of RDS on early fluid loss[42], the surfactant era has led to significantly improved early morbidity in very preterm infants, with severe RDS now mostly limited to those born near the limits of viability. Antenatal corticosteroids have a profound influence on early fluid shifts[43] and their widespread use is likely to have significantly altered the "normal" pattern of TBW changes since publication of most included studies.

Term Infants

A strength of this review is that it includes over 400 term-born infants, whereas individual studies have rarely recruited more than 100 infants. Meta-analysis of the percentage TBW of term-born infants is limited by heterogeneity between studies. This may be due to real differences in studied populations (including demographic and nutritional differences between mothers, and differences in antenatal care) or due to differences in study protocols and methods. Furthermore, there was some variability in the timing of TBW assessment. In most cases, measurements were made within the first two days of life, but it has previously been noted that there can be rapid fluid shifts even within this limited time period.[5] Subgroup analysis confirmed that measured values of TBW are influenced by the use of antipyrine compared to deuterium oxide. In addition, there were significant differences in TBW identified between SGA, AGA and LGA infants.

Preterm Infants at Birth

This review is limited by the presence of relatively few primary research articles measuring TBW in preterm newborn infants (with only 179 subjects in this section of the review). Furthermore, data were generally reported only as the mean or median gestation at birth. Therefore, meta-analysis was performed taking this mean as the independent variable. This reduced the power and precision of the meta-regression to identify the effect of gestation at birth. This shortcoming emphasises the importance of published individual patient data.[44] However, most included studies were more than two decades old and it was unfeasible to obtain individual patient data. Similar to the data for term infants, there was also heterogeneity in methods, antenatal care and timing of TBW assessment.

Preterm Infants During Growth

There was a general paucity of data tracking TBW of preterm infants during their growth and maturation, with only 59 patients considered in this section of the meta-analysis. In addition to the sources of heterogeneity listed above, there was also very limited information on the postnatal care of these infants. It was not possible to perform formal meta-analysis of these results, although narrative synthesis identified that most papers reported a fall in percentage TBW over time. A weighted linear regression model discounting gestational age and taking week of life as the independent variable confirmed this finding. Of note, the fall in TBW of approximately 1.3% per week was just under the 1.4% figure found in the meta-regression of preterm newborn infants.

Implications for Practice

Bearing in mind the international guidance[1,2] that the body composition preterm infants at term-equivalent age should mimic that of the term-born newborn infant, the value of TBW in term-born infants identified in this study may act as a guide. However, it should be remembered that TBW reflects only one element of body composition and that derangement of fluid and electrolyte imbalance may render this comparison unreliable. This study also defines the expected TBW value for newborn preterm

infants at birth in relation to their gestation. It may be expected that preterm infants would be expected to track along this trend if they are to reach a normal percentage TBW at term-corrected age.

The small amount of data pertaining to repeated measurements of TBW in preterm infants during growth makes it difficult for current *ex utero* body water changes to be assessed. Even if more such data were available, it is likely that widespread differences in medical and nutritional management of preterm infants (as well as differences in antenatal care) would frustrate attempts to meaningfully combine data to provide an overview of the *ex utero* changes occurring in the preterm infant during growth. Significant changes in body water soon after birth may mean that preterm infants deviate from the trend in TBW identified in cross-sectional assessments of newborn infants at different gestations.

Routine and reliable measurement of body composition in clinical practice currently remains out of reach. Non-invasive measures of body composition, such as bioelectrical impedance have been used but have stalled at the validation stage.[45] Our group have shown that limb circumference measurements are feasible in preterm infants and demonstrate a predictable pattern of growth,[46] but further work is needed to assess their clinical utility in reflecting changes in body composition.

CONCLUSION

This systematic review and meta-analysis identifies potential normal values for TBW in newborn preterm and term-born infants. These findings may be useful for clinicians trying to assess changes in, fluid status, electrolyte balance, growth and body composition after birth in preterm infants. Future work should focus on tracking the body composition of preterm infants as they grow and should be contextualised by detailed monitoring of fluid management, nutritional intake and growth. Reliable and validated bedside measures of body composition in preterm infants will be required before real-time targeting of body composition can be integrated into clinical practice.

ACKNOWLEDGEMENTS: None

COMPETING INTERESTS: All authors have no conflicts of interest to disclose.

FUNDING SOURCE: The study was supported by the National Institute for Health Research Biomedical Research Centre Southampton, UK.

DATA SHARING STATEMENT: All data relevant to the study are included in the article. Data enquiries should be made to Aneurin Young (a.young@soton.ac.uk).

CONTRIBUTORSHIP: AY contributed to the conception, design, data acquisition, data analysis, data interpretation and drafting of the work. LKB contributed to the acquisition, analysis and interpretation of data. SE contributed to critical revision of the work for important intellectual content. MJJ and RMB contributed to the conception and design of the work, and to critical revision of the work for important intellectual content.

REFERENCES

1. Agostoni C, Buonocore G, Carnielli VP, et al. Enteral nutrient supply for preterm infants: commentary from the European Society of Paediatric Gastroenterology, Hepatology and Nutrition Committee on Nutrition. *J Pediatr Gastroenterol Nutr* 2010;50(1):85-91. doi: 10.1097/MPG.0b013e3181adaee0 [published Online First: 2009/11/03]
2. American Academy of Pediatrics. Committee on Nutrition., Barness LA. Pediatric nutrition handbook. 6th ed. Elk Grove Village, IL: American Academy of Pediatrics 2009.
3. Johnson MJ, Wootton SA, Leaf AA, et al. Preterm birth and body composition at term equivalent age: a systematic review and meta-analysis. *Pediatrics* 2012;130(3):e640-9. doi: 10.1542/peds.2011-3379 [published Online First: 2012/08/15]
4. Rochow N, Raja P, Liu K, et al. Physiological adjustment to postnatal growth trajectories in healthy preterm infants. *Pediatr Res* 2016;79(6):870-9. doi: 10.1038/pr.2016.15 [published Online First: 2016/02/10]
5. Bonsante F, Iacobelli S, Chantegret C, et al. The effect of parenteral nitrogen and energy intake on electrolyte balance in the preterm infant. *Eur J Clin Nutr* 2011;65(10):1088-93. doi: 10.1038/ejcn.2011.79 [published Online First: 2011/05/19]
6. Andrews ET, Beattie RM, Johnson MJ. Measuring body composition in the preterm infant: Evidence base and practicalities. *Clin Nutr* 2019 doi: 10.1016/j.clnu.2018.12.033 [published Online First: 2019/02/10]
7. Ramel SE, Gray HL, Christiansen E, et al. Greater Early Gains in Fat-Free Mass, but Not Fat Mass, Are Associated with Improved Neurodevelopment at 1 Year Corrected Age for Prematurity in Very Low Birth Weight Preterm Infants. *J Pediatr* 2016;173:108-15. doi: 10.1016/j.jpeds.2016.03.003 [published Online First: 2016/04/09]
8. Parkinson JR, Hyde MJ, Gale C, et al. Preterm birth and the metabolic syndrome in adult life: a systematic review and meta-analysis. *Pediatrics* 2013;131(4):e1240-63. doi: 10.1542/peds.2012-2177 [published Online First: 2013/03/20]
9. Tinnion R, Gillone J, Cheetham T, et al. Preterm birth and subsequent insulin sensitivity: a systematic review. *Arch Dis Child* 2014;99(4):362-8. doi: 10.1136/archdischild-2013-304615 [published Online First: 2013/12/24]
10. Menon G, Davidson AL, Drake AJ, et al. Is preterm nutrition a trade-off between head and heart? *Arch Dis Child Fetal Neonatal Ed* 2019;104(3):F232-F34. doi: 10.1136/archdischild-2018-315672 [published Online First: 2018/10/14]
11. Uthaya S, Thomas EL, Hamilton G, et al. Altered adiposity after extremely preterm birth. *Pediatr Res* 2005;57(2):211-5. doi: 10.1203/01.PDR.0000148284.58934.1C [published Online First: 2004/12/22]
12. Friis-Hansen BJ, Holiday M, Stapleton T, et al. Total body water in children. *Pediatrics* 1951;7(3):321-27.
13. Edelman IS, Haley HB, Schloerb PR, et al. Further observations on total body water. I. Normal values throughout the life span. *Surgery Gynecology and Obstetrics* 1952;95(1):1-12.
14. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6(7):e1000097. doi: 10.1371/journal.pmed.1000097 [published Online First: 2009/07/22]
15. Ouzzani M, Hammady H, Fedorowicz Z, et al. Rayyan-a web and mobile app for systematic reviews. *Syst Rev* 2016;5(1):210. doi: 10.1186/s13643-016-0384-4 [published Online First: 2016/12/07]
16. Moola S, Munn Z, Tufanaru C, et al. Chapter 7: Systematic reviews of etiology and risk. In: Aromataris E, Munn Z, eds. Joanna Briggs Institute Reviewer's Manual: The Joanna Briggs Institute, 2017.
17. Balduzzi S, Rucker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health* 2019;22(4):153-60. doi: 10.1136/ebmental-2019-300117 [published Online First: 2019/09/30]
18. Wan X, Wang W, Liu J, et al. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med Res Methodol* 2014;14:135. doi: 10.1186/1471-2288-14-135 [published Online First: 2014/12/20]
19. RCPCH. UK-WHO growth charts - neonatal and infant close monitoring (NICM) London, UK: Royal College of Paediatrics and Child Health; 2019 [Available from: https://www.rcpch.ac.uk/resources/uk-who-growth-charts-neonatal-infant-close-monitoring-nicm.
20. Andersen JB. The effect of diuretics in late pregnancy on the newborn infant. *Acta Paediatrica Scandinavica* 1970;59(6):659-63.
21. Brans YW, Shannon DL, Hunter MA. Maternal diabetes and neonatal macrosomia. III. Neonatal body water estimates. *Early Human Development* 1983;8(3):307-16.
22. Cassady G. Effect of cesarean section on neonatal body water spaces. *N Engl J Med* 1971;285(16):887-91.
23. Cheek DB, Maddison TG, Malinek M, et al. Further observations on the corrected bromide space of the neonate and investigation of water and electrolyte status in infants born of diabetic mothers. *Pediatrics* 1961;28(6):861.
24. Cheek DB, Wishart J, MacLennan AH, et al. Cell hydration in the normally grown, the premature and the low weight for gestational age infant. *Early Human Development* 1984;10(1-2):75-84.
25. Cheek DB, Wishart J, MacLennan AH, et al. Hydration in the first 24 h of postnatal life in normal infants born vaginally or by ceasarean section. *Early Human Development* 1982;7(4):323-30.

26. Christian JR, Jackson H, Karazeris MP, et al. Total body water and exchangeable sodium in normal full-term newborn infants. *AMA J Dis Child* 1956;92(4):325-9.
27. Clapp WM, Butterfield LJ, O'Brien D. Body water compartmeate im the premature iaiaat, with special referece to the effect en tee respiratory duttrees syndrome aad off aaatenuu diabete mud toxemia. *Pediatrics* 1962;29(6):883-89.
28. Cochran WJ, Klish WJ, Wong WW, et al. Total body electrical conductivity used to determine body composition in infants. *Pediatric Research* 1986;20(6):561-4.
29. Llanos A, Segovia N, Gallegos R, et al. DEUTERIUM-OXIDE SPACE AS AN INDEX OF TOTAL-BODY WATER IN FETAL GROWTH RETARDED NEWBORN-INFANTS. *Journal of Pediatrics* 1976;89(2):323-24.
30. Offringa PJ, Boersma ER, Brunsting JR, et al. Weight loss in full-term negroid infants: relationship to body water compartments at birth? *Early Human Development* 1990;21(2):73-81.
31. Olhager E, Forsum E. Total energy expenditure, body composition and weight gain in moderately preterm and full-term infants at term postconceptional age. *Acta Paediatr* 2003;92(11):1327-34.
32. Osler M. Body water of newborn infants of diabetic mothers. *Acta Endocrinol (Copenh)* 1960;34:261-76.
33. Singhi SC, Ganguli NK, Bhakoo ON, et al. Body water distribution in newborn infants appropriate for gestational age. *Indian Journal of Medical Research* 1995;101:193-200.
34. Thornton CJ, Shannon DL, Hunter MA, et al. Body water estimates in neonatal polycythemia. *Journal of Pediatrics* 1983;102(1):113-7.
35. Baarsma R, Hof P, Zijlstra WG, et al. Measurement of total bodywater volume with deuterated water in newborn infants. *Biology of the Neonate* 1992;62(2-3):108-12.
36. Bauer K, Bovermann G, Roithmaier A, et al. Body composition, nutrition, and fluid balance during the first two weeks of life in preterm neonates weighing less than 1500 grams. *Journal of Pediatrics* 1991;118(4 Pt 1):615-20.
37. Bhatia J, Rassin DK. Growth and total body water in premature infants fed "in-utero" or "ex-utero". *Acta Paediatrica Scandinavica* 1988;77(3):326-31.
38. Hartnoll G, Betremieux P, Modi N. Body water content of extremely preterm infants at birth. *Archives of Disease in Childhood Fetal & Neonatal Edition* 2000;83(1):F56-9.
39. Heimler R, Doumas BT, Jendrzejczak BM, et al. Relationship between nutrition, weight change, and fluid compartments in preterm infants during the first week of life. *Journal of Pediatrics* 1993;122(1):110-4.
40. Raghavan CV, Super DM, Chatburn RL, et al. Estimation of total body water in very-low-birth-weight infants by using anthropometry with and without bioelectrical impedance and H2[(18)O]. *American Journal of Clinical Nutrition* 1998;68(3):668-74.
41. Singhi S, Sood V, Bhakoo ON, et al. Composition of postnatal weight loss & subsequent weight gain in preterm infants. *Indian Journal of Medical Research* 1995;101:157-62.
42. Tang W, Ridout D, Modi N. Influence of respiratory distress syndrome on body composition after preterm birth. *Arch Dis Child Fetal Neonatal Ed* 1997;77(1):F28-31. doi: 10.1136/fn.77.1.f28 [published Online First: 1997/07/01]
43. Dimitriou G, Kavvadia V, Marcou M, et al. Antenatal steroids and fluid balance in very low birthweight infants. *Arch Dis Child Fetal Neonatal Ed* 2005;90(6):F509-13. doi: 10.1136/adc.2005.071688 [published Online First: 2005/05/10]
44. Stewart LA, Parmar MK. Meta-analysis of the literature or of individual patient data: is there a difference? *Lancet* 1993;341(8842):418-22. doi: 10.1016/0140-6736(93)93004-k [published Online First: 1993/02/13]
45. Lingwood BE. Bioelectrical impedance analysis for assessment of fluid status and body composition in neonates - The good, the bad and the unknown. *European Journal of Clinical Nutrition* 2013;67(SUPPL. 1):S28-S33. doi: http://dx.doi.org/10.1038/ejcn.2012.162
46. Ashton JJ, Johnson MJ, Pond J, et al. Assessing the growth of preterm infants using detailed anthropometry. *Acta Paediatr* 2017;106(6):889-96. doi: 10.1111/apa.13804 [published Online First: 2017/02/28]

FIGURE 1

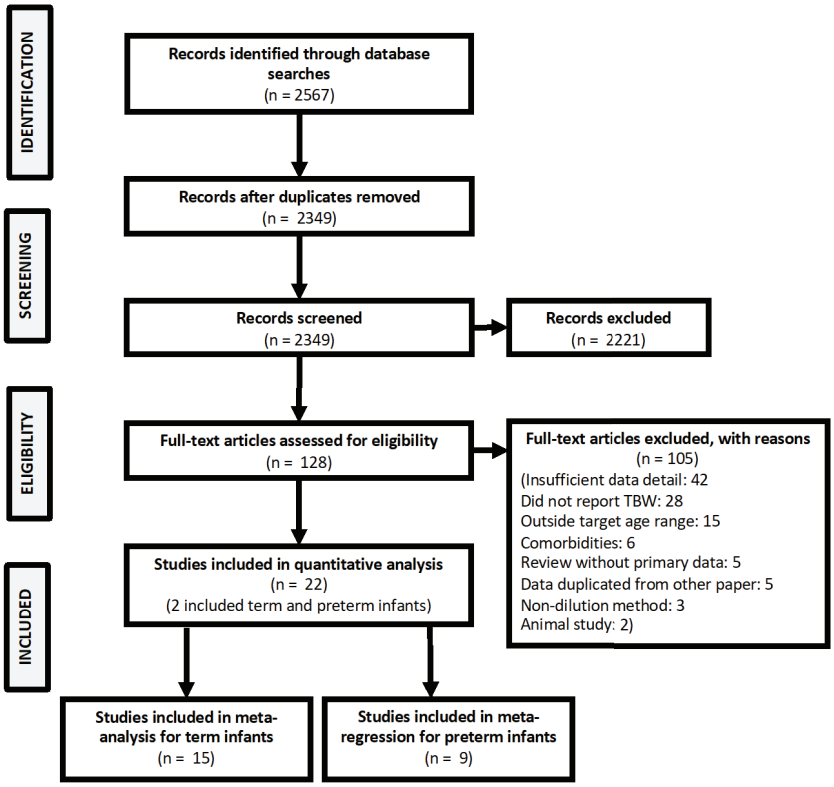


Figure 1. PRISMA flow chart of screening and selection of studies, with reasons for exclusion based on full text review.

FIGURE 2

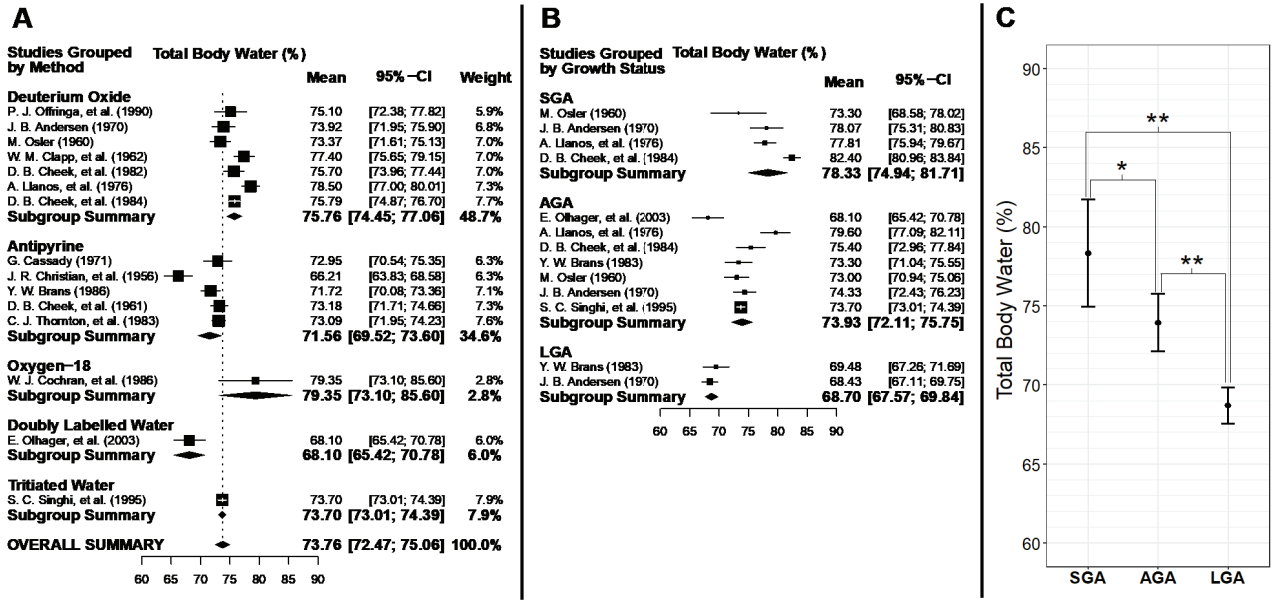


Figure 2. Total body water percentage in newborn full term infants. **A.** Forest plot with methods of TBW assessment as subgroups. **B.** Forest plot with weight for gestational age subgroups (overall test for subgroup differences: $p<0.01$) and **C.** Mean total body water percentage (±95% confidence interval). SGA, small for gestational age; AGA, appropriate for gestational age; LGA, large for gestational age. *$p<0.05$, **$p<0.01$.

FIGURE 3

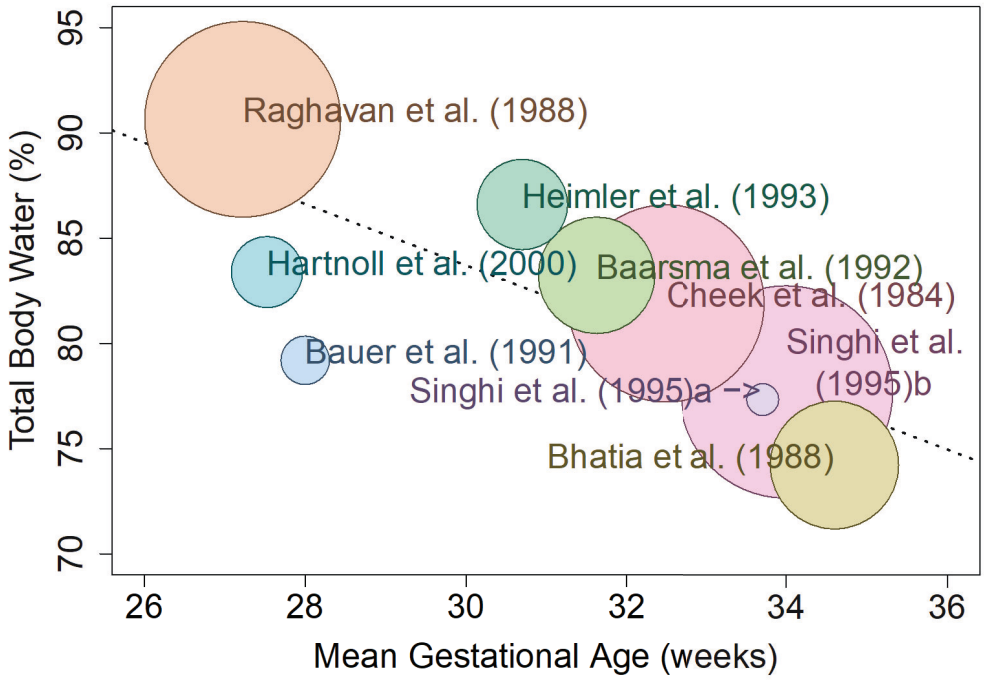


Figure 3. Bubble plot with fitted meta-regression line of total body water percentage at birth and mean gestational age in infants included in each study.

Circles are sized according to precision of the estimate, larger circles indicating greater precision. Regression equation: y=127-1.45x.

FIGURE 4

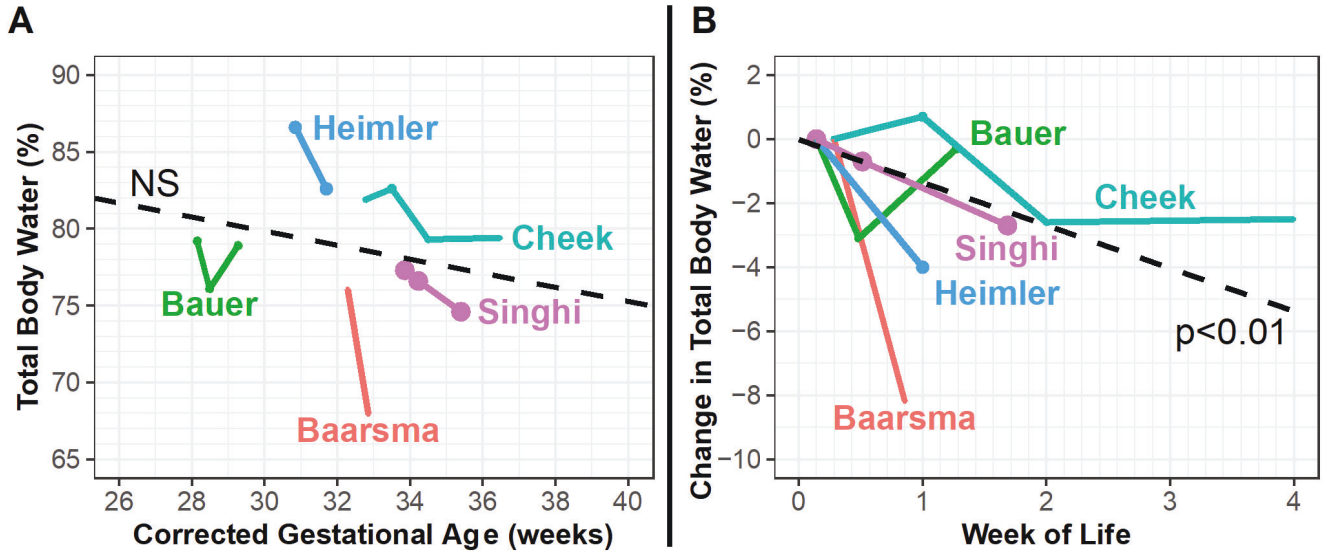


Figure 4.A. Total body water percentage in 58 preterm infants in 5 studies where repeated measurements of total body water were taken during their growth and maturation. Sizes of points indicate number of infants assessed. Points are plotted at the mean gestational age for infants in each study. Dashed line: linear regression line weighted for study size (slope value not significant). **B.** Change in total body water percentage in preterm infants after birth, grouped by study. Sizes of points indicate number of infants assessed. Dashed line: linear regression line weighted for study size (y=-0.13-1.31x).

Appendix 3

Published paper: Young A, Beattie RM, Johnson MJ. Optimising growth in very preterm infants: reviewing the evidence. Arch Dis Child Fetal Neonatal Ed. 2022

Optimising Growth in Very Preterm Infants - Reviewing the evidence

Aneurin Young[a,b] (ORCID: 0000-0003-3549-3813, Twitter: @AneurinYoung), R Mark Beattie[c] (0000-0003-4721-0577, @RMBeattie50), Mark J Johnson[a,b] (0000-0003-1829-9912, @MarkJohnson800).

[a]Department of Neonatal Medicine, Southampton Children's Hospital, University Hospital Southampton NHS Foundation Trust, UK; [b]NIHR Southampton Biomedical Research Centre, University Hospital Southampton NHS Foundation Trust and University of Southampton, UK; [c]Department of Paediatric Gastroenterology, Southampton Children's Hospital, University Hospital Southampton NHS Foundation Trust, UK;

Corresponding Author: Aneurin Young, Department of Neonatal Medicine, University Hospital Southampton NHS Foundation Trust, Coxford Road, Southampton, SO16 5YA, UK [a.young@soton.ac.uk] 02380777222

Word Count: 3143

This article has been accepted for publication in Archives of Disease in Childhood: Fetal and Neonatal Edition (2022) following peer review, and the Version of Record can be accessed online at: http://dx.doi.org/10.1136/archdischild-2021-322892 .

ABSTRACT

Infants born before 32 weeks postmenstrual age are at a high risk of growth failure. International guidelines have long recommended that they match the growth of an equivalent fetus, despite the challenges posed by ex-utero life and comorbidities of prematurity. Several groups have recently questioned the necessity or desirability of this target, shifting attention to aiming for growth which optimises important long-term outcomes. Specifically, recent research has identified the neurodevelopmental benefits of enhanced growth during the neonatal period, but work in term infant suggests that rapid growth may promote the metabolic syndrome in later life. In this context, defining a pattern of growth which optimises outcomes is complex, controversial and contested.

Even if an optimal pattern of growth can be defined, determining the nutritional requirements to achieve such growth is not straightforward, and investigations into the nutritional needs of the very preterm infant continue. Furthermore, each infant has individual nutritional needs and may encounter a number of barriers to achieving good nutrition.

This article offers a narrative review of recent evidence for the competing definitions of optimal growth in this cohort. It examines recent advances in the determination of macronutrient and micronutrient intake targets along with common barriers to achieving good nutrition and growth. Finally, key implications for clinical practice are set out and a recommendation for structured multidisciplinary management of nutrition and growth is illustrated.

INTRODUCTION

Preterm infants are vulnerable to poor growth. Transfer of nutrients from mother to infant during the third trimester of pregnancy is interrupted, and their postnatal nutritional intake is impaired by gut and metabolic immaturity. They are also prone to comorbidities which can also impair growth. Furthermore, defining what constitutes optimal growth (and conversely poor growth) has proven controversial.

In this narrative review, we aim to outline the importance of growth, to explore the definitions of optimal growth, to outline the evidence underpinning current nutritional best practice and to make practical recommendations for improving nutrition and growth.

Menon and colleagues have previously asked whether preterm nutrition is a trade-off between head and heart.[1] In other words, does enhanced nutrition improve neurodevelopmental outcomes at the expense of poorer cardiovascular health in later life? Over the last ten years, an increasing body of literature has shown that better weight gain in preterm infants is associated with better neurodevelopmental outcomes.[2-4] Belfort and co-workers have demonstrated that greater weight gain before term corrected age is associated with better neurodevelopmental outcomes.[3] They also identified that the most critical period for this effect is growth prior to and shortly after term equivalent age.[2]

At the same time, rapid weight gain in term infants is associated with development of the metabolic syndrome, a disease which is prevalent amongst ex-preterm infants. However, there are some data to suggest that rapid weight gain in early infancy for preterm infants is not associated with poorer metabolic outcomes, but that weight gain during subsequent phases of growth is a significant risk factor.[5] Recent results from the large observational EPICure study strongly suggest that the critical growth period for metabolic health is significantly later in childhood, between 2.5 and 6 years, whilst rapid growth prior to that period is not associated with an elevated risk.[6] Taken in conjunction, these neurodevelopmental and metabolic findings suggest that improved growth during early infancy, negating the impetus for later catch-up growth, may optimise neurodevelopmental outcomes without compromising metabolic health.

APPROACHES TO GROWTH STANDARDS

Guidelines set by the European Society of Paediatric Gastroenterology, Hepatology and Nutrition (ESPGHAN)[7] and the American Academy of Pediatrics (AAP)[8] recommend that growth mimics that of the equivalent fetus in utero. However, preterm infants have consistently been found to grow more slowly than fetuses of equivalent postmenstrual age,[9] although this effect can be mediated by different nutritional approaches.[10] Additionally, preterm infants show marked differences in body composition compared to term-born infants,[11] therefore failing to achieve either the rate of

growth or the composition of weight gain called for by the two societies. Our recent meta-analysis identified the pattern of changing total body water percentage in newborns depending on gestation, with total body water falling from 90% to around 75% between 26 weeks gestation and term.[12]

Preterm infant growth standards can most easily be defined by taking measurements of newborn preterm infants. However, several groups have recently questioned this approach and the concept that ex utero growth should mimic the growth of a fetus in utero, leading to interest in redefining expected growth using other techniques.

The INTERGOWTH-21st Project

In 2014, the INTERGROWTH-21st Consortium monitored fetal growth and birthweight across an international cohort for whom comorbidity data were collected and pregnancies with confirmed fetal growth restriction were excluded. Birthweight data were used to form expected growth curves (Figure 1) but included only 382 subjects below 32 weeks gestation (of whom only 82 were below 28 weeks).[13] Data for length and head circumference were similarly limited.

Another arm of the INTERGROWTH project took an empirical approach, targeting growth which matches those of preterm infants who have suffered minimal postnatal complications. They recommend using standards derived from longitudinal data in low-risk preterms (Figure 1).[14]

The Southampton Preterm Cohort

We have previously shown that the growth of infants in Southampton can keep up with birthweight-derived growth standards.[10] These local Southampton data were used to generate a growth standard from a cohort of over 200 very preterm infants.[15] The resulting growth standard is based on repeated measurement of preterm infants in the context of fully described nutritional intakes (fig 1). A web application was also created to support the resultant charts (www.bit.ly/sotongrowth). Growth curves generally tracked close to UK-WHO growth standards generated from cross-sectional birthweight data.

Individualised Growth Targets

Landau-Crangle and co-workers have approached the question physiologically, arguing that adaptive early loss of body water will place the optimally-growing preterm infant on a lower centile followed by growth at a velocity such that the birth centile is matched only after the equivalent water loss of the term infant.[16] That group recommends individualised growth targets taking these factors into account (Figure 1).

Figure 1 illustrates the expected pattern of weight gain for an example infant born at 27 weeks postmenstrual age weighing 1kg. It compares the expected growth trajectory for a range of

approaches detailed above. Many of these curves match closely, with the individualised curve anticipating slower growth at first with subsequent convergence with the other tracks, and the INTERGROWTH longitudinal curve recommending more rapid growth throughout.

Table 1. Guideline recommendations for enteral and parenteral nutritional intakes for stable preterm infants along with the range nutritional intakes from some typical feeds (at 150ml/kg/day) and a commercially available parenteral nutrition product.

		ENTERAL FEEDING GUIDELINES		ENTERAL FEEDS					PARENTERAL FEEDING GUIDELINES		PARENTERAL FEEDING PRODUCT
		Koletzko et al.[17]	ESPGHAN[7]	Breastmilk (AAP[8])	Fortified Breastmilk	Preterm Formula	Preterm Follow-on Formula	Term Formula	ESPGHAN (2018)[18]	NICE[19]	Commercially Available Preterm PN
Nutrient	**Unit**	(per kg/day)	(per kg/day)	(per 150ml)	(per 150ml)	(per 150ml)	(per 150ml)	(per 150ml)	(per kg/day)	(per kg/day)	(per kg/day)
Energy	kcal	110-130	110-135	~100	120-130	120	110	100	90-120	75-120	116
Macronutrients											
Protein / Amino Acids	g	3.5-4.5	4-4.5 (<1kg body weight) 3.5-4 (>1kg body weight)	~1.4	3-4.5	4-4.5	3	2-2.2	2.5-3.5	3-4	3.8
Lipid	g	4.8-6.6	4.8-6.6	~5	5-6	5.5-6	5.5-6	5	3-4	3-4	3.2
Carbohydrate	g	11.6-13.2	11.6-13.2	~12	13-16	12-13	10.5-11.5	11-12	8-10	9-16	17
Micronutrients											
Sodium	mg	69-115	69-115	~20-40	70-90	80-100	40-50	30-50	46-115	Responsive*	165
	mmol	3-5	3-5	~1-2	3-4	3.5-4.5	1.8-2.4	1-2	2-5		2.8
Potassium	mg	78-195	66-132	~60-80	120-150	120-170	120-130	100-140	78-117	Responsive*	103
	mmol	2-5	1.7-3.4	1.5-2	2.5-3.5	3-4.5	3-3.5	2.5-3.5	2-3		2.6
Calcium	mg	120-200	120-140	~30-40	130-150	150-180	125	70	64-140	60-80	65
	mmol	3-5	3-3.5	0.7-1	3.5-4	4-4.5	3.1	1.7	1.6-3.5	1.5-2	1.6
Phosphorus	mg	60-140	60-90	~20	80-90	100-120	70-75	40-50	50-108	60	50
	mmol	2-4.5	2-3	0.6	2.5-3	3-4	2-2.5	1.3-1.6	1.6-3.5	2	1.6
Iron	mg	2-3	2-3	~0.05-0.14	0-3**	2-2.5	1.1-1.8	0.5-0.8	0	0	0
Zinc	mg	1.4-2.5	1.1-2	~0.15-0.5	1-2	1.5-2	1.3	0.8	0.4-0.5	NG	0.25
Vitamin A	µg RE	400-1100	400-1000	~50-90	400-600	500-550	100-150	90	227-455	NG	276
Vitamin D	IU	400-1000	800-1000	~0.05	250-300	180-200	100	90	80-400	NG	160
Vitamin E	mg α-TE	2.2-11	2.2-11	~0.5-1.2	5-7	5.5-7	1.3-3	1.7-2.1	2.8-3.5	NG	2.6
Vitamin K1	µg	4.4-28	4.4-28	~0.3-0.5	10-15	9-10	6-9	5-7	10	NG	80
Choline	mg	8-55	8-55	NG	3-3.5	30-40	35	33	NG	NG	NG
DHA	mg	55-60	12-30	NG	2-2.5	30-40	30	25	42	NG	NG

RE, retinol equivalents; α-TE, α-tocopherol equivalents; NG, no guidance given or not defined in product information

* Guidance recommends adjusting intake based on clinical sampling. ** Some fortifiers do not contain iron

ESTABLISHING NUTRITIONAL REQUIREMENTS

Notwithstanding difficulties in defining optimal growth, in practice it is necessary to set target nutrient intakes which may then be adjusted for individual infants based on their clinical status and growth trajectory. Table 1 presents several international guidelines for nutrient intake along with the nutritional provision of some typical feeds and parenteral nutrition products. Different methods have been used to define the recommended nutrient intakes. Historically, the factorial method aimed to estimate nutritional requirements by studying the changing composition of fetuses at different gestations.[20] More recently, experimental studies have examined the impact of changes in nutritional approach.

Experimental Approaches to Macronutrient Intakes

Several recent trials have randomized preterm infants to receive differing nutritional regimens. The NEON study assessed the impact of immediate or incremental increases in amino acid intakes and lipid emulsions during the first few days of life, with the control arm receiving up to 2.7g/kg/day amino acid by day 3 of life and the intervention arm receiving 3.6g/kg/day amino acid from day 1.[21] This study did not identify any differences in its primary outcomes of non-adipose mass and intrahepatocellular lipid. The group with higher early provision of amino acids had lower growth of head circumference 0.8 cm adjusted mean difference ($p=0.02$). The study was criticised for the minimal effect of allocation on actual macronutrient intake over the course of neonatal stay.[22]

The SCAMP Trial randomized infants to a control parenteral nutrition product or an intervention product which delivered more amino acid, lipid and glucose with a resultant higher energy provision throughout the period of parenteral nutrition.[23] Similarly to the NEON study, the control group were managed with PN containing amino acid at 2.8g/kg/day and the intervention group received amino acid at 3.8g/kg/day. However, the exposure to the allocated PN was for a significantly longer period, meaning that the total differences in nutrient intake were more pronounced. The group receiving more macronutrients had better head circumference growth to day 28 of life (the primary outcome, mean difference 5mm, $p<0.001$) and this persisted to 36 weeks corrected gestational age. Weight was unaffected as were all other tested clinical outcomes.

Similarly, a Norwegian group demonstrated an increased growth velocity to 36 weeks corrected GA in response to an enhanced supply of amino acid (comparing 3.2g/kg/day and 4g/kg/day amino acid intake), lipid and energy (17.4 vs 14.3 g/kg/day, $p<-0.001$).[24] A Dutch group randomized infants to a range of intakes of amino acids (2.4-3.6g/kg/day) and types of lipid, with one group (high amino acids and mixed lipid emulsion) demonstrating greater weight gain at two years corrected age but without any differences in neurodevelopment (the primary outcome).[25]

Each of these studies were performed before the most recent ESPGHAN parenteral nutrition guidelines. At the time of the studies, ESPGHAN recommended 1.5-4g/kg/day amino acid intake.[26] Therefore, they compared amino acid intakes within the recommended range at the time of the studies. It remains possible that substantially higher nutrient intakes than this may precipitate excess growth or detrimental derangement of body composition, although this possibility has not been tested in large studies to date. Taken together, these studies suggest that higher rates of nutritional intake lead to improved growth, although an impact on neurodevelopment remains unclear.

Experimental Approaches to Specific Micronutrients

A number of micronutrients have recently come under scrutiny, either for a possible general effect on growth or as targets for improving other specific clinical outcomes.

Trials of high doses of vitamin D supplementation have been shown to improve radiological markers of bone mineralization and to increase weight (13.6 vs 16.4g/day, $p<0.01$) and length gain (0.69 vs 0.79cm/week, $p=0.02$)[27]. Vitamin A supplementation has been shown to improve a marker of retinal function.[28]

Choline and docosahexaenoic acid are implicated in phosphatidylcholine metabolism and are found in high concentrations in fetal plasma, falling rapidly after preterm birth. A small trial has shown that supplementation with choline is practical and can restore plasma choline to near fetal concentrations although further work is needed to assess any potential impact on growth or neurodevelopment.[29]

A Cochrane review of LCPUFA supplementation in preterms found no proven effect.[30] Zinc has also risen to prominence in recent years, with a Cochrane review suggesting that enteral supplementation with zinc is likely to improve growth and reduce mortality.[31] This is particularly significant given that commonly available parenteral and enteral nutrition products typically deliver markedly insufficient amounts of zinc (Table 1), and zinc deficiency is common in preterm infants.[32]

BARRIERS TO ACHIEVING NUTRITIONAL TARGETS

Enteral Feed Tolerance

Feed intolerance is common and can be considered to be universal to some extent in the extremely preterm infant, most of whom require parenteral nutrition support at least in the first few weeks. However, parenteral nutrition presents significant risks, including central line associated bloodstream infection and cholestasis, as well as carrying significant financial cost. These problems have led to recent interest in accelerating enteral feed increments and restricting

the range of infants who receive parenteral nutrition. A recent systematic review highlighted the difficulties in defining feed intolerance, with inconsistent definitions frustrating attempts to formulate a consensus definition of feed intolerance in the preterm population.[33]

The SIFT trial concluded that increasing feeds at 30ml/kg/day (compared to 18ml/kg/day) reduced time to reaching full feeds.[34] The primary outcome of survival without moderate or severe neurodisability showed no difference between the groups. However, faster feeding was associated with an increased risk of moderate or severe motor impairment (adjusted effect 1.48, CI 1.02-2.14). An associated cost analysis identified that this excess of motor impairment meant that faster feeds are both clinically and economically undesirable.[35]

Observational data have demonstrated an association between early passage and clearance of meconium with improved enteral feed tolerance.[36] However, meta-analysis of studies aiming to improve enteral food tolerance by the prophylactic use of enemas or suppositories identified no effect of these interventions.[37]

The FEED1 trial is currently investigating whether giving full enteral feeds from the first day of life will decrease length of stay for infants born from 30^{+0} to 32^{+6} weeks gestation.[38]

Metabolic Tolerance

Metabolic disturbance is more common in the most preterm infants and in those with intrauterine growth restriction. The substantial energy needs of the preterm infant (Table 1) require the delivery of a significant load of carbohydrate, protein and lipid. However, these infants are prone to hyperglycaemia in the early neonatal period, which in turn have been associated with an increased risk of death, poor growth and most major morbidities associated with prematurity, although it is difficult to prove a causal link given the presence of likely confounding factors.[39] Technological advancements in continuous glucose monitoring have been shown to improve glycaemic control but the impact on outcomes remain uncertain.[40] Similarly, hypertrigyceridaemia is common at intravenous lipid delivery levels meeting nutritional requirements and is associated with poorer clinical outcomes.[41,42]

Parenteral Nutrition Limitations

Current formulations of parenteral nutrition often do not meet target or recommended nutrient requirements, especially for micronutrient minerals, particularly calcium and phosphate. In part, this is due to concerns about stability of these substances in solution and the possibility of precipitation. Studies continue in this area, especially as there is a pressing need to optimise calcium and phosphate delivery to prevent metabolic bone disease of prematurity.[43]

Sepsis and Inflammation

Preterm infants frequently experience episodes of inflammation, both from infections and from other causes, including surgical interventions. Infection is common, with around 10% of preterm infants experiencing late onset infection (Vermont Oxford Network VLBW cohort).[44] Acute inflammation profoundly alters the metabolic state of the preterm infant, driving catabolism, insulin resistance and suppression of growth factors such as IGF-1.[45] This is likely to lead to less effective nutrient metabolism with usual or increased nutrition in this context likely to drive hyperglycaemia and hypertriglyceridaemia without contributing to growth. This theoretical problem is reflected in well-established findings in critically ill adults and children, where early aggressive parenteral nutrition during acute illness is deleterious.[46,47]

Fluid Restriction

Newborn infants have a limited capacity for diuresis and so fluid intake is often limited during the first few days of life. In addition, fluid restriction may be part of medical management, for example in the presence of patent ductus arteriosus. Even once total fluid restriction is relaxed, there is often a period during which breastmilk replaces much more energy-dense parenteral nutrition. These multiple restrictions of fluid intake inevitably limit delivery of nutrition. These difficulties may be addressed by strategies including increasing the concentration of parenteral nutrition (as recommended by NICE)[48] and by earlier initiation of breastmilk fortification.

Nutritional Content of Breastmilk

Mother's own breastmilk provides substantial benefits to the preterm infant and is recommended as the ideal basis for enteral feeding.[7,8] Using breastmilk in preference to formula also significantly reduces the risk of necrotising enterocolitis. However, breastmilk alone cannot provide adequate nutritional intakes and hence multicomponent fortification has been widely adopted. A Cochrane review in 2004 recommended routine fortification as it improves short-term growth and identified no increase in adverse events related to its use, albeit with insufficient long-term follow-up data to reach a conclusion on neurodevelopmental outcomes.[49]

Breastmilk fortifier is typically formulated using extensively hydrolysed cow's milk protein. During the last decade, milk fortification products based on donated human milk have been developed. Initial studies establishing the use of the first of these products were troubled by design flaws and there is ongoing controversy surrounding its potential benefits and costs.[50] A recent systematic review and meta-analysis concluded that there is a suggestion of decreased risk of necrotising enterocolitis with human milk-based fortifier but that the overall quality of evidence is low and so its routine use cannot currently be recommended.[51]

Individualised Breastmilk Fortification

Maternal milk constitution and infant nutritional requirements are both highly variable. Therefore, attempts have been made to personalise breastmilk fortification to adjust breastmilk nutritional contents to prespecified values[52] or in response to infant blood urea level,[53] or both.[54] A study adjusting fortification in response to breastmilk analysis demonstrated an improvement in weight gain and a trend towards improved linear growth in the intervention arm.[52] Altering fortification in response to blood urea has shown promise in improving growth,[53] although there was significantly higher protein provision to the intervention group, meaning that it is difficult to know whether increased protein or personalisation per se was the important factor. A Cochrane review identified that targeted fortification improved weight, length and head growth during initial neonatal stay but that there was insufficient evidence for other outcomes.[55]

Donor Breastmilk

Milk banking has increased the availability of donated breastmilk throughout North America and Europe during the last decade. A recent Cochrane review addressed many questions relating to the relative safety and efficacy of fortified donor breastmilk compared to preterm formula.[56] Weight and length gain were better in the formula-fed group, with no difference in head growth or neurodevelopmental outcomes. Necrotising enterocolitis was more common in the formula fed group (risk ratio 1.87, 95%CI 1.23-2.85).

PRACTICAL MANAGEMENT OF NUTRITION AND GROWTH

Monitoring the growth of very preterm babies presents a number of challenges. Clinically unstable infants may be difficult to remove from their incubators to measure, although incubator scales can be used effectively.[57] Head circumference measurements may be difficult in the presence of respiratory support and length measurements may be difficult to perform accurately. In addition, fluid shifts may cause difficulty in interpreting weight gains and losses.

Monitoring the growth of preterm infants and tailoring their nutritional interventions is a complex task which requires the expertise of a multidisciplinary team, involving doctors, nurses, dietitians, pharmacists and other team members depending on the individual infant. There have been successful efforts to implement routine nutritional risk assessment, growth reviews and multidisciplinary shared decision-making based on comprehensive nutritional guidelines[58] with a resultant weight gain pattern which more closely follows birthweight-derived UK-WHO growth curves.[10,15] Figure 2 provides an outline of the factors to be considered during multidisciplinary review of growth and nutrition.

Implications for Practice

This evidence review highlights several practical interventions which have been shown to improve growth in the very preterm infant:

- Weight, length and head circumference should be routinely measured, at least weekly for length and more frequently for weight and head circumference
- Measurements should be plotted on a growth chart derived from a growth standard appropriate to the population in question. Further research is required to confirm whether individualised growth trajectories are preferable to standard birthweight-derived charts.
- There should be a standardised approach to the provision of enteral and parenteral nutrition, which is designed to meet published nutritional requirements (see table 1).
- Enteral or parenteral nutrition should be started as soon after birth as feasible
- Enteral feeding should be based on human breastmilk whenever possible.
- Breastmilk fortifier should be used to supplement human breastmilk so that it meets published nutritional requirements. Current data are insufficient to recommend human milk-based fortifier in preference to cows milk-based fortifier.
- Feeding increments of 18-30ml/kg/day seem reasonable, but the lower end of this range should be used for the most premature infants.
- Parenteral nutrition products should be formulated at concentrations which optimise nutritional intake, especially when total fluid intake is restricted.
- In the absence of strong evidence in neonatal populations, a temporary reduction in nutritional intake during acute inflammatory states should be considered in term infants based on evidence from children and adults. More data is needed to determine the best approach in preterm infants.
- There should be regular multidisciplinary monitoring of growth and planning of nutritional management to infants at risk of nutritional compromise.

CONCLUSION

Optimal growth is difficult to define and to deliver to very preterm infants. Whilst it is possible to set general targets for growth and nutritional intake, the requirements of any individual infant are defined by a set of complex and interacting factors. Future research is likely to focus on determining factors which can be used to tailor individualised approaches to nutrition, thereby optimising growth, avoiding morbidity and promoting health and neurodevelopment into childhood.

FIGURE LEGENDS

Figure 1.Examples of expected growth curves for a male infant born at 27 weeks weighing 1kg, generated by tracking a constant centile from the INTERGROWTH birthweight and postnatal growth standards,[14] the Fenton growth chart,[59] UK-WHO growth chart[60], growth chart generated from a Southampton cohort[15] and by calculating an individualised growth trajectory as per Landau-Crangle et al.[16] Figure created by the first author.

Figure 2. A multidisciplinary approach to growth assessment and decision-making. Figure created by the first author.

REQUIRED STATEMENTS

ACKNOWLEDGEMENTS: None

COMPETING INTERESTS: All authors have no conflicts of interest to disclose.

FUNDING SOURCE: The study was supported by the National Institute for Health Research Biomedical Research Centre Southampton, UK.

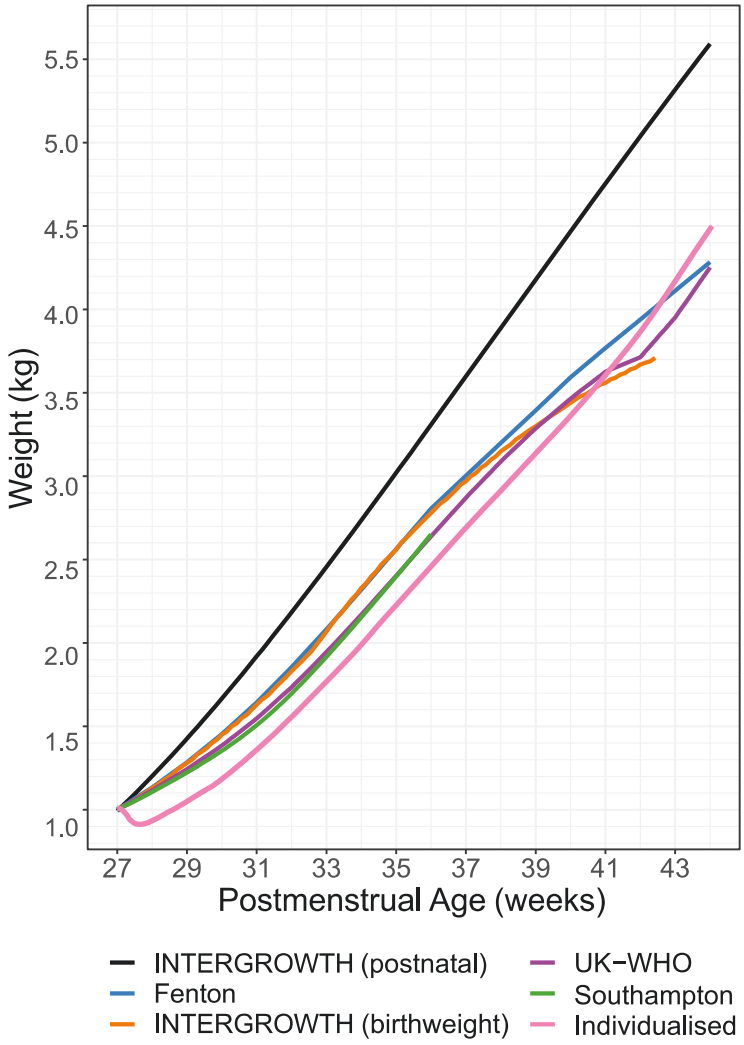
CONTRIBUTORSHIP: AY contributed to the conception, design and drafting of the work. RMB and MJJ contributed to the conception and design of the work, and to critical revision of the work for important intellectual content.

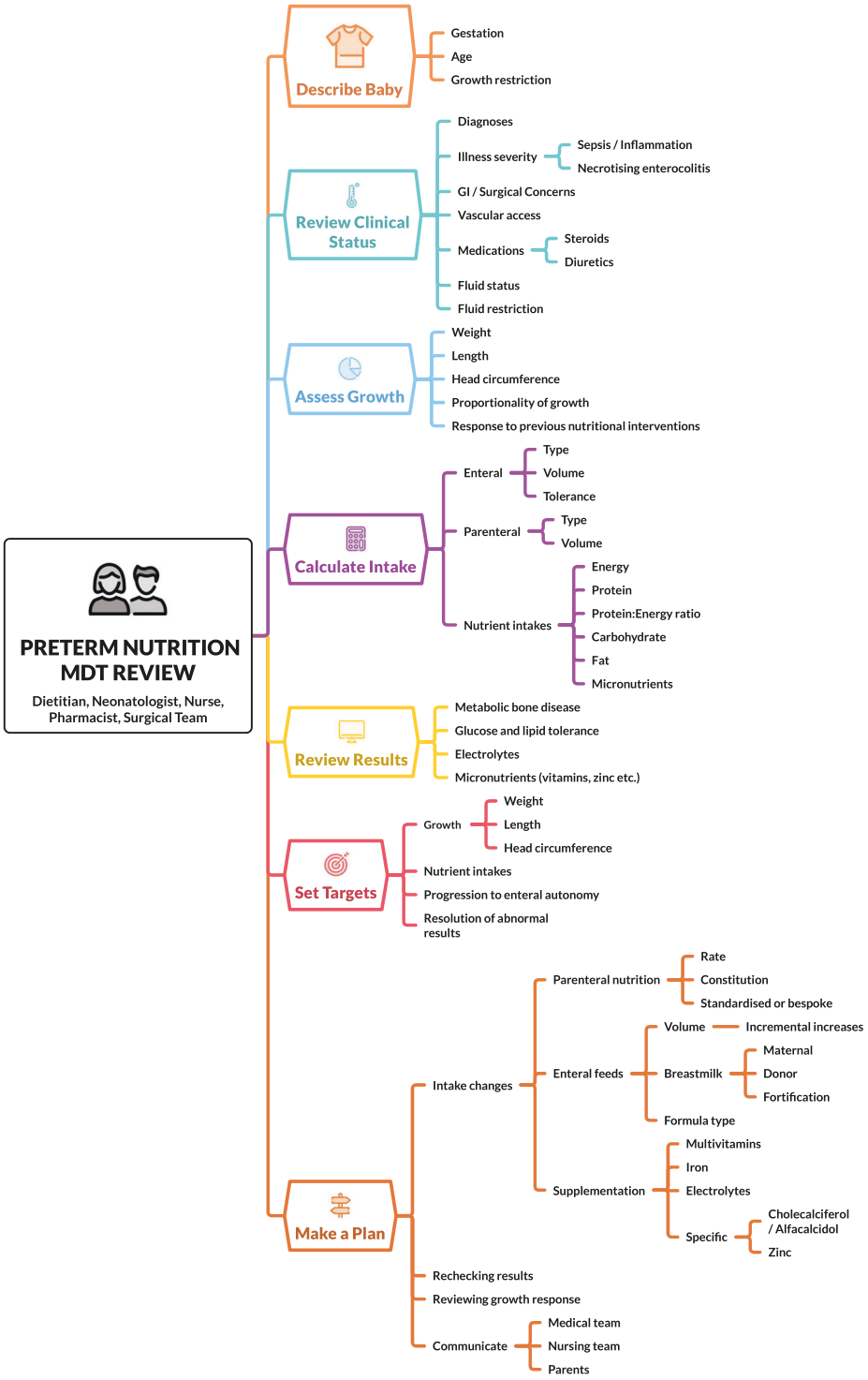
REFERENCES

1. Menon G, Davidson AL, Drake AJ, et al. Is preterm nutrition a trade-off between head and heart? *Arch Dis Child Fetal Neonatal Ed* 2019;104(3):F232-F34. doi: 10.1136/archdischild-2018-315672 [published Online First: 2018/10/14]
2. Leppanen M, Lapinleimu H, Lind A, et al. Antenatal and postnatal growth and 5-year cognitive outcome in very preterm infants. *Pediatrics* 2014;133(1):63-70. doi: 10.1542/peds.2013-1187 [published Online First: 2013/12/18]
3. Belfort MB, Rifas-Shiman SL, Sullivan T, et al. Infant growth before and after term: effects on neurodevelopment in preterm infants. *Pediatrics* 2011;128(4):e899-906. doi: 10.1542/peds.2011-0282 [published Online First: 2011/09/29]
4. Belfort MB, Martin CR, Smith VC, et al. Infant weight gain and school-age blood pressure and cognition in former preterm infants. *Pediatrics* 2010;125(6):e1419-26. doi: 10.1542/peds.2009-2746 [published Online First: 2010/05/19]
5. Embleton ND, Korada M, Wood CL, et al. Catch-up growth and metabolic outcomes in adolescents born preterm. *Arch Dis Child* 2016;101(11):1026-31. doi: 10.1136/archdischild-2015-310190 [published Online First: 2016/06/12]
6. Ni Y, Beckmann J, Hurst JR, et al. Size at birth, growth trajectory in early life, and cardiovascular and metabolic risks in early adulthood: EPICure study. *Arch Dis Child Fetal Neonatal Ed* 2021;106(2):149-55. doi: 10.1136/archdischild-2020-319328 [published Online First: 2020/08/17]
7. Agostoni C, Buonocore G, Carnielli VP, et al. Enteral nutrient supply for preterm infants: commentary from the European Society of Paediatric Gastroenterology, Hepatology and Nutrition Committee on Nutrition. *J Pediatr Gastroenterol Nutr* 2010;50(1):85-91. doi: 10.1097/MPG.0b013e3181adaee0 [published Online First: 2009/11/03]
8. American Academy of Pediatrics Committee on Nutrition. Nutritional Needs of the Preterm Infant. In: Kleinman RE, Greer FR, eds. Pediatric Nutrition. 8th edition. ed. Itasca, IL: American Academy of Pediatrics, 2019.
9. Cole TJ, Statnikov Y, Santhakumaran S, et al. Birth weight and longitudinal growth in infants born below 32 weeks' gestation: a UK population study. *Arch Dis Child Fetal Neonatal Ed* 2014;99(1):F34-40. doi: 10.1136/archdischild-2012-303536 [published Online First: 2013/08/13]
10. Andrews ET, Ashton JJ, Pearson F, et al. Early postnatal growth failure in preterm infants is not inevitable. *Arch Dis Child Fetal Neonatal Ed* 2019;104(3):F235-F41. doi: 10.1136/archdischild-2018-315082 [published Online First: 2018/08/24]
11. Johnson MJ, Wootton SA, Leaf AA, et al. Preterm birth and body composition at term equivalent age: a systematic review and meta-analysis. *Pediatrics* 2012;130(3):e640-9. doi: 10.1542/peds.2011-3379 [published Online First: 2012/08/15]
12. Young A, Brown LK, Ennis S, et al. Total body water in full-term and preterm newborns: systematic review and meta-analysis. *Arch Dis Child Fetal Neonatal Ed* 2021 doi: 10.1136/archdischild-2020-321112 [published Online First: 2021/04/02]
13. Villar J, Giuliani F, Fenton TR, et al. INTERGROWTH-21st very preterm size at birth reference charts. *Lancet* 2016;387(10021):844-5. doi: 10.1016/S0140-6736(16)00384-6 [published Online First: 2016/02/24]
14. Villar J, Giuliani F, Bhutta ZA, et al. Postnatal growth standards for preterm infants: the Preterm Postnatal Follow-up Study of the INTERGROWTH-21st Project. *Lancet Global Health* 2015;3(11):E681-E91. doi: Doi 10.1016/S2214-109x(15)00163-1
15. Young A, Andrews ET, Ashton JJ, et al. Generating longitudinal growth charts from preterm infants fed to current recommendations. *Arch Dis Child Fetal Neonatal Ed* 2020 doi: 10.1136/archdischild-2019-318404 [published Online First: 2020/05/27]
16. Landau-Crangle E, Rochow N, Fenton TR, et al. Individualized Postnatal Growth Trajectories for Preterm Infants. *JPEN J Parenter Enteral Nutr* 2018;42(6):1084-92. doi: 10.1002/jpen.1138 [published Online First: 2018/02/09]
17. Koletzko B, Poindexter B, Uauy R. Recommended nutrient intake levels for stable, fully enterally fed very low birth weight infants. *World Rev Nutr Diet* 2014;110:297-9. doi: 10.1159/000360195 [published Online First: 2014/04/23]
18.
20. Ziegler EE, O'Donnell AM, Nelson SE, et al. Body composition of the reference fetus. *Growth* 1976;40(4):329-41. [published Online First: 1976/12/01]
21. Uthaya S, Liu X, Babalis D, et al. Nutritional Evaluation and Optimisation in Neonates: a randomized, double-blind controlled trial of amino acid regimen and intravenous lipid

composition in preterm parenteral nutrition. *Am J Clin Nutr* 2016;103(6):1443-52. doi: 10.3945/ajcn.115.125138 [published Online First: 2016/04/22]
22. Embleton ND, Van Den Akker CH. Early parenteral amino acid intakes in preterm babies: does NEON light the way? *Arch Dis Child Fetal Neonatal Ed* 2018;103(2):F92-F94. doi: 10.1136/archdischild-2017-313406 [published Online First: 2017/10/04]
23. Morgan C, McGowan P, Herwitker S, et al. Postnatal head growth in preterm infants: a randomized controlled parenteral nutrition study. *Pediatrics* 2014;133(1):e120-8. doi: 10.1542/peds.2013-2207 [published Online First: 2014/01/01]
24. Strommen K, Haag A, Moltu SJ, et al. Enhanced nutrient supply to very low birth weight infants is associated with higher blood amino acid concentrations and improved growth. *Clin Nutr ESPEN* 2017;18:16-22. doi: 10.1016/j.clnesp.2017.01.003 [published Online First: 2017/11/15]
25. Roelants JA, Vlaardingerbroek H, van den Akker CHP, et al. Two-Year Follow-up of a Randomized Controlled Nutrition Intervention Trial in Very Low-Birth-Weight Infants. *JPEN J Parenter Enteral Nutr* 2018;42(1):122-31. doi: 10.1177/0148607116678196 [published Online First: 2016/11/23]
26. Koletzko B, Goulet O, Hunt J, et al. 1. Guidelines on Paediatric Parenteral Nutrition of the European Society of Paediatric Gastroenterology, Hepatology and Nutrition (ESPGHAN) and the European Society for Clinical Nutrition and Metabolism (ESPEN), Supported by the European Society of Paediatric Research (ESPR). *J Pediatr Gastroenterol Nutr* 2005;41 Suppl 2:S1-87. doi: 10.1097/01.mpg.0000181841.07090.f4
27. Mathur NB, Saini A, Mishra TK. Assessment of Adequacy of Supplementation of Vitamin D in Very Low Birth Weight Preterm Neonates: A Randomized Controlled Trial. *J Trop Pediatr* 2016;62(6):429-35. doi: 10.1093/tropej/fmv110 [published Online First: 2016/06/22]
28. Mactier H, McCulloch DL, Hamilton R, et al. Vitamin A supplementation improves retinal function in infants at risk of retinopathy of prematurity. *J Pediatr* 2012;160(6):954-9 e1. doi: 10.1016/j.jpeds.2011.12.013 [published Online First: 2012/01/31]
29. Bernhard W, Bockmann K, Maas C, et al. Combined choline and DHA supplementation: a randomized controlled trial. *Eur J Nutr* 2020;59(2):729-39. doi: 10.1007/s00394-019-01940-7 [published Online First: 2019/03/13]
30. Moon K, Rao SC, Schulzke SM, et al. Longchain polyunsaturated fatty acid supplementation in preterm infants. *Cochrane Database Syst Rev* 2016;12:CD000375. doi: 10.1002/14651858.CD000375.pub5 [published Online First: 2016/12/21]
31. Staub E, Evers K, Askie LM. Enteral zinc supplementation for prevention of morbidity and mortality in preterm neonates. *Cochrane Database Syst Rev* 2021;3:CD012797. doi: 10.1002/14651858.CD012797.pub2 [published Online First: 2021/03/13]
32. Vazquez-Gomis R, Bosch-Gimenez V, Juste-Ruiz M, et al. Zinc concentration in preterm newborns at term age, a prospective observational study. *BMJ Paediatr Open* 2019;3(1):e000527. doi: 10.1136/bmjpo-2019-000527 [published Online First: 2019/10/28]
33. Weeks CL, Marino LV, Johnson MJ. A systematic review of the definitions and prevalence of feeding intolerance in preterm infants. *Clin Nutr* 2021;40(11):5576-86. doi: 10.1016/j.clnu.2021.09.010 [published Online First: 20210917]
34. Dorling J, Abbott J, Berrington J, et al. Controlled Trial of Two Incremental Milk-Feeding Rates in Preterm Infants. *N Engl J Med* 2019;381(15):1434-43. doi: 10.1056/NEJMoa1816654 [published Online First: 2019/10/10]
35. Tahir W, Monahan M, Dorling J, et al. Economic evaluation alongside the Speed of Increasing milk Feeds Trial (SIFT). *Arch Dis Child Fetal Neonatal Ed* 2020;105(6):587-92. doi: 10.1136/archdischild-2019-318346 [published Online First: 2020/04/04]
36. Mihatsch WA, Franz AR, Lindner W, et al. Meconium passage in extremely low birthweight infants and its relation to very early enteral nutrition. *Acta Paediatr* 2001;90(4):409-11.
37. Deshmukh M, Balasubramanian H, Patole S. Meconium Evacuation for Facilitating Feed Tolerance in Preterm Neonates: A Systematic Review and Meta-Analysis. *Neonatology* 2016;110(1):55-65. doi: 10.1159/000444075 [published Online First: 20160407]
38. Unit NCT. Fluids exclusively enteral from day one in premature infants: Nottingham Clinical Trials Unit, 2019.
39. van der Lugt NM, Smits-Wintjens VE, van Zwieten PH, et al. Short and long term outcome of neonatal hyperglycemia in very preterm infants: a retrospective follow-up study. *BMC Pediatr* 2010;10:52. doi: 10.1186/1471-2431-10-52 [published Online First: 2010/07/22]
40. Beardsall K, Thomson L, Guy C, et al. Real-time continuous glucose monitoring in preterm infants (REACT): an international, open-label, randomised controlled trial. *Lancet Child Adolesc Health* 2021 doi: 10.1016/S2352-4642(20)30367-9 [published Online First: 2021/02/13]

41. Giretti I, D'Ascenzo R, Correani A, et al. Hypertriglyceridemia and lipid tolerance in preterm infants with a birth weight of less than 1250 g on routine parenteral nutrition. *Clin Nutr* 2021 doi: 10.1016/j.clnu.2020.12.039 [published Online First: 2021/02/03]
42. Alburaki W, Yusuf K, Dobry J, et al. High Early Parenteral Lipid in Very Preterm Infants: A Randomized-Controlled Trial. *J Pediatr* 2021;228:16-23 e1. doi: 10.1016/j.jpeds.2020.08.024 [published Online First: 2020/08/18]
43. Wang HJ, Hsieh YT, Liu LY, et al. Use of sodium glycerophosphate in neonatal parenteral nutrition solutions to increase calcium and phosphate compatibility for preterm infants. *Pediatr Neonatol* 2020;61(3):331-37. doi: 10.1016/j.pedneo.2020.02.004 [published Online First: 2020/03/23]
44. Vermont Oxford Network. Vermont Oxford Network Annual Report 2019. Burlington, VT, USA, 2019:5.
45. Ramel SE, Brown LD, Georgieff MK. The Impact of Neonatal Illness on Nutritional Requirements-One Size Does Not Fit All. *Curr Pediatr Rep* 2014;2(4):248-54. doi: 10.1007/s40124-014-0059-3 [published Online First: 2015/02/28]
46. Casaer MP, Mesotten D, Hermans G, et al. Early versus late parenteral nutrition in critically ill adults. *N Engl J Med* 2011;365(6):506-17. doi: 10.1056/NEJMoa1102662 [published Online First: 2011/07/01]
47. Fivez T, Kerklaan D, Mesotten D, et al. Early versus Late Parenteral Nutrition in Critically Ill Children. *N Engl J Med* 2016;374(12):1111-22. doi: 10.1056/NEJMoa1514762 [published Online First: 2016/03/16]
48. Excellence NIfHaC. Neonatal parenteral nutrition. London, UK, 2020.
49. Kuschel CA, Harding JE. Multicomponent fortified human milk for promoting growth in preterm infants. *Cochrane Database Syst Rev* 2004(1):CD000343. doi: 10.1002/14651858.CD000343.pub2 [published Online First: 2004/02/20]
50. Embleton ND, King C, Jarvis C, et al. Effectiveness of human milk-based fortifiers for preventing necrotizing enterocolitis in preterm infants: case not proven. *Breastfeed Med* 2013;8(4):421. doi: 10.1089/bfm.2013.0049 [published Online First: 2013/06/12]
51. Grace E, Hilditch C, Gomersall J, et al. Safety and efficacy of human milk-based fortifier in enterally fed preterm and/or low birthweight infants: a systematic review and meta-analysis. *Arch Dis Child Fetal Neonatal Ed* 2021;106(2):137-42. doi: 10.1136/archdischild-2020-319406 [published Online First: 2020/09/19]
52. Rochow N, Fusch G, Ali A, et al. Individualized target fortification of breast milk with protein, carbohydrates, and fat for preterm infants: A double-blind randomized controlled trial. *Clin Nutr* 2021;40(1):54-63. doi: 10.1016/j.clnu.2020.04.031 [published Online First: 20200506]
53. Arslanoglu S, Moro GE, Ziegler EE. Adjustable fortification of human milk fed to preterm infants: does it make a difference? *J Perinatol* 2006;26(10):614-21. doi: 10.1038/sj.jp.7211571 [published Online First: 2006/08/04]
54. Brion LP, Rosenfeld CR, Heyne R, et al. Optimizing individual nutrition in preterm very low birth weight infants: double-blinded randomized controlled trial. *J Perinatol* 2020;40(4):655-65. doi: 10.1038/s41372-020-0609-1 [published Online First: 2020/02/20]
55. Fabrizio V, Trzaski JM, Brownell EA, et al. Individualized versus standard diet fortification for growth and development in preterm infants receiving human milk. *Cochrane Database Syst Rev* 2020;11:CD013465. doi: 10.1002/14651858.CD013465.pub2 [published Online First: 2020/11/24]
56. Quigley M, Embleton ND, McGuire W. Formula versus donor breast milk for feeding preterm or low birth weight infants. *Cochrane Database of Systematic Reviews* 2019(7) doi: 10.1002/14651858.CD002971.pub5
57. El-Kafrawy U, Taylor RJ. Optimizing care of ventilated infants by improving weighing accuracy on incubator scales. *J Neonatal Perinatal Med* 2016;9(4):377-83. doi: 10.3233/NPM-161623 [published Online First: 2016/12/24]
58. Johnson MJ, Leaf AA, Pearson F, et al. Successfully implementing and embedding guidelines to improve the nutrition and growth of preterm infants in neonatal intensive care: a prospective interventional study. *BMJ Open* 2017;7(12):e017727. doi: 10.1136/bmjopen-2017-017727 [published Online First: 2017/12/09]
59. Fenton TR, Kim JH. A systematic review and meta-analysis to revise the Fenton growth chart for preterm infants. *BMC Pediatr* 2013;13:59. doi: 10.1186/1471-2431-13-59 [published Online First: 2013/04/23]
60. RCPCH. UK-WHO growth charts - neonatal and infant close monitoring (NICM) London, UK: Royal College of Paediatrics and Child Health; 2019 [Available from: .


Weight (kg)
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
27
29
31
33
35
37
39
41
43
Postmenstrual Age (weeks)
INTERGROWTH (postnatal)
Fenton
INTERGROWTH (birthweight)
UK−WHO
Southampton
Individualised


PRETERM NUTRITION MDT REVIEW
Dietitian, Neonatologist, Nurse, Pharmacist, Surgical Team
Describe Baby
Gestation
Age
Growth restriction
Review Clinical Status
Diagnoses
Illness severity
Sepsis / Inflammation
Necrotising enterocolitis
GI / Surgical Concerns
Vascular access
Medications
Steroids
Diuretics
Fluid status
Fluid restriction
Assess Growth
Weight
Length
Head circumference
Proportionality of growth
Response to previous nutritional interventions
Calculate Intake
Enteral
Type
Volume
Tolerance
Parenteral
Type
Volume
Nutrient intakes
Energy
Protein
Protein:Energy ratio
Carbohydrate
Fat
Micronutrients
Review Results
Metabolic bone disease
Glucose and lipid tolerance
Electrolytes
Micronutrients (vitamins, zinc etc.)
Set Targets
Growth
Weight
Length
Head circumference
Nutrient intakes
Progression to enteral autonomy
Resolution of abnormal results
Make a Plan
Intake changes
Parenteral nutrition
Rate
Constitution
Standardised or bespoke
Enteral feeds
Volume
Incremental increases
Breastmilk
Maternal
Donor
Fortification
Formula type
Supplementation
Multivitamins
Iron
Electrolytes
Specific
Cholecalciferol / Alfacalcidol
Zinc
Rechecking results
Reviewing growth response
Communicate
Medical team
Nursing team
Parents

Appendix 4

Published paper: Young A, Cole TJ, Cheng G, Ennis S, Beattie RM, Johnson MJ. Changes in the growth of very preterm infants in England 2006-2018. Arch Dis Child Fetal Neonatal Ed. 2022

Changes in the Growth of Very Preterm Infants in England 2006-2018

Aneurin Young [1,2][ORCiD: 0000-0003-3549-3813], Tim J Cole[3][0000-0001-5711-8200], Guo Cheng[4], Sarah Ennis[4], R Mark Beattie[2,5][0000-0003-4721-0577], Mark J Johnson[1,2][0000-0003-1829-9912].

1. Department of Neonatal Medicine, University Hospital Southampton NHS Foundation Trust, Southampton, UK
2. NIHR Southampton Biomedical Research Centre, University of Southampton and University Hospital Southampton NHS Foundation Trust, Southampton, UK
3. UCL Great Ormond Street Institute of Child Health, London, UK
4. Human Genetics and Genomic Medicine, University of Southampton, Southampton, UK
5. Department of Paediatric Gastroenterology, University Hospital Southampton NHS Foundation Trust, Southampton, UK

Corresponding author:

Aneurin Young
Neonatal Unit,
University Hospital Southampton NHS Foundation Trust,
Tremona Road,
Southampton,
SO16 6YD.
aneurin.young@nhs.net

Word count: 2,498

This article has been accepted for publication in Archives of Disease in Childhood: Fetal & Neonatal following peer review, and the Version of Record can be accessed online at http://dx.doi.org/10.1136/archdischild-2022-324584.

ABSTRACT

Objective

To compare weight gain from birth to term equivalent age in very preterm infants in England born during two eras (2006-2011 and 2014-2018). To assess demographic and care factors influencing weight gain.

Methods

Data for infants born before 32 weeks of gestation during 2014-2018 in England were obtained (29,687 infants). Weight gain modelled using SuperImposition by Translation And Rotation (SITAR), with infants grouped by gestational week. A cohort from 2006-2011 was used for comparison (3,288 infants). Multiple linear regression was used to assess factors influencing change in weight SD score from birth to 36 weeks postmenstrual age (PMA).

Results

Weight gain velocity (termed "intensity" in SITAR models) was greater in the more recent cohort for all gestation groups born before 30 weeks of gestation. After adjustment for gestation, birthweight and other perinatal factors, care elements associated with faster weight gain included delivery in a Level 3 unit (0.09 SD less weight gain deficit, 95%CI 0.07 to 0.10) and parenteral nutrition initiation during the first day of life (0.08 SD, 95%CI 0.06 to 0.10). Factors associated with slower weight gain included early ventilation (-0.07 SD, 95%CI: -0.08 to -0.05) and less deprived neighbourhood (-0.012 SD per index of multiple deprivation decile, 95%CI: -0.015 to -0.009).

Conclusions

Weight gain for extremely preterm infants was faster during 2014-2018 than during 2006-2011. Early initiation of parenteral nutrition and birth in a level 3 unit may contribute to faster weight gain.

WHAT IS ALREADY KNOWN?

- Very preterm infants in England during 2006-2011 demonstrated a pattern of weight crossing centiles downwards.
- Poor growth during the neonatal period is associated with adverse neurodevelopmental outcomes.

WHAT THIS STUDY ADDS

- Extremely preterm infants in England gained weight faster during the period 2014-2018 compared with 2006-2011, but a fall in weight SD score between birth and 36 weeks postmenstrual age remained.
- This fall in weight SD score was smaller for infants born in a Level 3 neonatal unit and those who received early parenteral nutrition, and was greater for infants from less deprived neighbourhoods.

HOW THIS STUDY MIGHT AFFECT RESEARCH AND PRACTICE

- This study identifies care factors associated with faster weight gain in very preterm infants. Further research is required to assess their effects on important outcomes such as neurodevelopment.

INTRODUCTION

Analysis of weight trajectories of very and extremely preterm infants managed in English neonatal units during the period 2006-2011 demonstrated that downward crossing of weight centiles was common.[1] Mean growth curves of the infants fell by more than two centile spaces between birth and term equivalent age (corresponding to a fall in standard deviation score of around 1.5). This work used the SuperImposition by Translation And Rotation (SITAR) growth curve model to summarise the growth of groups of infants into a single summary curve.[2]

The extent to which this pattern of growth is physiological or desirable remains contested. Early water loss during postnatal adaptation may cause weight loss followed by growth along a lower trajectory. However, the appropriate degree of early weight loss at different gestations is difficult to define given that all very preterm infants are subject to active fluid management in neonatal units. Previous work using the SITAR method with infants born in England demonstrated a lack of early weight loss in the most preterm infants[1], whereas early weight loss was identified at all gestations in a study of a cohort of German and Canadian infants.[3]

An increasing body of evidence has demonstrated an association between faster growth during the neonatal period and improved later neurodevelopmental outcomes for preterm infants.[4-6] Work in term-born infants has identified that low birthweight followed by rapid weight gain in infancy is associated with adverse cardiometabolic outcomes.[7] Recent findings have suggested that growth during the first few weeks of life in preterm infants has little impact on metabolic health.[8,9] Taken together, these data suggest that increasing growth in the early postnatal period may improve neurodevelopmental outcomes without adversely impacting cardiometabolic health.

Since 2011 there have been changes in neonatal practice, prioritising early nutrition including the early introduction of parenteral nutrition and ongoing improved nutritional provision to preterm infants, with a

recent study by our group demonstrating that early postnatal growth of infants born very preterm can be improved and brought more into line with in-utero growth.[10]

This paper has two aims: to assess changes in growth patterns of preterm infants between the 2006-2011 cohort and a later cohort born during 2014-2018; and to assess the impact of demographic features at birth, complications of prematurity and level of neonatal care unit on the growth of preterm infants.

PATIENTS AND METHODS

Data were obtained from the National Neonatal Research Database (NNRD).[11] Weight was selected as a marker of growth as it had the most complete data entry in the database.

Data for all infants born before 32 weeks of gestation and cared for during the period from 2014 to 2018 in England were obtained from the NNRD. Data included gestation, birthweight, sex, multiplicity of pregnancy, level of neonatal unit at the place of birth, Apgar score, lower super output area (LSOA) of the mother's address, serial weight measurements, and daily ventilation and nutrition type information. The LSOA is a geographical area which is associated with an index of multiple deprivation (IMD) score in UK government publications,[12] with a higher IMD score denoting lower levels of deprivation in that neighbourhood. Weight and sex information from the published 2006-2011 cohort were used.[1,12]

Visual examination of weights revealed a handful of extreme outliers above 60,000g or below 60g which were excluded. Bands of weights far above and far below the expected range were identified as transposition errors caused by incorrect decimal places and were shifted accordingly. The data cleaning protocol is given as Supplementary File A. Upper age limits were set so that at least 10% of infants for each completed week of gestation remained in the dataset at that age, to maintain adequate data for modelling. This approach is consistent with previous reports.[1]

This project was approved by the Oxford A research ethics committee (20/SC/0073).

SITAR Analysis

Weight gain was summarised using the SITAR growth curve model with the `sitar` package in R.[13] Cleaned data from the 2006-2011 cohort were also re-analysed to compare the results for the two eras.

The SITAR model fits a shape-invariant mean growth curve which is assumed to underlie each infant's curve. Individual curves then deviate from the mean curve in terms of three random effects: *size* reflects the overall weight of the infant, *timing* is the age when the infant has the fastest weight gain velocity, and *intensity* is the overall rate of weight gain. The era of the study (i.e. 2006-2011 versus 2014-2018) was included as a fixed effect in combined models using all infants, to test its effect on size, timing and intensity.

SITAR curves were fitted for male and female infants combined as their growth patterns are known to be similar.[1,14] Separate growth curves were formed for each week of gestational age at birth, with 22 and 23 weeks of gestation combined due to small numbers. This matched the approach of the earlier cohort. Weight measurements more than four residual standard deviations from these mean growth curves were then excluded. For each individual growth curve, weight measurements were examined as triplets, velocity of weight gain was calculated for each measurement and those with an implausible velocity were excluded. Final SITAR models were formed from the resultant data.

Reanalysed UK1990 growth centiles[15,16] were illustrated using reference data within the `childsds` package for R, with the centiles for males and females averaged.[17]

Regression Analysis

Weight SD scores were calculated using the reanalysed UK1990 data, which excluded infants born prior to 23 weeks.[15] Weight SD scores were calculated at birth, at seven days of life and at 36 weeks PMA. Weight change was defined as the difference in weight SD scores between two time points.

Multiple linear regression models were used to assess demographic and clinical factors which had been identified by clinical researchers as likely to influence growth. The Bayesian information criterion (BIC) was used to assess model efficiency and to select the most appropriate variables for inclusion. For each model

variable the *t*-statistic was calculated by dividing the regression coefficient by its standard error; *t*-statistics exceeding 2 in absolute value indicate statistical significance at $P < 0.05$.[18]

RESULTS

Weight data were available for 29,687 infants born 2014-2018 and 3,288 infants born 2006-2011 (161 units contributed data during 2014-2018; 40 were included in the database in 2006-2011).

SITAR Growth Curve Analysis

Figure 1A shows SITAR curves for infants born 2014-2018 grouped by completed weeks of gestation and superimposed on UK-WHO weight centiles with the sexes averaged. Figure 1B shows the same data with weight on a logarithmic scale, so that the slope of each curve represents relative weight gain (*i.e*. g/kg body weight/day).[19] Each curve is close to median weight for gestation at birth but then crosses centiles downwards.

Comparison with Data from 2006-2011

Figure 2 compares SITAR curves from the two eras, 2006-2011 and 2014-2018. For infants born from 24 to 28 weeks of gestation, weight at birth was similar. Birthweight was lower in the more recent era for infants born at 22-23 weeks, likely reflecting a trend towards initiating intensive care more commonly in the smallest and most premature infants.

For each gestation group before 28 weeks, SITAR curves for the two eras diverged slightly, with infants in the more recent era exhibiting faster weight gain signified by a steeper SITAR curve. This effect was not so apparent for infants born after 28 weeks PMA. For the more mature infants, the SITAR curve for the 2006-2011 infants demonstrated early weight loss, followed by increasing weight. This pattern was much less prominent for the 2014-2018 cohort, with very little early weight loss seen in these infants. Instead, there was a period of static weight followed by acceleration in growth velocity.

Table 1 shows the effect of being in the 2014-2018 cohort on the size (in grams), timing (in weeks) and intensity (in percentage change) fixed effects of the SITAR models for each gestation group, with the associated *t*-statistics. The size value was generally higher (although not always significantly so). This indicates that mean weight was greater during the 2014-2018 cohort than the 2006-2011 cohort (see Figure 2). The timing parameter reflects the age at peak weight gain, and the negative effect for the later cohort indicates peak weight gain occurring earlier (although not significantly so for most groups). Intensity was significantly greater for every gestation group born before 30 weeks of gestation, indicating more rapid weight gain in the later cohort.

Table 1. The effect of birth in the 2014-2018 era (compared to the 2006-2011 era) on mean size, timing and intensity of SITAR models, grouped by gestation.

Gestation group (weeks)	Effect of later era on size, grams (95% CI)	*t*-statistic	Effect of later era on timing, weeks (95% CI)	*t*-statistic	Effect of later era on intensity, % (95% CI)	*t*-statistic
22-23	20 (-35 to 74)	0.7	-0.4 (-1 to 0.3)	-1.2	12 (5 to 19)	3.3
24	62 (31 to 93)	3.9	-0.2 (-0.6 to 0.1)	-1.3	15 (11 to 19)	7
25	11 (-17 to 40)	0.8	-0.4 (-0.7 to -0.2)	-3.2	5 (1 to 8)	2.5
26	61 (38 to 84)	5.1	-0.1 (-0.3 to 0.2)	-0.5	13 (10 to 16)	8
27	53 (33 to 73)	5.1	0.0 (-0.2 to 0.1)	-0.3	9 (7 to 12)	6.7
28	50 (32 to 68)	5.3	0.1 (-0.1 to 0.2)	1.1	6 (4 to 9)	4.8
29	24 (6 to 42)	2.6	0.1 (0 to 0.1)	2.1	5 (3 to 7)	4.3
30	11 (-7 to 29)	1.2	-0.3 (-0.6 to -0.1)	-2.8	-2 (-3 to 0)	-2.2
31	12 (-6 to 30)	1.3	-1 (-1.3 to -0.8)	-7.2	-3 (-4 to -2)	-7.4

Early Weight Loss in the 2014-2018 Cohort

Infants born at 22-23 weeks of gestation had the smallest drop in weight (mean loss 1.8%) and those at 31 weeks had the largest (mean loss 4.3%). Over the whole cohort, 49% of infants had no early weight loss.

Influences on Weight Gain in the 2014-2018 Cohort

For the analysis of influences on weight gain in the 2014-2018 cohort, infants were included who had weight measured at birth, during the second week of life and at 36 weeks PMA. Demographic details of included infants are given in Table 2. As expected, there were more infants born at later gestations (Supplementary Figure 1). The mean change in weight SD score from birth to 36 weeks PMA was -0.94. This consisted of a

drop of 0.66 SD scores during the first week of life followed by a drop of 0.28 SD scores between the second week of life and 36 weeks PMA.

Table 2. Summary details of included infants

Number of infants		27505
Gestational age (weeks+days) (median (IQR))		29^{+2} (27^{+3} to 30^{+6})
Birthweight (g) (median (IQR))		1156 (895 to 1430)
Birthweight z-score (mean (SD))		-0.45 (0.90)
Small for gestational age (% (*n*))		18% (4826)
Male (% (*n*))		54% (14914)
Singleton (% (*n*))		73% (19968)
Neonatal Unit Level at Birth Centre	Level 1 (% (*n*))	9% (2423)
	Level 2 (% (*n*))	38% (10571)
	Level 3 (% (*n*))	53% (14511)
Change in Weight SD score from birth to 36 weeks PMA (mean (SD))		-0.94 (0.66)
Change in Weight SD score from birth to the second week of life (mean (SD))		-0.66 (0.39)
Change in Weight SD score from the second week of life to 36 weeks PMA (mean (SD))		-0.28 (0.62)

Table 3 summarises the multiple linear regression of weight change in the 2014-2018 cohort on selected perinatal factors. The regression coefficient column shows the effect of a one unit change in each factor on the change in weight SD score between birth and 36 weeks PMA.

Infants with a higher Apgar score exhibited faster weight gain and those who required ventilation during the first day of life had slower weight gain, suggesting that sicker infants had slower weight gain. Birth in a level 3 unit was associated with an increase in weight SD change of 0.09 (95%CI 0.07 to 0.10), i.e. infants born in a level 3 unit gained weight faster, with a mean 0.09 SD score smaller weight gain deficit, than infants born in a level 1 or level 2 unit. Initiation of parenteral nutrition during the first day of life was also associated with greater weight gain (0.08 SD, 95%CI 0.06 to 0.10). Birth to a mother residing in a less deprived neighbourhood was associated with slower weight gain. These effects were not materially altered when weight change was calculated from the second week of life to 36 weeks PMA (Supplementary Table 1).

However the effect of birthweight was different depending on whether weight gain was calculated from birth or from the second week of life. When birth was taken as the starting point, gestational age was positively associated with weight change, and birthweight negatively, indicating that weight change was generally more positive for later born and smaller-for-dates infants. There was a positive interaction between birthweight and gestational age at birth, indicating that the effect of birthweight (i.e. faster growth for smaller-for-dates infants) was greater for more preterm infants. When weight during the second week of life was taken as the starting point, birthweight was positively associated with weight change and the interaction term was negative. This reflects greater early weight loss in larger-for-dates infants.

Table 3. Results of multiple linear regression of change in weight SD score from birth to 36 weeks postmenstrual age on factors around the time of birth

	Regression coefficient – weight change (95% CI)	***t*-statistic**
Gestational age (weeks)	0.14 (0.13 to 0.15)	23
Birthweight (kg)	-1.5 (-1.8 to -1.1)	-8
Interaction term for gestational age (weeks) and birthweight (kg)	0.02 (0.006 to 0.029)	3
Sex (female)	-0.12 (-0.14 to -0.10)	-15
IMD decile*	-0.012 (-0.015 to -0.009)	-8
Born in level 3 unit (vs born in level 1 or level 2 unit)	0.09 (0.07 to 0.10)	10
Apgar score at 5 minutes	0.005 (0.002 to 0.009)	2.1
Ventilated on day of birth (vs not)	-0.07 (-0.08 to -0.05)	-7
Parenteral nutrition during first day of life (vs not)	0.08 (0.06 to 0.10)	9
Birth year (effect of birth one year later)	0.035 (0.029 to 0.040)	12

*Index of multiple deprivation decile for mother's postcode (higher is less deprived)

Supplementary Table 1. Results of multiple linear regression of change in weight SD score from the second week of life to 36 weeks postmenstrual age on factors around the time of birth

	Regression coefficient – weight change (95% CI)	***t*-statistic**
Gestational age (weeks)	0.09 (0.08 to 0.10)	15
Birthweight (kg)	1.0 (0.7 to 1.4)	6
Interaction term for gestational age (weeks) and birthweight (kg)	-0.03 (-0.05 to -0.02)	-6
Sex (female)	-0.04 (-0.06 to -0.03)	-5
IMD decile*	-0.013 (-0.016 to -0.010)	-9
Born in level 3 unit (vs born in level 1 or level 2 unit)	0.08 (0.06 to 0.09)	10
Apgar score at 5 minutes	0.008 (0.004 to 0.013)	4
Ventilated on day of birth (vs not)	-0.08 (-0.10 to -0.06)	-9
Parenteral nutrition during first day of life (vs not)	0.04 (0.02 to 0.06)	4
Birth year (effect of birth one year later)	0.022 (0.016 to 0.027)	8

*Index of multiple deprivation decile for mother's postcode (higher is less deprived)

DISCUSSION

Changes in Weight Gain Pattern

This paper describes the weight gain pattern of very preterm infants in England born during 2014-2018. In comparison to infants born during 2006-2011, the most preterm infants (born before 30 weeks of gestation) exhibited faster weight gain. Early weight loss, which was apparent for later gestations in 2006-2011, was much less pronounced in 2014-2018, being replaced by a short period of static weight. The findings of this study are strengthened by the large number of infants included. Comparisons between the eras are limited by the inclusion of far more neonatal units in the more recent cohort. Weight data were gathered during clinical care and were not subject to standardisation or error checking. There were insufficient data to analyse linear growth or head circumference growth.

The findings presented here are in good agreement with a recent paper using different methods to examine growth in an overlapping cohort.[23] That group found that early postnatal weight loss had decreased over time (as seen in the more mature groups in this analysis) and that subsequent weight gain was faster. Similarly to the data presented here, they identified that weight gain continues to be less than the rate needed to keep pace with the equivalent fetus in utero.

The optimal pattern of weight change during the first two weeks of life remains disputed. In term-born infants, there is a well-recognised pattern of early weight loss, caused by the term contraction of extracellular spaces (TeCES).[24] The extent to which there is a preterm correlate to this effect is unclear. Data from a cohort of preterm infants with minimal comorbidities found greater percentage weight loss in the most preterm infants than in more mature infants[3], with an average early weight loss of 11% in infants born before 30 weeks of gestation and 7% in those born at 30-34 weeks of gestation. Conversely, we found less weight loss in the most immature infants, at only 4%. The discrepancy between these results may be due to differing fluid management and nutritional strategies. Alternatively, the low mean early weight loss seen in these English data may be due to fluid overload in critically unwell infants who would have been excluded from the selective cohort of preterm infants with few complications. SITAR curves provide an objective summary of weight change over a period of time, although the smoothing process may reduce the impact of rapid weight changes, especially early weight loss.

Influences on Growth

Multiple linear regression identified demographic and care factors which were associated with changes in growth from birth to 36 weeks PMA (Table 3). The mean change in weight SD score from birth to 36 weeks PMA was -0.94, consisting of a fall in weight SD score of around 0.66 during the first week of life, followed by a more gradual drop of 0.28 between the second week of life at 36 weeks PMA (as reflected in the SITAR charts). Change in weight SD score was positively associated with later gestation, corroborating the reduced deviation from birthweight centile seen in SITAR models (Figure 2). There was also a positive association with year of birth, confirming that growth has increased over time. Birth in a hospital offering level 3 neonatal care was associated with faster weight gain, suggesting an impact of antenatal transfer to a specialised setting.[20] Early parenteral nutrition was also associated with a reduced weight gain deficit.

CONCLUSIONS

Despite modest increases in weight change in the most preterm infants, very preterm infants in England continue to exhibit a pattern of weight gain which falls short of the equivalent fetal growth pattern. Birth in a level 3 unit was associated with faster weight gain, as was early initiation of parenteral nutrition, whilst a marker of lower deprivation was associated with slower weight gain.

FIGURE LEGENDS

Figure 1. SITAR curves for weight by completed weeks of gestational age for the 2014-18 cohort superimposed on reanalysed *UK1990* centile lines (sexes averaged) plotted with **A.** an absolute weight scale; and **B.** a logarithmic scale of weight with a velocity fan demonstrating weight gain in g/kg/day.

Figure 2. SITAR curves by gestational age, comparing infants born during the era of 2006-2011 and the era of 2014-2018. Grey lines are the reanalysed *UK1990* growth chart 2nd, 50th and 98th centile lines for comparison.[15]

Supplementary Figure 1. Number of infants born at each week of completed gestation for included infants in the 2014-18 cohort.

Contributorship

Aneurin Young contributed to the conception and design of the work, analysed the data and drafted the manuscript. Tim Cole contributed to the conception and design of the work, supported data analysis and revised the manuscript. Guo Cheng contributed to the analysis of the data. Sarah Ennis, R Mark Beattie and Mark Johnson contributed to the conception of the work and revised the manuscript.

Funding

This study was funded by the NIHR Southampton Biomedical Research Centre. Aneurin Young is supported by a research fellowship issued by the NIHR Southampton Biomedical Research Centre (no grant number supplied).

Acknowledgements

The authors acknowledge the Neonatal Data Analysis Unit (Imperial College London) as the source of the National Neonatal Research Database (NNRD) data used in this publication, the neonatal units contributing data to the NNRD and the patients included on the database. The authors also acknowledge the support of the NIHR Southampton BRC Data Science unit and Dr Hang Phan. Data analysis was performed using the IRIDIS High Performance Computing Facility and associated services at the University of Southampton.

Competing Interests

Aneurin Young and Mark Johnson occupy posts supported by NIHR Southampton Biomedical Research Centre.

Data Sharing and Data Availability

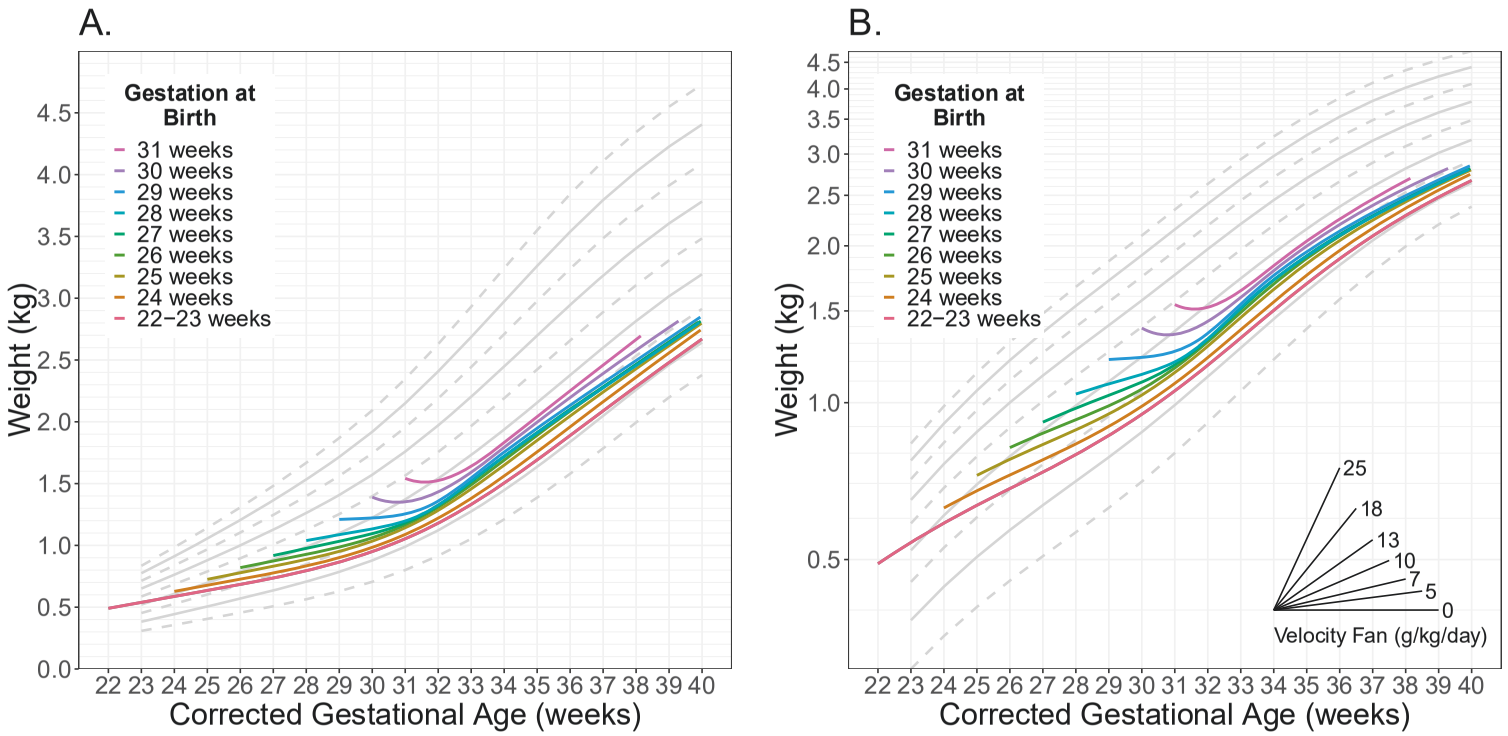
Data from the NNRD can be acquired for research upon application.

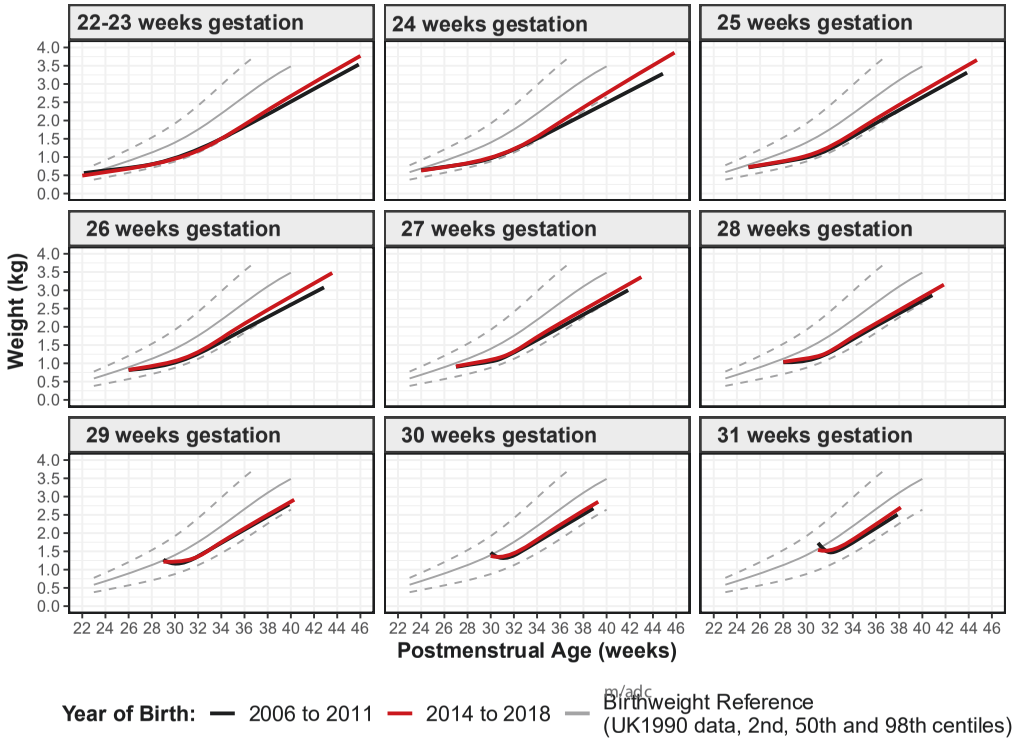
Ethics Approval Statement

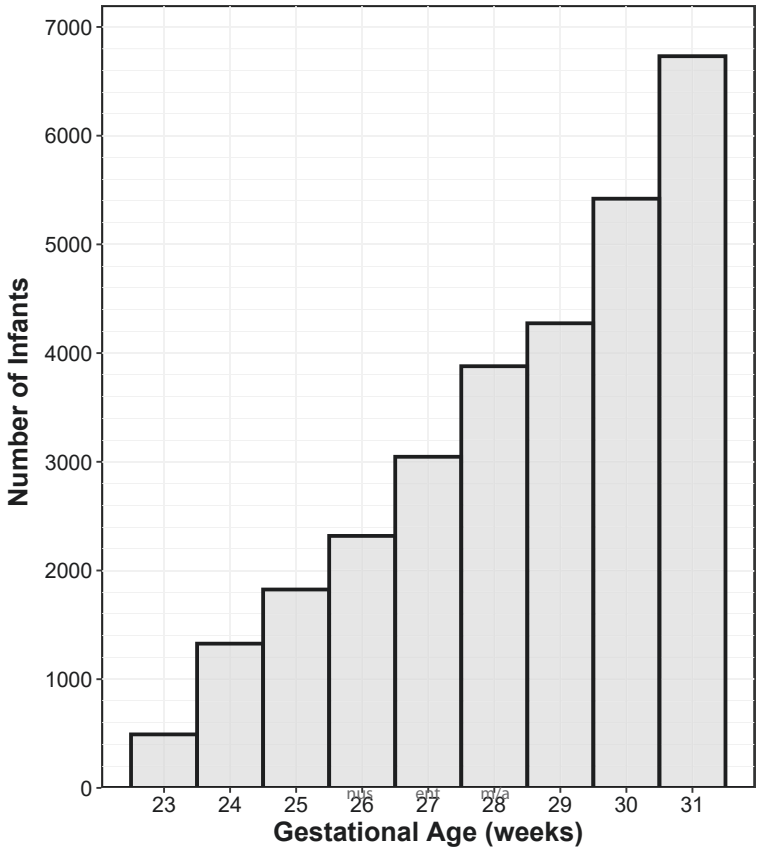
This study involves human participants and was approved by the NHS HRA Oxford A Research Ethics Committee (20/SC/0073).

REFERENCES

1. Cole TJ, Statnikov Y, Santhakumaran S, et al. Birth weight and longitudinal growth in infants born below 32 weeks' gestation: a UK population study. *Arch Dis Child Fetal Neonatal Ed* 2014;99(1):F34-40. doi: 10.1136/archdischild-2012-303536 [published Online First: 2013/08/13]
2. Cole TJ, Donaldson MD, Ben-Shlomo Y. SITAR--a useful instrument for growth curve analysis. *Int J Epidemiol* 2010;39(6):1558-66. doi: 10.1093/ije/dyq115 [published Online First: 2010/07/22]
3. Rochow N, Raja P, Liu K, et al. Physiological adjustment to postnatal growth trajectories in healthy preterm infants. *Pediatr Res* 2016;79(6):870-9. doi: 10.1038/pr.2016.15 [published Online First: 2016/02/10]
4. Belfort MB, Rifas-Shiman SL, Sullivan T, et al. Infant growth before and after term: effects on neurodevelopment in preterm infants. *Pediatrics* 2011;128(4):e899-906. doi: 10.1542/peds.2011-0282 [published Online First: 2011/09/29]
5. Ehrenkranz RA, Dusick AM, Vohr BR, et al. Growth in the neonatal intensive care unit influences neurodevelopmental and growth outcomes of extremely low birth weight infants. *Pediatrics* 2006;117(4):1253-61. doi: 10.1542/peds.2005-1368 [published Online First: 2006/04/06]
6. Cordova EG, Cherkerzian S, Bell K, et al. Association of Poor Postnatal Growth with Neurodevelopmental Impairment in Infancy and Childhood: Comparing the Fetus and the Healthy Preterm Infant References. *J Pediatr* 2020;225:37-43 e5. doi: 10.1016/j.jpeds.2020.05.063 [published Online First: 2020/06/12]
7. Vaag A. Low birth weight and early weight gain in the metabolic syndrome: consequences for infant nutrition. *Int J Gynaecol Obstet* 2009;104 Suppl 1:S32-4. doi: 10.1016/j.ijgo.2008.11.026 [published Online First: 2009/01/22]
8. Embleton ND, Korada M, Wood CL, et al. Catch-up growth and metabolic outcomes in adolescents born preterm. *Arch Dis Child* 2016;101(11):1026-31. doi: 10.1136/archdischild-2015-310190 [published Online First: 2016/06/12]
9. Ni Y, Beckmann J, Hurst JR, et al. Size at birth, growth trajectory in early life, and cardiovascular and metabolic risks in early adulthood: EPICure study. *Arch Dis Child Fetal Neonatal Ed* 2021;106(2):149-55. doi: 10.1136/archdischild-2020-319328 [published Online First: 2020/08/17]
10. Andrews ET, Ashton JJ, Pearson F, et al. Early postnatal growth failure in preterm infants is not inevitable. *Arch Dis Child Fetal Neonatal Ed* 2019;104(3):F235-F41. doi: 10.1136/archdischild-2018-315082 [published Online First: 2018/08/24]
11. Gale C, Morris I, Neonatal Data Analysis Unit Steering B. The UK National Neonatal Research Database: using neonatal data for research, quality improvement and more. *Arch Dis Child Educ Pract Ed* 2016;101(4):216-8. doi: 10.1136/archdischild-2015-309928 [published Online First: 20160311]
12. Ministry of Housing Communities and Local Government. Lower Super Output Area (LSOA) IMD 2019 (OSGB1936). In: Ministry of Housing Communities and Local Government, ed. London, UK, 2021.
13. sitar - Super Imposition by Translation and Rotation Growth Curve Analysis [program]. 1.2.0 version. London, UK, 2021.
14. Young A, Andrews ET, Ashton JJ, et al. Generating longitudinal growth charts from preterm infants fed to current recommendations. *Arch Dis Child Fetal Neonatal Ed* 2020 doi: 10.1136/archdischild-2019-318404 [published Online First: 2020/05/27]
15. Cole TJ, Williams AF, Wright CM, et al. Revised birth centiles for weight, length and head circumference in the UK-WHO growth charts. *Ann Hum Biol* 2011;38(1):7-11. doi: 10.3109/03014460.2011.544139 [published Online First: 2010/12/24]
16. RCPCH. UK-WHO growth charts - neonatal and infant close monitoring (NICM) London, UK: Royal College of Paediatrics and Child Health; 2019 [Available from: https://www.rcpch.ac.uk/resources/uk-who-growth-charts-neonatal-infant-close-monitoring-nicm.
17. childsds - Data and Methods Around Reference Values in Pediatrics [program]. 0.7.6 version, 2020.
18. Jiang J. Applied medical statistics. Hoboken, NJ: John Wiley & Sons, Inc.,, 2022:1 online resource.
19. Cole TJ, Altman DG. Statistics Notes: Percentage differences, symmetry, and natural logarithms. *BMJ* 2017;358:j3683. doi: 10.1136/bmj.j3683 [published Online First: 20170816]
20. Marlow N, Bennett C, Draper ES, et al. Perinatal outcomes for extremely preterm babies in relation to place of birth in England: the EPICure 2 study. *Arch Dis Child Fetal Neonatal Ed* 2014;99(3):F181-8. doi: 10.1136/archdischild-2013-305555 [published Online First: 2014/03/08]
21. Agostoni C, Buonocore G, Carnielli VP, et al. Enteral nutrient supply for preterm infants: commentary from the European Society of Paediatric Gastroenterology, Hepatology and Nutrition Committee on Nutrition. *J Pediatr Gastroenterol Nutr* 2010;50(1):85-91. doi: 10.1097/MPG.0b013e3181adaee0 [published Online First: 2009/11/03]
22. National Institute for Health and Care Excellence. Neonatal parenteral nutrition. London, UK, 2020.
23. Greenbury SF, Angelini ED, Ougham K, et al. Birthweight and patterns of postnatal weight gain in very and extremely preterm babies in England and Wales, 2008-19: a cohort study. *Lancet Child Adolesc Health* 2021;5(10):719-28. doi: 10.1016/S2352-4642(21)00232-7 [published Online First: 2021/08/28]
24. Landau-Crangle E, Rochow N, Fenton TR, et al. Individualized Postnatal Growth Trajectories for Preterm Infants. *JPEN J Parenter Enteral Nutr* 2018;42(6):1084-92. doi: 10.1002/jpen.1138 [published Online First: 2018/02/09]
25. Andrews ET, Beattie RM, Johnson MJ. Measuring body composition in the preterm infant: Evidence base and practicalities. *Clin Nutr* 2019 doi: 10.1016/j.clnu.2018.12.033 [published Online First: 2019/02/10]


A.
B.
Gestation at Birth
31 weeks
30 weeks
29 weeks
28 weeks
27 weeks
26 weeks
25 weeks
24 weeks
22–23 weeks
Weight (kg)
Corrected Gestational Age (weeks)
25
18
13
10
7
5
0
Velocity Fan (g/kg/day)


22-23 weeks gestation
24 weeks gestation
25 weeks gestation
26 weeks gestation
27 weeks gestation
28 weeks gestation
29 weeks gestation
30 weeks gestation
31 weeks gestation
Weight (kg)
0.0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0
22 24 26 28 30 32 34 36 38 40 42 44 46
Postmenstrual Age (weeks)
Year of Birth: 2006 to 2011 2014 to 2018 Birthweight Reference
(UK1990 data, 2nd, 50th and 98th centiles)


Number of Infants
7000
6000
5000
4000
3000
2000
1000
0
23
24
25
26
27
28
29
30
31
Gestational Age (weeks)

SUPPLEMENTARY FILE A

DATA CLEANING PROTOCOL

Reported weights (g) were filtered using the following protocol to correct transcription errors and discard implausible values:

If weight is greater 60000 or less than 60, discard it.
If weight is above a line defined by $\frac{2500}{11}CGA - \frac{1700}{11}$ *where CGA is the corrected gestational age in weeks, divide it by 10 (correcting for transcription errors).*
If weight is below a line defined by $\frac{125}{7}CGA - \frac{2500}{7}$*, multiply it by 10 (correcting for transcription errors).*
If weight has not been adjusted or discarded by the above steps and it lies above the line $\frac{650}{3}CGA - \frac{11500}{3}$ *or below the line* $\frac{400}{13}CGA - \frac{6700}{13}$*, discard it.*
Otherwise, retain it.

This protocol is coded as a function in R as follows:

```
function(weight,cga){
  if (weight>60000 | weight<60) NA else
  if (weight>(2500/11*cga-1700/11)) weight/10 else
  if (weight<(125/7*cga-2500/7)) weight*10 else
  if (weight>(650/3*cga-11500/3) | weight<(400/13*cga-6700/13))
NA else
  weight
}
```

SITAR models were then formed for each gestational age group using these filtered data. Weights with a standardised residual exceeding ±4 in the relevant SITAR model were excluded from the final SITAR model.

Appendix 5

Guideline for the nutritional care of infants on the neonatal unit of University Hospital Southampton

University Hospital Southampton NHS
NHS Foundation Trust

Nutritional Care of Infants in the Neonatal Unit Guideline

Version: 3.0

Date Issued:	Feb 2021
Review Date:	Feb 2024
Document Type:	Clinical Guideline

Executive Summary

Good nutrition is important at all stages of life. Babies are born at a time of rapid growth and formation of body tissues and organs, yet immature metabolism means they are unable to cope with either excess or lack of nutrients. Detail in both the quantity and quality of nutrients is critically important.

There is clear evidence that mother's breast milk confers many advantages to both babies and mothers. As well as providing optimal nutrition for human development, breast milk contains many factors which promote immune function and enable healthy intestinal development. Breast milk and breast-feeding should be the preferred milk feed and all mothers should be encouraged and supported to breast feed.

Preterm infants and those with congenital abnormalities or metabolic disorders may require nutrient supplements or special feeds, and may require a period of intravenous nutrition until the gut is able to support their needs.

Measuring growth and monitoring biochemical well-being is crucial to optimising nutrition in high risk individuals.

1 Scope and Purpose

These guidelines aim to provide both practical and theoretical guidance for the optimal nutrition for all sick and preterm infants on the neonatal unit in Southampton.

2 Definitions

AREDF	Absent or Reversed End Diastolic Flow (in umbilical artery, seen on antenatal scans)
AXR	Abdominal X-Ray
Babiven	Babiven start-up – does not contain NaCl and the glucose concentration is 10%
BMF	Breast Milk Fortifier
CPAP	Continuous Positive Airways Pressure
D/C	Discharge
DBM	Donor Breast Milk
DH	Department of Health
ELBW	Extremely Low Birth Weight (birth weight <1000g)
FBC	Full Blood Count
g	grams
IU	International Units
IUGR	Intrauterine Growth Restriction
IV	Intravenous
Kcal	kilocalories
Kg	kilogram
LBW	Low Birth Weight (birth weight <2500g)
LFT	Liver Function Tests
MBM	Maternal Breast Milk
mg	milligram
ml	millilitre
mmol	millimole
NBM	Nil By Mouth
NEC	Necrotising Enterocolitis
NICU	Neonatal Intensive Care Unit
NNU	Neonatal Unit
PBP	Potentially Better Practice
PDA	Patent Ductus Arteriosus
PDF	Post Discharge Formula
PN	Parenteral Nutrition
RCT	Randomised Controlled Trial
SD	Standard Deviation
TAT	Trans-anastamotic Tube
TPN	Total Parenteral Nutrition
U&E	Urea and Electrolytes
VLBW	Very Low Birth Weight (birth weight <1500g)
VON	Vermont Oxford Network

3 Details of procedure to be followed

Assessment and Monitoring of Growth

Growth Measurement
Regular measurements are vital to guide nutritional care and allow subsequent progress to be monitored. Weight measurements alone are not sufficient to determine adequate growth. Changes in weight in the early days of life usually reflect fluid balance: aim for weight loss of no more than 10% from birth weight. Once baby is stable and growing, aim for gain of 15-20 grams/kg/day. Head circumference and length: normally expect increase of 0.75 cm/week. Measurements will populate growth charts on Metavision.

Admission: Measure Weight, Length and Head circumference for ALL babies.

Weekly (Sunday night): Measure Weight, Length and Head circumference for all babies provided they are well enough. Infants on CPAP should have their hats removed and their head measured weekly unless they are very unstable.

Babies are also routinely weighed on Tuesday night and Thursday night to guide clinical care. Daily weights are often required for babies with renal failure.

Nutritional risk assessment
Nutritional screening is important to identify infants at risk of poor growth. There is a nutritional screening tool on Metavision based on the risk criteria below, the results of which guides rates of PN and enteral feeds (see appendix for flow chart). The nutrition team will review all high risk patients in ITU/HDU (Tuesday 9-10:30 neonatal seminar room) and others by request.

High risk
- Preterm <28 weeks
- ELBW < 1000g
- Severe IUGR (weight < 2nd centile with AREDF) <35 weeks
- Infant establishing feeds after episode of NEC or GI perforation
- Infants with severe congenital GI malformation: e.g. gastroschisis
- Infants with complex congenital heart disease

Moderate risk
- Preterm 28-31^{+6} weeks, otherwise well
- VLBW 1000 – 1500g
- Moderate IUGR (weight < 9th centile and AREDF) <35 weeks
- Baby on inotropes
- Baby on indomethacin/ibuprofen (NB avoid concomitant treatment with steroids)
- Baby >1500g with illness or congenital anomaly which may compromise feeding
- Perinatal hypoxia / ischaemia
- Symptomatic polycythaemia, with PCV $\geq$ 70%

Low risk
- Preterm 32-36^{+6} weeks, otherwise well
- AREDF / IUGR $\geq$35 weeks
- Term Infants >37 weeks

Nutrition Support for the Preterm Infant

OVERVIEW - GETTING STARTED - EARLY TPN AND TROPHIC MILK FEEDS

See following appendices for further information

- **Nutrient intake recommendation- Appendix A**
- **Nutritional content of common feeds and stock PN - -Appendix B**
- **Stock PN flow rates- appendix C**
- **PN monitoring including lipid adjustment for triglyceride levels – Appendix D**
- **High and Moderate risk infant flow chart for feeds and PN- Appendix E**
- **Suggested enteral feed volumes – Appendix F**

HIGH AND MODERATE RISK (see appendix E)
Aim to introduce milk feeds gradually while maintaining calorie and nutrient intake with PN Before starting or increasing milk ensure baby is clinically stable and abdomen soft. Small gastric residuals can be tolerated if baby well. Passage of meconium and then changing stools is an important indication of gut motility. Glycerine suppositories may help if no stool passed for 24 hours. Ensure mother has lactation support to start expressing (see breastfeeding care pathway)

High and moderate preterm (<31^{+6}; <1500g; moderate/severe IUGR/AREDFV <35 weeks)

First 24 hours	Start Babiven PN at 60-90 ml/kg/day via UVC or long line, as soon as possible unless baby very unstable. Ideally within 6 hours of birth. Give fresh colostrum as mouth care or as trophic feeds
24-72 hours	Continue colostrum as available until sufficient MBM is supplied to start trophic feeds as MBM 1 ml/kg 2-4 hourly. If at 48 hours there is no MBM as mother is unable or unwilling to provide MBM, then DBM can be used
48-72 hours	Change to Stock Preterm Concentrated. Increase milk by 10-20 ml/kg/day as tolerated (see table in Appendix D); aim to decrease PN flow rates with feeds only once baby on total fluids of 150ml/kg/day. Commence Abidec 0.6ml once daily when tolerating 60ml/kg/day enteral feed to provide additional vitamin A and D.

Special considerations

- If infant is felt to be at especially high risk of NEC (eg congenital cardiac anomaly) or if infants cannot have fortifier for some reason (eg post NEC surgical patients), then an alternative to fortification is a prolonged period on PN until 34 weeks
- Similarly, a higher total volume of fluids of **180ml/kg/day** may be used to enable a prolonged combination of both PN and unfortified/half fortified feeds if appropriate
- Ask for nutrition team review for complex or unusual situations.

LOW RISK
Low risk

First 24 hours	Commence milk feeds 30-60 ml/kg/day, supplemented by IV fluids if necessary
Beyond 72 hours	Increase milk feeds by 30 ml/kg/day as tolerated

Parenteral Nutrition

PN comprises an aqueous solution (glucose, amino acids, electrolytes and trace elements) and a lipid solution (which contains both fat- and water-soluble vitamins). For adequate nutrition it is important that the lipid is always run alongside the aqueous solution (except when well advanced on enteral feeds - see below). It is given via a **central venous line** (UVC, percutaneous longline or other central line). Rarely in cases where vascular access is a problem, lower concentration bespoke PN can be made and given by peripheral line – this is a consultant decision and requires careful consideration of the risks and benefits.

Indications for PN

- High or Moderate risk infants as described above (start within 6 hours)
- Birth weight >1500g – if enteral feeding contra-indicated, start PN by
    - 48 hours in 1500-2500g
    - 72 hours in 2500-3500g if NBM
- Stock PN is available to start at any time
- Term infants >3.5kg without risk factors should not routinely be started on PN until 5 days of inadequate enteral intake (excluding gastroschisis and other surgical problems- see surgical section in appendix K for more info including post-operative timing of PN)

Stock PN

- Babiven – For preterm infants (<37/40 gestation) for the first 48-72 hours.
- Preterm Concentrated- For preterm infants (<37/40 gestation) requiring maintenance sodium. **This should be the PN of choice for the majority of preterm infants after the first few days following birth**, as it contains more protein. Note that this bag may still be appropriate if the serum sodium is high due to a relative fluid deficit.
- Term PN – for Term infants (≥37 weeks gestation) at any point after birth.

Special prescription ('bespoke') PN

- Bespoke PN may also be appropriate where infants have electrolyte requirements than cannot be met with Stock PN and is prescribed in conjunction with the unit pharmacist. It is not available out of normal working hours.

Blood tests required (see appendix E)

First week of PN:
- Full TPN Profile daily (Renal, bone and liver profiles, inorganic phosphate, magnesium on eQuest) this includes U&E's, Calcium, magnesium phosphate and LFTs)
- FBC twice weekly

Second and subsequent weeks of PN:
- Full TPN Profile and FBC twice weekly if stable
- Triglycerides should be measured weekly (ideally Sunday night) when on IV lipid (see appendix E)
- If on PN for 3 weeks or more, measure Trace elements (Zn, Cu, Se, Mn – use special blood bottle in Dr's Office) and Vitamins (A, D and E) and repeat monthly if a baby is still on PN.
- When on enteral feeds, Infants in the High and Medium risk categories need weekly FBC, U&Es, LFTs and Bone profiles, Magnesium and inorganic Phosphate once they are off PN

and fully enterally fed. This can be extended to once fortnightly when babies are moved into Special Care.

Cautions on PN

SEPSIS - may affect lipid metabolism; measure triglycerides and if >3mmol/L adjust lipid (see appendix D)

CHOLESTATIC JAUNDICE – total and prolonged PN increases the risk, so try to give some enteral feed if at all possible; other risk factors include IUGR, sepsis and short bowel syndrome. SMOF lipid solutions may be beneficial in cases of cholestasis, and should be considered in high risk babies if expected to be on PN for 4 weeks or more.

SURGERY – see surgical section below in appendix K

Reducing PN as enteral feeds increase

- **Only once the infant is receiving 150ml/kg/day total fluids should the aqueous PN solution be decreased as enteral feeds increase** (unless there is a clinical decision to restrict fluids).
- Note that if there is no plan to fortify the infant or if fortifier is felt to be contraindicated, high and moderate risk preterm infants will need to work to a total of 180ml/kg/d total feeds rather than 150ml/kg/d. Also, even at 180ml/kg/d unfortified breast milk will not meet the requirements of preterm infants born at <1.8kg, so thought must be given to a longer period on PN or other supplementation in order to maintain growth. Using unfortified breast milk as the only feed is nutritionally inadequate for preterm infants so should only be considered if there are no other viable options
- Once the infant is on 90ml/kg/day enteral feeds, the rate of lipid infusion should be halved, and then stopped when the infant reaches 120ml/kg/day enteral feeds. Any shortfall in total fluid volume due to the reduction in lipid should be made up by increasing the aqueous PN solution, to allow maximum protein to be delivered to the infant (though do not exceed the maximum prescribed rate). This is important when infants are on Stock PN, but for those on bespoke PN, the reduction in lipid may have already been done/accounted for by the pharmacists when the PN was prescribed so may not be necessary (check with the pharmacists first). **Remember that once the lipid is reduced, vitamin intake will be inadequate until vitamin supplements (Abidec or Dalivit) are started.**

Enteral Feeds

Mother's breast milk (MBM) is almost always the feed of first choice, unless contraindicated by maternal illness, drugs or maternal reasons. If no MBM is available pasteurised donor breast milk (DBM) may be used for high and moderate risk babies with parental consent – see **appendix G for DBM guidance**. Preterm formula (Nutriprem 1) is indicated for infants with birth weight <1.8kg grams; Post discharge formula (Nutriprem 2) is indicated for preterm infants either as sole diet or in addition to breast-feeding from around 36 weeks (or discharge) up to 6 months corrected age.

Other formulas which may be used in special circumstances are summarised in the table below. Note that ideally term infant formulas should not routinely be used for preterm infants unless under specialist advice.

Feed name	Characteristic
• Nutriprem 1	**Preterm** infant formula – whole protein nutrient enriched
• Nutriprem 2	**Preterm** discharge formula – whole protein nutrient enriched
• Hydrolysed Nutriprem 1	**Preterm** extensively hydrolysed whey/casein protein, lactose containing
• Infatrini, SMA High Energy, Similac High Energy	High energy nutrient dense feed for **term infants**, whole whey protein, lactose containing
• Infatrini Peptisorb	High energy nutrient dense feed for **term infants** extensively hydrolysed whey protein, MCT 50%, lactose free
• Nutramigen, Peptijunior, Althera, Aptamil Pepti	Extensively hydrolysed **term infant** formula whey/ casein protein, MCT containing, lactose free
• Alfamino, Neocate, PurAmino	Amino acid **term infant** fromulas, MCT containing, lactose free
• Monogen, Kindergen	Specialist **term infant** formula, whole protein, lactose containing

Nutritional supplementation

BREAST MILK FORTIFIER

'Multi-component' fortifier provides additional calories (carbohydrate), extensively hydrolysed protein (cows' milk based), minerals and vitamins in a powder which is added to MBM.

- It should be routine for all moderate and high risk babies together with those in the late preterm group who are less than 1.8kg at birth and who are exclusively breast fed
- It should also be considered for late preterm infants (<35 weeks gestation) whose mother's are intending to breastfeed or if growth is poor.

Once babies are starting to breast feed, the fortifier can be prescribed as a bolus (see below) which can then be continued on discharge until 48-52 weeks CGA. This should be done under dietetic supervision and can be supported by the Neonatal Home Team.

Infants who are fed with a mixture of fortified breast milk and preterm formula do not need fortifier to be added if they are on 50% or greater preterm formula (assuming they are on 180ml/kg/day)

BMF SUPPLEMENTS AT HOME

This is given as 4 sachets BMF mixed in 40mls of freshly expressed breast milk. This should be given via cup or using a 5ml syringe, giving 5ml 8 times a daily with a breastfeed to support weight gain until 48-52 weeks CGA(1). This should be done under dietetic supervision with the support of the Neonatal Home Team. Parents should be given the UHS parent information sheet on how to make up BMF supplements, which can be found here on staffnet.

MUM PLANNING TO FORMULA FEED

- Babies <34 weeks gestation, with birthweight <1.8kg can be considered for discharge on Post-Discharge Formula (PDF) – Nutriprem 2. This should be continued until 3 to 6 months corrected age.
- ELBW and VLBW babies who have been on Nutriprem 1 should be changed to PDF at approximately 36 weeks corrected age, or taking most feeds by bottle. For those with very poor growth, continuing with Nutriprem 1 formula to 40 weeks corrected age may be appropriate.
- Babies discharged on PDF should have **Abidec 0.6 ml, but not Sytron**.
- If changing to term formula, prescribe Abidec 1 ml (continue until at least one year post term) and Sytron 1ml (continue until 6 month post term)

SOLIDS

Can be introduced at 5-8 months chronological age (or 4-6 months corrected gestational age), depending on developmental stage and degree of prematurity.

VITAMINS, IRON, ZINC, TRACE ELEMENTS AND OTHER SUPPLEMENTS

Vitamins and Iron

Breast milk provides insufficient vitamins (particularly vitamin A and D) and iron for preterm infants. Preterm infants have low levels of fat soluble vitamins, particularly A and D.

- Preterm infants <36 weeks should therefore be commenced on Abidec 0.6ml once daily once tolerating 60ml/kg/day enteral feed (different types or amounts of vitamin supplements may be recommended by the nutrition team in special circumstances).
- All preterm infants <36 weeks should then be switched to Abidec 1ml OD once intravenous lipid is stopped or on reaching full feeds, whichever is sooner, in order to ensure adequate vitamin D provision (unless they are on preterm formula, in which case 0.6ml OD will suffice). 1ml Abidec should be continued until 1 year of age and then a standard dose multivitamin preparation containing vitamin D should be continued until 5 years in line with current Department of Health advice.
- Sytron (iron) should be started at 1ml OD from day 28 in preterm infants on breast milk (even if fortified) or term formula until 1 year of age. Preterm formula is fortified with iron, so iron supplements are not required.

Zinc and Trace elements

An enteral intake of Zinc of 1.4 – 2.5mg/kg/day is recommended and a parental intake around 400µg/kg/day (Appendix A). In patients with significant enterostomy fluid losses, plasma zinc should be regularly reviewed, since they have a high risk of zinc deficiency (2).

When infants are gaining weight rapidly and are sequestering nutrients into tissues, intakes of not only zinc but copper and other nutrients will increase. For infants who have serum zinc levels <11.2 with C-reactive protein <10, zinc supplementation should be commenced at a dose of 2mg/kg/day for 4 – 8 weeks (3-5) .

For Preterm Infants

Zinc (trace element) levels should be measured on or around day 21 of life in:

- All babies who are still on PN at day 21 of life
- All babies who are 'high risk' according to the nutrition guidelines (ie <1000g or <28 weeks at birth)

If infants have low zinc levels (below the reference range on eQuest <11.2) then they should start zinc supplements of **2mg/kg/day** for a period of 4 weeks. After the 4 weeks of

supplementation, they do not need their zinc levels rechecking unless they are displaying poor growth or clinical signs of zinc deficiency (e.g. skin lesions)

For Surgical infants
Many surgical infants will fall into the above criteria for zinc monitoring and supplementation. In addition to this, any baby with an ileostomy who is 21 days or older should have their zinc (trace element) levels tested. If they are deficient, the should also start zinc supplements of **2mg/kg/day** for a period of 4 weeks as above.

However, for infants with ileostomies, supplements should only be started once full feeds have been established, other electrolyte supplements have been started where necessary (and tolerated) and Abidec and Sytron have been commenced and tolerated. Once all these things are addressed, zinc supplements can be started, but need to be done so in isolation (e.g. not with a lot of other feed/supplement changes) so it is possible to be sure that infants are tolerating the zinc from a stoma output perspective.

Supplements and Metabolic bone disease of prematurity (MBDP)

Metabolic bone disease of prematurity (MBDP) is caused by the under mineralisation of the preterm infant's skeleton due to inadequate intake calcium and phosphorus due the perinatal period. Poor growth and fragility fractures may be a consequence of inadequate intake of calcium and phosphorus although the appropriate use of breastmilk fortifier to maternal/ donor breastmilk or the use of preterm formula can help to reduce the incidence.

Biochemical screening e.g. monitoring alkaline phosphatase (ALP), calcium, phosphorus, vitamin D and parathyroid hormone allows for early identification and appropriate nutritional support/ mineral supplementation (see Appendix N for Flow diagram for the management of MBD in prematurity).

Management of Common Gut and Feeding Problems

a. **Gastric aspirates / residuals** – preterm infants have immature gut motility, and aspirates/residuals and small vomits are not uncommon. Large volume aspirates or dark green bile stained aspirates, particularly in association with abdominal distension and / or tenderness are a cause for concern. However small milky / yellow aspirates up to 2-3 mls are frequently normal. They can be replaced, and feeds continued (see Appendix H)

b. **Abdominal distension** – this is another common feature in preterm infants, due to poor gut motility. It tends to be more common in babies on CPAP, with high volumes of air flowing into the upper airway and oesophagus. Tenderness, or systemic symptoms and signs such as apnoea, tachycardia or temperature instability should raise concern. If baby is otherwise well, a small glycerol (glycerin) suppository may help to stimulate peristalsis, and enable feeds to be continued (see Appendix H).

c. **Suspected NEC** – classical features are blood and mucous in stools, bile stained aspirates and abdominal tenderness. Systemic signs such as tachycardia and hypotension occur in severe NEC. X-ray might show intramural gas (pneumatosis coli), dilated loops of bowel, free air, or a gas-less bowel. In suspected NEC feeds should be stopped, and urgent attention paid to supporting ventilation, circulation and fluid balance.

d. **Gastro-Oesophageal Reflux (GOR)** – See Appendix I

e. **Cow's Milk Protein Allergy (CMPA)** – See Appendix J

Other Special Cases

a. **Nutritional Management of Surgical Infants** – See Appendix K

b. **Nutritional Management of Infants with Congenital Heart Disease** – See Appendix L

c. **Nutritional management of neonates/infants with liver dysfunction** – See Appendix M

d. **Nutritional management of metabolic bone disease of prematurity** -See Appendix N

e. **Nutritional management of infants with short bowel syndrome** – See Appendix O

4 Roles and Responsibilities

BREAST-FEEDING AND LACTATION SUPPORT

- All staff: awareness of Trust Policy and NNU Guidelines
- NNU lactation support team – Lead Charlotte Oates: expert guidance for mothers breast-feeding and/or expressing milk in NNU

PARENTERAL NUTRITION

- All staff: awareness of need for PN in high risk infants
- Nursing staff: awareness of location of 'stock' PN in NNU and knowledge and skills for PN administration appropriate to nursing skill level
- Medical staff: awareness of PN supplies available and how to prescribe; awareness of potential complications of PN and how to avoid
- Pharmacists: expertise in detailed composition of PN solutions and provision of PN in different situations on NNU

ENTERAL NUTRITION

- All staff: support for mothers in informed choice of feeding, recognising that breast milk is the preferred feed for all infants, particularly those born preterm
- All staff: awareness of choices for enteral nutrition: maternal breast milk / breast-feeding; donor breast milk / milk bank; standard infant formula; formulas for preterm infants; special formulas for infants with specific gut or feeding problems
- Neonatal Dietitian: expert knowledge of composition of breast milk and alternatives and guidance on making appropriate choices
- Surgical team: expert knowledge on potential feeding challenges in infants with congenital or acquired abnormalities of the gut, particularly following surgery.

FEEDING DIFFICULTIES

- All staff: awareness of common feeding difficulties of preterm infants and those with neurological complications
- Speech and language therapist: expert knowledge of structure and function of upper gastro- intestinal tract and how to optimise feeding potential of vulnerable babies

GROWTH MONITORING

- All staff: Awareness of importance of making accurate and regular measurements and plotting them on appropriate charts to monitor growth
- Nursing staff: Weigh babies at intervals as indicated by clinical condition (ideally three times per week) and head circumference and length weekly.

SPECIAL CASES

- Neonatal Nutrition Team: Will review high risk medical, surgical or complex patients on weekly nutrition ward round

5 Related Trust Policies and documents

Lactation and Breastfeeding in the NNU Guideline
Breastfeeding Term Infants: Guideline
PIER Guideline Gastro-Oesophageal Reflux
PIER Guideline Cow's Milk Protein Allergy
Milk free diet for breast feeding mothers - information leaflet

6 Implementation

Information for all staff on induction and regular updates as necessary.

7 Process for Monitoring Compliance/Effectiveness

Key aspects of the procedural document that will be monitored:

What aspects of compliance with the document will be monitored	What will be reviewed to evidence this	How and how often will this be done	Detail sample size (if applicable)	Who will co-ordinate and report findings	Which group or report will receive findings
None required					

8 Arrangements for Review of the Policy

This guideline will be reviewed every 3 years.

9 References

1. Marino LV, Fudge C, Pearson F. Home use of breast milk fortifier to promote postdischarge growth and breast feeding in preterm infants: a quality improvement project. 2018.
2. Koletzko B, Poindexter B, Uauy R. Nutritional Care of Preterm Infants: Scientific Basis and Practical Guidelines. Basel: Karger; 2014.
3. Islam MN, Chowdhury MA, Siddika M, Qurishi SB, Bhuiyan MK, Hoque MM, et al. Effect of oral zinc supplementation on the growth of preterm infants. Indian pediatrics. 2010;47(10):845-9.
4. Golden MH. Proposed recommended nutrient densities for moderately malnourished children. Food and nutrition bulletin. 2009;30(3 Suppl):S267-342.
5. Formulary BN. ZINC SULFATE. National Institute for Health and Care Excellence [Internet]. 2018 13th July 2018. Available from: https://bnf.nice.org.uk/drug/zinc-sulfate.html.
6. Onyeador N, Paul SP, Sandhu BK. Paediatric gastroesophageal reflux clinical practice guidelines. Arch Dis Child Educ Pract Ed. 2014;99(5):190-3.
7. Terrin G, Passariello A, De Curtis M, Manguso F, Salvia G, Lega L, et al. Ranitidine is Associated With Infections, Necrotizing Enterocolitis, and Fatal Outcome in Newborns. Pediatrics. 2011;129(1):e40-e5.
8. Guillet R, Stoll BJ, Cotten CM, Gantz M, McDonald S, Poole WK, et al. Association of H2-blocker therapy and higher incidence of necrotizing enterocolitis in very low birth weight infants. Pediatrics. 2006;117(2):e137-42.
9. Grimshaw KE, Bryant T, Oliver EM, Martin J, Maskell J, Kemp T, et al. Incidence and risk factors for food hypersensitivity in UK infants: results from a birth cohort study. Clinical and translational allergy. 2015;6:1.
10. Venter C, Brown T, Meyer R, Walsh J, Shah N, Nowak-Wegrzyn A, et al. Better recognition, diagnosis and management of non-IgE-mediated cow's milk allergy in infancy: iMAP-an international interpretation of the MAP (Milk Allergy in Primary Care) guideline. Clinical and translational allergy. 2017;7:26.

11. Fisher HR, Du Toit G, Bahnson HT, Lack G. The challenges of preventing food allergy: Lessons learned from LEAP and EAT. Annals of allergy, asthma & immunology : official publication of the American College of Allergy, Asthma, & Immunology. 2018;121(3):313-9.
12. Pierro A, Eaton S. Metabolism and nutrition in the surgical neonate. Seminars in pediatric surgery. 2008;17(4):276-84.
13. Turi RA, Petros AJ, Eaton S, Fasoli L, Powis M, Basu R, et al. Energy metabolism of infants and children with systemic inflammatory response syndrome and sepsis. Annals of surgery. 2001;233(4):581-7.
14. Fivez T, Kerklaan D, Mesotten D, Verbruggen S, Wouters PJ, Vanhorebeek I, et al. Early versus Late Parenteral Nutrition in Critically Ill Children. N Engl J Med. 2016;374(12):1111-22.
15. van Puffelen E, Vanhorebeek I, Joosten KFM, Wouters PJ, Van den Berghe G, Verbruggen S. Early versus late parenteral nutrition in critically ill, term neonates: a preplanned secondary subgroup analysis of the PEPaNIC multicentre, randomised controlled trial. Lancet Child Adolesc Health. 2018;2(7):505-15.
16. Westin V, Stoltz Sjostrom E, Ahlsson F, Domellof M, Norman M. Perioperative nutrition in extremely preterm infants undergoing surgical treatment for patent ductus arteriosus is suboptimal. Acta paediatrica (Oslo, Norway : 1992). 2014;103(3):282-8.
17. Denne SC PB, Leitch CA. Nutrition and metabolism in the high risk neonate. In: Martin RJ, Fanarof AA, Walsh MC. Philadelphia, PA: Mosby-Elsivier; 2006.
18. Garza JJ, Shew SB, Keshen TH, Dzakovic A, Jahoor F, Jaksic T. Energy expenditure in ill premature neonates. Journal of pediatric surgery. 2002;37(3):289-93.
19. Tsang R UR, Koletzko B, Zlotkin S. Nutrition of the Preterm Infant: Scientific Basis and Practical Guidelines: Digital Educational Publishing Inc; 2005.
20. Owens JL, Musa N. Nutrition support after neonatal cardiac surgery. Nutrition in clinical practice : official publication of the American Society for Parenteral and Enteral Nutrition. 2009;24(2):242-9.
21. Dinleyici EC, Kilic Z, Buyukkaragoz B, Ucar B, Alatas O, Aydogdu SD, et al. Serum IGF-1, IGFBP-3 and growth hormone levels in children with congenital heart disease: relationship with nutritional status, cyanosis and left ventricular functions. Neuro endocrinology letters. 2007;28(3):279-83.
22. Medoff-Cooper B, Ravishankar C. Nutrition and growth in congenital heart disease: a challenge in children. Current opinion in cardiology. 2013;28(2):122-9.
23. Meyer R, Venter C, Fox AT, Shah N. Practical dietary management of protein energy malnutrition in young children with cow's milk protein allergy. Pediatric allergy and immunology : official publication of the European Society of Pediatric Allergy and Immunology. 2012;23(4):307-14.
24. Meyer R ML. Critically illness. Oxford: Blackwell Publishing; 2014.
25. Pillo-Blocka F, Adatia I Fau - Sharieff W, Sharieff W Fau - McCrindle BW, McCrindle Bw Fau - Zlotkin S, Zlotkin S. Rapid advancement to more concentrated formula in infants after surgery for congenital heart disease reduces duration of hospital stay: a randomized clinical trial. (0022-3476 (Print)).
26. Marino LV, Venkatesh P, Ho A, Beattie RM, Bharucha T. Hypophosphataemia in infants with CHD treated with amino acid infant formula. Cardiology in the young. 2018:1-5.
27. Marino LV, Johnson MJ, Hall NJ, Davies NJ, Kidd CS, Daniels ML, et al. The development of a consensus-based nutritional pathway for infants with CHD before surgery using a modified Delphi process. Cardiology in the young. 2018;28(7):938-48.
28. Dhawan A. Acute liver failure in children and adolescents. Clinics and research in hepatology and gastroenterology. 2012;36(3):278-83.
29. Dhawan A, Cheeseman P, Mieli-Vergani G. Approaches to acute liver failure in children. Pediatric transplantation. 2004;8(6):584-8.
30. Clinical Paediatric Dietetics. Shaw V LM, editor. Oxford: Blackwell Synergy; 2014.

31. Welling L, Bernstein LE, Berry GT, Burlina AB, Eyskens F, Gautschi M, et al. International clinical guideline for the management of classical galactosemia: diagnosis, treatment, and follow-up. Journal of inherited metabolic disease. 2017;40(2):171-6.
32. Kyle UG, Shekerdemian LS, Coss-Bu JA. Growth failure and nutrition considerations in chronic childhood wasting diseases. Nutrition in clinical practice : official publication of the American Society for Parenteral and Enteral Nutrition. 2015;30(2):227-38.
33. Leonis MA, Balistreri WF. Evaluation and management of end-stage liver disease in children. Gastroenterology. 2008;134(6):1741-51.
34. Sultan MI, Leon CD, Biank VF. Role of nutrition in pediatric chronic liver disease. Nutrition in clinical practice : official publication of the American Society for Parenteral and Enteral Nutrition. 2011;26(4):401-8.
35. Yang CH, Perumpail BJ. Nutritional Needs and Support for Children with Chronic Liver Disease. 2017;9(10).
36. Hurtado-Lopez EF, Vasquez-Garibay EM, Trujillo X, Larrosa-Haro A. Body Composition Predicts Growth in Infants and Toddlers With Chronic Liver Disease. Journal of pediatric gastroenterology and nutrition. 2017;65(6):e117-e9.
37. Widodo AD, Soelaeman EJ, Dwinanda N, Narendraswari PP, Purnomo B. Chronic liver disease is a risk factor for malnutrition and growth retardation in children. Asia Pacific journal of clinical nutrition. 2017;26(Suppl 1):S57-s60.
38. Berry GT. In: Adam MP AH, Pagon RA, et al. Classic Galactosemia and Clinical Variant Galactosemia. Feb 4 [Updated 2017 Mar 9]. . Seattle; 1993-2019. : University of Washington,Seattle (WA):. Available from: https://www.ncbi.nlm.nih.gov/books/NBK1518/.
39. Ficicioglu C, Hussa C, Yager C, Segal S. Effect of galactose free formula on galactose-1-phosphate in two infants with classical galactosemia. European journal of pediatrics. 2008;167(5):595-6.
40. Zlatunich CO, Packman S. Galactosaemia: early treatment with an elemental formula. Journal of inherited metabolic disease. 2005;28(2):163-8.
41. Chinoy A, Mughal MZ, Padidela R. Metabolic bone disease of prematurity: causes, recognition, prevention, treatment and long-term consequences. Arch Dis Child Fetal Neonatal Ed. 2019;104(5):F560-F6.

Appendix A: Nutrient Requirements

Nutrient requirements for Term and Preterm infants in the first weeks of life are summarised below. The figures shown below are based on the parenteral requirements for the first week, and the enteral requirements for the subsequent weeks. It should be noted that these are average recommendations and some infants may require additional nutrients such as sodium, potassium and zinc as dictated by the results of blood tests.

Term infants – based on intake in 150 ml/kg breast milk; preterm infants based on recommendations in Koletzko 2014 unless otherwise stated(2).

There are no specific guidelines for those babies born over 1.5kg and under term weight (2.5 kg) but it can be anticipated that their nutritional needs will be between those of preterm infants and term infants. Nutritional support should therefore aim to deliver nutrient intakes in this area.

Nutrient(Unit/kg/day)	Term infant	Preterm VLBW <1500g (enteral)
Energy (kcal)	90-100	110-130 (85-95IV)
Protein (g)	1.5-2.1	3.5-4.5
Nitrogen (g)	0.24-0.34	0.56-0.72
Sodium (mmol)	1.4	2.4-5.0
Potassium (mmol)	2.0	2.0-5.0
Calcium (mmol)	1.25	3.0-5.0
Phosphate (mmol)	1.3	1.9-4.5
Vitamin D IU*	340	400-1100
Vitamin A IU**	1150	1300-3614
Iron (umol)	17.9	35.7-53.6

*Vitamin D = dose quoted is total daily dose (340 IU = 8.5 mcg Vit D)

**Vitamin A = dose quoted is total daily dose (1150 IU = 350 mcg of Vitamin A retinol equivalent)

Recommendations for micronutrient intakes (per/kg/day) in ELBW and VLBW infants (2)

Nutrient	**Enteral (per/kg/day)**	**Parenteral (per/kg/day)**	**Content in 2ml Peditrace**	**Maternal breastmilk**		**Fortified breastmilk**		**Nutriprem 1**		**Alfamino**			**Infatrin/ Infatrini Peptisorb**	
				100ml	**150ml**	**100ml**	**150ml**	**100ml**	**150ml**	**100ml**	**150ml**		**100ml**	**150ml**
Iron (mg)	2 – 3	0 – 0.25	-	0	0	0	0	1.4	2.1	0.7	1		1.0	1.5
Zinc (mg)	**1.4 – 2.5**	**0.4***	**0.5**	**0.3**	**0.45**	**0.9**	**1.35**	**0.9**	**1.35**	**0.7**	**1**		**0.9**	**1.35**
Copper (µg)	100 – 230	40*	40	40	60	75	112.5	80	120	56	84		60	90
Selenium (µg)	5 – 10	5 – 7*	4	2.2	3.3	4	6	1.9	2.85	1.8	2.7		2.0	3
Manganese (µg)	1 – 15	1*	2	0.44	0.66	8.54	12.8	10	15	7	10.5		0.08	0.12
Iodine (µg)	10 – 55	10*	2	16.8	25.2	27.8	41.7	25	37.5	11	16.5		14	21
Chromium (µg)	0.03 – 2.25	0.05 – 0.3*	-	-	-	-	-	-	-	-	-		4	6
Molybdenum (µg)	0.35	0.25*	-	-	-	-	-	-	-	-	-		6	9

*Approximate values. Iodine recommendations assumes no use of iodine containing antiseptic

University Hospital Southampton NHS
NHS Foundation Trust

Appendix B: Nutritional content of common feeds and stock PN

Nutrient Content of Commonly Used Products per 100ml

Typical Values are used and are correct at 06/05/2020

*Based on Cow and Gate Nutriprem Breast Milk Fortifier

Fluid Name / Nutrient	Babiven Stock PN	Preterm concentrated Stock PN	Term Stock PN	Stock Lipid	Dextrose 10%	MBM/DBM	MBM with Full Fortifier*	Alfamino	Peptijunior	LBW Formula (Nutriprem 1)	Post D/C Formula (Nutriprem 2)	Term formula	Infantrini
Energy (kcal)	50.7	84.8	69.8	171	40	69	80	70	66	80	75	66	100
Protein (g)	2.68	4.18	2.5	0	0	1.3	2.6	1.9	1.8	2.6	2	1.3	2.6
Carbohydrate (g)	10	17	14.9	0	0	7.2	**9.6**	7.9	6.8	8.4	7.4	7.3	10.3
Fat (g)	0	0	0	17.1	0	4.1	3.5	3.4	3.5	3.9	4	3.5	5.4
Sodium(mmol)	0	6.05	2.8	0	0	0.7	2.2	1.1	0.9	3	1.2	0.7	1.1
Potassium (mmol)	1	2.4	1.9	0	0	1.5	2.1	2	1.7	2.1	2	1.6	2.4
Calcium(mmol)	1	1.2	0.92	0	0	0.8	2.5	1.4	1.2	2.3	2.2	1.2	2
Phosphorous (mmol)	1	2.7	0.95	1.5	0	0.5	1.7	1.3	0.9	2	1.5	0.9	1.3
Magnesium	0.2	0.25	0.19	0	0	0.12	0.33	0.27	0.21	0.33	0.29	0.21	0.34
Iron (umol)	0	0	0	0	0	1.3	1.3	12.5	13.8	25.1	17.9	9.5	21.5
Zinc	4	3.9	3.6	0	0	4.6	13.8	10.7	7.6	16.8	13.8	7.6	13.8
Vitamin A (IU)	0	0	0	3997.6	0	213	985.6	307	173.2	599.4	269.7	183.2	333
Vitamin D (IU)	0	0	0	695.2	0	0	200	40	52	120	68	48	68
Volume (ml/kg) required to reach recommended protein intake (ELBW infants)	130	82.5	140	Contains no protein	Contains no protein	292	152	195	211	146	190	292	146

Nutritional Content of Stock PN based on ml/kg/day.

Babiven

Fluid ml/kg/day	60	90	120	130
Babiven bag (ml/**kg**/day)	55	82.5	110	117.5
Lipid syringe (ml/**kg**/day)	5	7.5	10	12.5
Provides (per kg):				
Nitrogen (g)	0.24	0.35	0.47	0.5
Protein (g)	1.5	2.2	2.9	3.1
Glucose (g)	5.5	8.3	11	11.8
Fat (g)	1	1.5	2	2.5
Kcal/kg	38	57	75.6	84.6
Sodium (mmol)	0	0	0	0
Potassium (mmol)	0.6	0.8	1.1	1.2
Total Phosphate (mmol)	0.6	0.9	1.3	1.4
Calcium (mmol)	0.6	0.8	1.1	1.2
Magnesium (mmol)	0.1	0.16	0.22	0.23
Total chloride (mmol)	0	0	0	0
Zinc (micromol)	2.2	3.3	4.4	4.7
Acetate (mmol)	0	0	0	0
Cal:1gN ratio	133	138	136	144

Preterm Concentrated

Fluid ml/kg/day	77.5	90	100
Pre-term Concentrated (ml/**kg**/day)	67.5	75	82.5
Lipid syringe (ml/**kg**/day)*	10	15	17.5
Provides (per kg):			
Nitrogen (g)	0.45	0.49	0.55
Protein (g)	2.81	3.09	3.43
Glucose (g)	11.48	12.6	14
Fat (g)	2	3	3.5
Kcal/Kg	77.2	92.8	104.7
Sodium (mmol)	4.09	4.49	4.99
Potassium (mmol)	1.64	1.8	2
Total phosphate (mmol)	1.97	2.23	2.49
Calcium (mmol)	0.81	0.89	0.99
Magnesium (mmol)	0.172	0.19	0.21
Total chloride (mmol)	1.03	1.13	1.25
Peditrace	0.68	0.75	0.83
Acetate (mmol)	1.64	1.8	2
Cal:1g N ratio	147	164	166

Term bags

Fluid requirement ml/kg/day	60	90	120	135
Term bag (ml/**kg**/day)	55	80	105	120
Lipid syringe (ml/**kg**/day)	5	10	15	15
Provides (per kg):				
Nitrogen (g)	0.22	0.32	0.42	0.48
Protein (g)	1.38	2.00	2.63	3.00
Glucose (g)	8.2	11.9	15.7	17.9
Fat (g)	1	2	3	3
Kcal/kg	48.3	75.6	103.3	113.6
Sodium (mmol)	1.53	2.23	2.93	3.34
Potassium (mmol)	1.04	1.51	1.98	2.27
Total phosphate (mmol)	0.6	0.9	1.22	1.36
Calcium (mmol)	0.50	0.73	0.96	1.10
Magnesium (mmol)	0.11	0.15	0.20	0.23
Total chloride (mmol)	1.79	2.60	3.42	3.91
Peditrace	0.51	0.74	0.98	1.12
Acetate (mmol)	1.04	1.51	1.98	2.27
Cal: 1gN ratio	195	211	221	212

*Please note total phosphate is received when the patient is receiving full lipid - Clinoleic 20% contains 1.5mmol PO4/100ml

Appendix C: Flow rates of Stock PN Solutions

The flow rates given in the tables below are per kg and therefore will need to be multiplied by the baby's weight.

Babiven Start-Up Bags (sodium-free)
Suitable for <30/40 weeks gestation

Day of parenteral nutrition	1	2	3	4
Babiven Start-Up (ml/**kg**/day)	55	82.5	110	117.5
Lipid Syringe (ml/**kg**/day)	5	7.5	10	12.5
Total Fluid (ml/**kg**/day)	60	90	120	130

Preterm concentrated can be started from 48 hours onward depending on electrolytes

Preterm concentrated
Suitable for transferring from Babiven Start-up at Day 3

Day of parenteral nutrition	**3**	4	5+
Preterm concentrated (ml/**kg**/day)	67.5	75	82.5
Lipid Syringe (ml/**kg**/day)	10	15	17.5
Total Fluid (ml/**kg**/day)	77.5	90	100

If a larger fluid volume is required additional 5% glucose may be run alongside the PN to bring the total fluids up to the desired total (ensure other infusions are also considered). 5% glucose can also be used to deliver additional electrolytes if required.

Term Bags

Day of parenteral nutrition	1	2	3	4+
Term Bag (ml/**kg**/day)	55	80	105	120
Lipid Syringe (ml/**kg**/day)	5	10	15	15
Total Fluid (ml/**kg**/day)	60	90	120	135

Prescribing
The stock PN should be prescribed, including a maximum rate of infusion, on Metavision. Two prescriptions are required, one for the aqueous nutrition bag and one for the lipid syringe.

Notes

- The lipid infusions should remain at a constant daily rate.
- The rate of the nutrition bag should be adjusted to take into account other continuous infusions.
- "Day of parenteral nutrition" in tables above is a rough guide only. Clinical need and fluid status should be taken into account when deciding on PN regimen.
- No additions should be made to stock PN.
- Each bag may be used for a maximum of 48hours
- Lipid must be changed every 24 hours
- Babiven & Term stock bags - If PN is being started at a rate of 120ml/kg/day or greater, the lipid should be prescribed at half-rate for 24hours, the additional volume may be made up from the aqueous bag
- Preterm concentrated stock bag – If PN is being started for the first time at a rate of 90ml/kg/day or greater, the lipid should be prescribed at half-rate for 24 hours. The additional volume may be made up by infusing additional glucose 5%.

Appendix D: Monitoring Requirements whilst on Parenteral Nutrition

Daily (Until stable and then twice weekly)

	Target levels for preterm infants (NB not same as lab normal range)
Sodium	136-144mmol/l
Potassium	3.5-5mmol/l
Corrected Calcium	2.15-2.6mmol/l
Ionised Calcium (blood gas)	1.0 – 1.3mmol/l
Phosphate	1.6-2.8mmol/l
Creatinine	30-65 micromol/l
Urea	2.9-7.1mmol/l
Albumin	25-45g/l
Bilirubin	<200 micromol/l
Blood Glucose	2.6-5.5mmol/l

Weekly

* Preterm-infants should have their triglyceride levels measured on day 3 of PN.

	Target levels for preterm infants (NB not same as lab normal range)
Magnesium	0.74-1.03mmol/l
Triglycerides*	< 3 mmol/l
Alk. Phos.	150-425 IU/l
ALT	10-40 IU/l
Weight	At least weekly

From day 21 on PN and then Monthly

	Target levels for preterm infants (NB not same as lab normal range)
Zinc	11-24 mol/l
Selenium	0.2-0.9micromol/litre
Copper	3-11 micromol/l
Manganese	120 – 325 nmol/l
Vitamin D	> 50 nmol/l
Vitamin A	0.7-1.5 micromol/l
Vitamin E	7-21 micromol/l

Impact of acute phase response on serum trace elements levels.

Trace Elements	Effect of Acute Phase Response
Copper	Increased
Ferritin	Increased
Iron	Decreased
Zinc	Decreased
Plasma Selenium †	Decreased
Chromium	Decreased
Manganese	No effect
Iodine, Molybdenum	Unknown

† Red cell selenium is not affected by acute phase response.

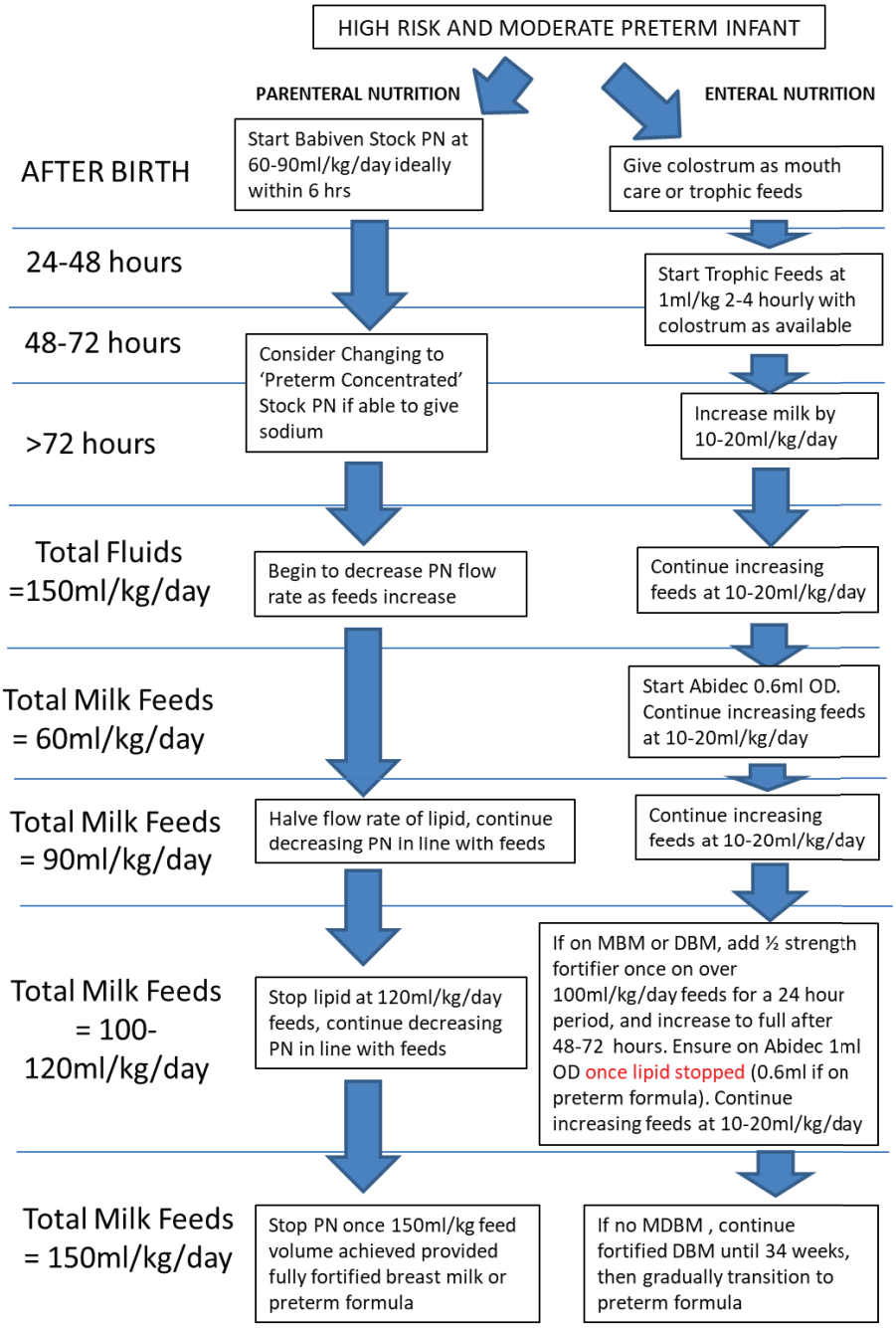
Table from Zemrani B, McCallum Z, Bines JE. Trace Element Provision in Parenteral Nutrition in Children: One Size Does Not Fit All. Nutrients. 2018;10(11).

Triglyceride Monitoring and Management for Neonates receiving PN.

- Triglyceride (TG) levels should be measured at least weekly, preterm babies should have their TG levels measured on day 3 of PN.
- Triglyceride levels may need to be measured more frequently in cases of sepsis, catabolism, hyperglycaemia, thrombocytopenia and liver impairment.
- Hypertriglyceridaemia is defined as > 3mmol/l in neonates.
- If the level is > 3mmol/l do NOT stop the lipid but reduce the rate of the infusion (see recommendations below).
- To prevent essential fatty acid deficiency, pre-term infants require a minimum intake of 1g/kg/day and term infants requires a minimum intake of 0.5g/kg/day.

Triglyceride level	Recommended Action
< 3 mmol/l	Continue with lipid infusion as per guideline.
> 3 – 5 mmol/l	• Reduce lipid infusion by ¼ and recheck TG after 24 hours
> 5 – 8 mmol/l	• Reduce lipid infusion by ½ and recheck TG after 24 hrs
> 8 mmol/l	• Reduce lipid to 1g/kg/day (5ml/kg/day) and recheck TG after 24 hours. If still raised reduce lipid infusion to 2.5ml/kg/day, if pump allows. • Continue to monitor TG and once < 3mmol/l increase lipid to 10ml/kg/day and recheck TG after 24 hours. ○ If this level is > 3mmol/l reduce back to 5ml/kg/day and when the level is < 3mmol increase more cautiously to 7.5ml/kg/day. ○ If level is < 3mmol/l increase lipid to 15ml/kg/day ○ Do not increase lipid further without repeating TG level.

Appendix E: Managing parenteral nutrition and feeds – High and Moderate risk infants



Appendix F: Preterm Infant Feed Volume Tables

a. Starting and Increasing Feeds

i. High Risk Infants (based on increases of 10-20ml/kg/day)

Weight (kg)	Start at (hourly)	Start at (2 hourly)	Increase hourly feed volume by*	Increase 2hourly feed volume by
less than 0.6	N/A	0.5	0.25ml every 24 hours	0.5ml every 24 hours
0.6-0.9	0.5	1	0.5ml every 24 hours	1ml every 24 hours
0.9-1.2	0.75	1.5	0.5ml every 12 hours	1ml every 12 hours
1.2-1.5	1	2	0.5ml every 8 hours	1ml every 8 hours
1.5-1.8	1.25	2.5	0.5ml every 6 hours	1ml every 6 hours
1.8-2	1.5	3	1ml every 12 hours	2ml every 12 hours

*Note that this refers to the actual feed volume based on 1 hourly feeds. Therefore if baby is 2 hourly fed then multiply the amount on this table by 2 to give the increase on the feed volume, if on 3 hourly feeds multiply by 3 and so on.

Appendix G: Donor Breast Milk Guidelines

A mother's own breast milk is always preferable to infant formula or donor breast milk (except in very rare circumstances for example maternal HIV or chemotherapy) and every effort should be made to support the mother in producing milk for her baby. Infants who are at risk of gut complications

Donor breast milk (DBM) is a human body fluid and, as such, carries risks of transmission of infective agents. Donors are screened and the milk is pasteurised to minimise risk. Written consent must be obtained for the use of donor breast milk. Handling, testing and documentation of the milk in the donor milk bank and specialist feed unit is carried out according to Food Standards Agency Guidelines 2007

Indications for use

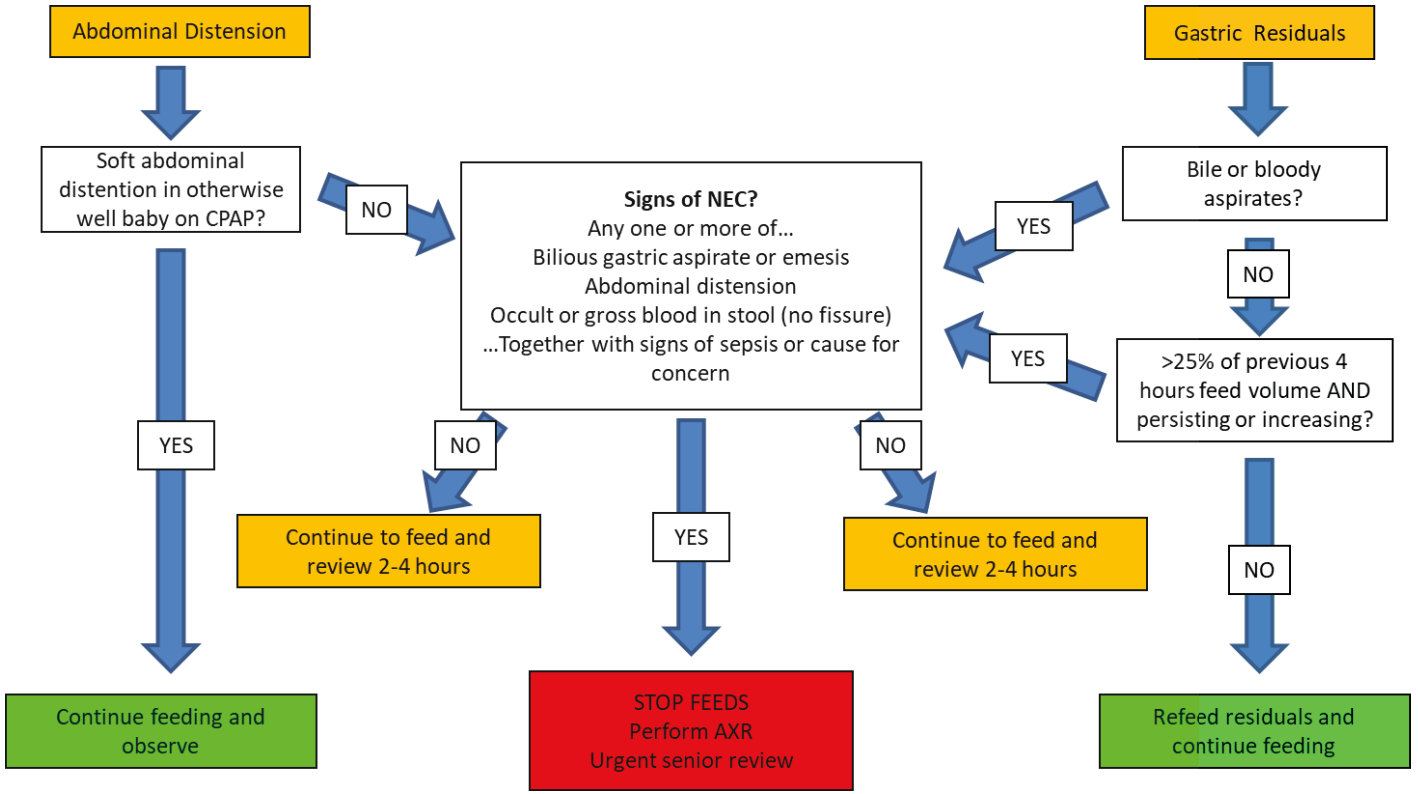
- High or moderate risk infants (as assessed by nutritional screening tool)
- Any infant with surgical pathology
- Any infant with cardiac pathology
- Consultant decision to use

Duration of Use

Although donor breast milk may have reduced nutritional content because of the processing required, breast milk fortifier and/or additional supplementation may ameliorate this and support adequate growth until such a time that the risk of NEC has reduced. Fortified DBM can therefore be continued until 34 weeks when a gradual change to preterm formula can take place. If fortification is contra-indicated (for example, in a baby with a stoma) then DBM can be continued as part of a nutrition plan with parenteral nutrition making up the shortfall. These complex cases will be managed by the nutrition team.

University Hospital Southampton NHS
NHS Foundation Trust

Appendix H: Management of Gastric Residuals/Aspirates and Abdominal Distension



Appendix I: Management of Gastro-Oesophageal Reflux (GOR) in Infants

This guidance is informed by NICE Guideline: Gastro-oesophageal reflux disease in children and young people: diagnosis and management. The PIER guideline for term infants can be accessed here

GOR is a **normal physiological phenomenon.** It is very common and manifests as effortless milky vomiting with no discomfort. It does not affect growth and development and resolves spontaneously. GOR disease (GORD) exists when GOR results in frequent regurgitation with significant distress and at its most extreme leads to faltering growth. It needs to be distinguished from vomiting due to non-GORD disorders such as pyloric stenosis, metabolic causes, infections, and so on. Some conditions are associated with an increased risk of GORD including: congenital diaphragmatic hernia and history of oesophageal atresia. GORD **only rarely** causes apnoea or acute life threatening events. Significant weight loss or faltering of growth in any age group is a red flag and should prompt further evaluation. Rarely, Cow's milk protein allergy may be suspected.

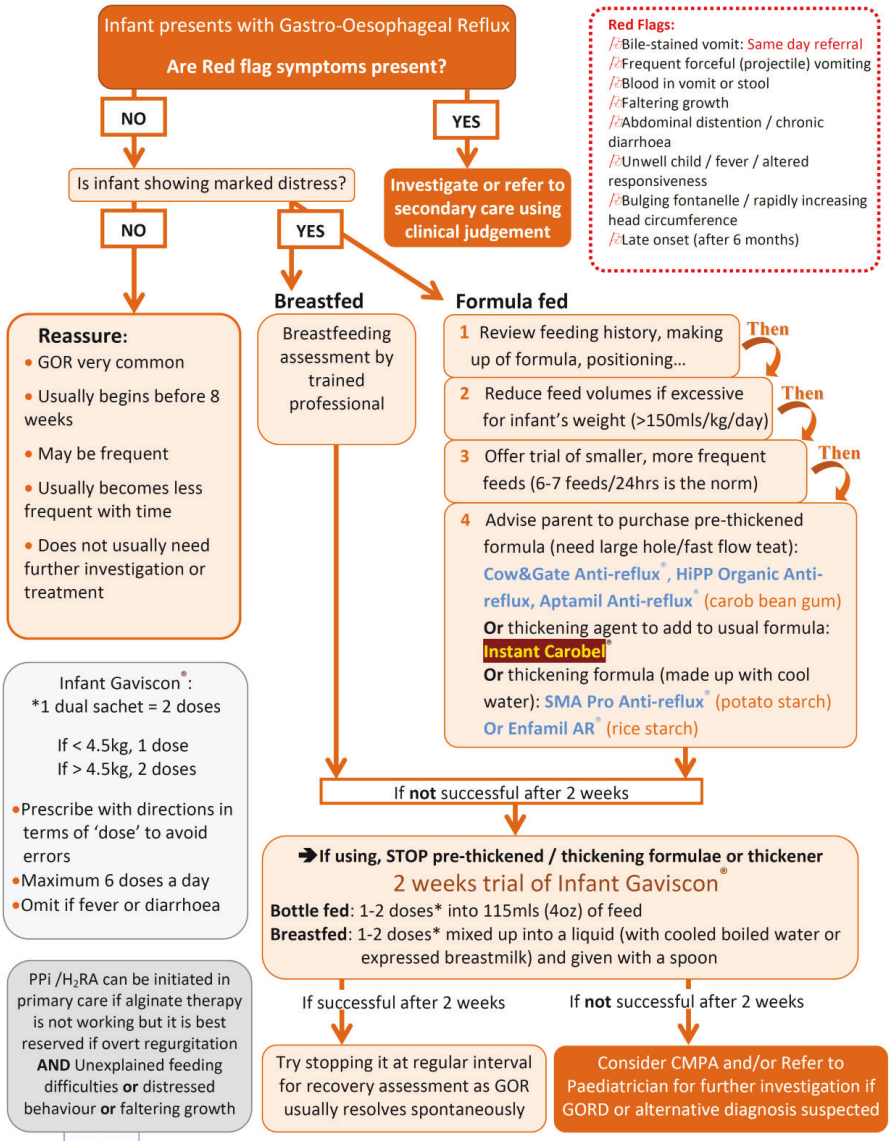
Preterm Babies with suspected GORD:
Preterm babies are at increased risk of GORD and benefit from positioning including elevated side-lying for feeding and raising the head end of the bed to 30° (6). Parental reassurance should be given.

- GORD only rarely causes apnoea or acute life threatening events. For preterm babies where GORD is thought to be having a major adverse impact on the babies clinical condition, continuous NG feeds or nasojejunal feeds may be used.
- Pharmacological treatments used in children and adults (prokinetics, alginates and sucralfate) have **not** shown to be of any benefit in preterm infants and there is evidence of harm (increased risk of NEC and late onset sepsis with the use of PPIs (Omeprazole) and H2 receptor antagonists (Ranitidine)) so these are **contra-indicated** (7, 8).
- PPIs have also been associated with an increased risk of vitamin and mineral deficiencies including calcium, phosphorous, iron and magnesium. For these reasons, if these drugs are used, they should be regularly reviewed.
- A trial of gaviscon can be considered for a short period but should not be used where milk is already fortified, thickened or the baby is at risk of bowel obstruction.
- Feed thickeners such as carobel should **not be used**

Role of PPI/H2RA and jejunal feeding:

- Under the advice of gastroenterology, Ranitidine or Omeprazole can be offered for a period of 4 weeks if one or more of the following are present in term infants without another explanation: unexplained feeding difficulties, distressed behaviour, faltering growth.
- Continuous jejunal feeding is recommended if the baby is at risk of reflux-related pulmonary aspiration.
- The prokinetics **metoclopramide, domperidone or erythromycin should not be offered** in view of the risks of serious side effects. These drugs are contra-indicated in symptoms of neurological origin as they may worsen symptoms.

Term infants (>36 weeks) with suspected GORD (adapted from PIER guidance- see here)



Appendix J: Suspected Cow's Milk Protein Allergy (CMPA)

A link to the PIER guidelines can be found here

Incidence of CPMA in infants is rare - estimated is 2.4% for IgE mediated and 1.7% for non-IgE mediated (9). CMPA can occur in young infants either breast fed or formula fed. Symptoms may include severe regurgitation, vomiting, constipation, peri-anal rash and macroscopic blood in stools, with the latter being most significant (as the others can occur normally in infants) and suggestive of a need to try a cow's milk free feed. Whilst microscopic haematuria is associated with CMPA, it is less sensitive. Note that **urine dipsticks should not be used to test for microscopic blood in stools** (faecal occult blood, FOB) as they are subject to high levels of false positives. An approved bedside test for faecal blood should be used if this is required. Non-intestinal features may include skin rash – atopic eczema, and colic.

If CMPA is thought to be the cause of symptoms, it is recommended that cow's milk protein be excluded from diet:

- If breast feeding, mother should exclude both cows' milk and egg products from her diet for two – four weeks, while continuing to breast feed. Mothers should be encouraged to continue breastfeeding. A parent information leaflet is available here
- Formula fed term infants should be tried on extensively hydrolysed infant formula as per the PIER Wessex Infant Feeding Guidelines. : NB: Amino acid infant formula is **not** the first line treatment.
- Preterm infants - hydrolysed Nutriprem 1 should be used in preterm infants in the first instance, moving to extensively hydrolysed infant formula (e.g. Nutramigen 1 or Peptijunior) only if symptoms persist and are felt to be due to CMPA. Amino acid formula is nutritionally inadequate for these patients.

If improvement is seen, national guidance is that a staged reintroduction should be carried out after 2-4 weeks in order to formally make the diagnosis of CMPA. However, in high risk or nutritionally compromised infants, the need for and timing of this can be done on an outpatient basis. If no improvement is seen on definite exclusion diet, CMPA is unlikely. If exclusion diet is difficult to maintain, a trial of extensively hydrolysed formula may be appropriate for breast fed infants, but where possible breastfeeding should be continued as this is associated with lower risk of food allergy particularly when introducing complementary food. See review by Venter C et al. (10, 11)

Note that any infant on a non-standard formula should have a paediatric dietician involved in their management.

Appendix K: Nutritional Management of Surgical Neonates

The neonatal unit cares for babies with a wide range of surgical pathology. For those with additional co-morbidities such as prematurity and congenital cardiac disease, nutritional care can be complex as, while there are well established recommended nutritional intakes, there is a lack of robust data on which to base recommendation about different strategies. However, the overarching aim is to meet the nutritional needs of this group of patients to maximise their growth and neurodevelopmental outcomes and avoid complications associated with malnutrition.

Options for delivering nutritional requirements include parenteral nutrition alone, combination of parenteral nutrition and enteral feed, supplemented breast milk or preterm formula. There is no evidence to support one of these approaches over another, and such decisions should be made in a multidisciplinary setting as part of the weekly nutrition ward round. Specialist prescription formula may be required. These babies will be reviewed weekly during a combined surgical/gastro/nutrition team MDT meeting.

Nutrient requirements
It is important to note that these infants have the same nutritional requirements as other infants of the same gestation, so have the same targets laid out above on page 6.
It should be noted that these are just recommendations, and some infants may require more of certain nutrients such as sodium and potassium as dictated by the results of blood and urine tests.

Parenteral nutrition around the time of surgery:
Term infants, have a brief hypermetabolic response, which peaks at 6 hours post operatively(12). Compared to healthy term controls, infants undergoing surgery do not show increased energy expenditure, so do not need additional calories. Administering excess calories in this group may therefore lead to excess fat deposition (13). More recently the PEPaNIC study (14, 15) has shown that infants in paediatric intensive care who receive PN during the immediate post-operative period, are more at risk of morbidity such as prolonged ventilation and risk of infection. Conversely, in preterm infants the metabolic responses to surgery are attenuated. As preterm infants have limited reserves, evidence suggests they may benefit from having full nutrition support continued throughout the perioperative period. We therefore recommend the following for infants undergoing surgery:

- Term infants>3.5kg without risk factors should not routinely be started on PN following major surgery before day 5 (this is based on data from the PEPaNIC study). 'Major surgery' is defined as that which is likely to induce a significant systemic inflammatory response post-operatively (and would for example, not include a routine gastroschisis closure).
- Preterm infants, smaller term infants (<3.5kg) and term infants who have undergone minor surgery or where it is felt stopping PN is contraindicated, should continue to receive full parenteral nutrition (16-19), during the perioperative period. If post-operatively there are concerns regarding hyperglycaemia then a 6 – 12 hour reduction by 25 - 50% of the previous PN could be considered.

Appendix L: Nutritional Management of Cardiac Infants

Congenital Heart Disease (CHD) – nutritional requirements

Growth failure in CHD is likely to be multi-factorial arising from low levels of growth hormone and other growth factors, undefined genetic polymorphisms, insufficient nutrition support, feeding difficulties including gastro-oesophageal reflux disease, vocal cord palsy and feeding aversions (20-22).

1. **Nutritional requirements;** pre-surgical requirements are estimated at being 10% higher than otherwise healthy infants. The amount of energy a term infant requires is dependent on the type of cardiac lesion and ranges 110 – 130kcal/kg day, but occasionally up to 150kcal/kg is required to support growth. Although additional calories and protein are often prescribed it is common for infants not to achieve feeding targets either due to feeding difficulties or fluid restrictions leading to growth failure. Post-surgery catch up growth may be required. To achieved 10g/kg per day 128kcal/kg and 2.8g/ kg of protein is required until weight/ length gain goals are met (23). It is important to ensure micronutrient and electrolytes are given in sufficient amounts to support lean body mass acquisition and growth (4) and as such vitamin supplementation should be given to all CHD infants. Medication particularly diuretics reduces total body stores, particularly sodium, affecting growth and as such supplementation may be required (22, 24).
2. **Breast milk is best;** where possible breast milk should be used. It is usual for most infants with heart failure to be on some fluid restriction and combination feeding of fortified breast milk and energy/nutrient dense formula feeds may help promote growth (24, 25). Where there are feeding issues, such as reflux, a ready to feed extensively hydrolysed protein/energy/nutrient dense feed may be better tolerated. Where possible amino acid infant feeds should not be used as these are associated with poor weight gain and hypophosphatemia(26).
3. **Start nutrition support early;** in those infants with CHD lesion requiring pulmonary artery banding or staged surgical repair, early nutrition support in the form of nutrient-energy dense feeds and breastmilk is important in order to prevent growth faltering. Please refer to the pre-surgical CHD pathway (here) or contact a cardiac dietitian to provide an appropriate nutrition care plan e.g. A, B or C (27)
4. **Refer ALL infants with CHD (excluding neonatal PDA ligations) to the Paediatric Cardiac Dietitians**; to ensure they are included in the nutrition home monitoring program before discharge and that the Cardiac dietitian is able to meet the parents(27).
5. **Methods of feed administration;** where infants are not able to breast feed or finish bottles easily the following can be tried;
   - Offer smaller oral feeds more frequently i.e. 2 – 3 hourly
   - Put the remainder of an oral feed via an nasogastric tube (NGT) – always try to give some feed orally to maintain oral feeding skills
   - Provide an overnight feed via an enteral feeding pump and small day time boluses orally

If weight gain remains poor then give feeds continuously over 20 – 22 hours via an enteral feeding pump, if vomiting is an issue then consider feeding via a nasojejunal tube (24).

Appendix M: Nutritional management of neonates/ infants with liver dysfunction (28-40)

<table>
<tr><td>Background</td><td colspan="3">• Liver dysfunction occurs as an acute episodes arising from a previously undiagnosed disease e.g. inborn error of metabolism (IMD) or as a result of prematurity or medical management.
• Jaundice results in dark urine and pale stools; occurs when there is an increase total bilirubin of which more than 20% is conjugated (normal = 5 %).
• Associated with this is decreased fat emulsification and digestion resulting in; malabsorption of fat, fat-soluble vitamins and some minerals.
• In severe cases of liver disease steatorrhea may occur leading to growth failure and metabolic bone disease.
• The use of medium chain triglyceride (MCT) containing infant formula in infants with liver dysfunction should be considered as MCT do not require emulsification. Supplementation with fat soluble vitamins including vitamin K is essential.</td></tr>
<tr><td>Nutrition requirements</td><td colspan="3">Depending on gestational age and previous growth trajectory:
• Energy 90 – 130kcal/kg
• Protein 2.5 – 4g/kg
• Fat should not be restricted and a percentage of the total fat as MCT’s e.g. 30-50 % may be required</td></tr>
<tr><td>Management of fat soluble vitamin deficiency</td><td colspan="3"><table>
<tr><th>Vitamin A</th><th>Vitamin D</th><th>Vitamin E</th><th>Vitamin K</th></tr>
<tr><td colspan="4">Laboratory marker of micronutrient status</td></tr>
<tr><td>• Serum retinol</td><td>• Serum 25-OHD</td><td>• Serum tocopherol</td><td>• Prothrombin/INR</td></tr>
<tr><td colspan="4">Recommendations for supplementation based on serum levels</td></tr>
<tr><td>• 1000IU/kg/day up to 25,000IU water miscible formulation.
• <10kg start with 5,000IU/day (0.6ml OD Dalavit)
• >10kg start with 10,000IU/day</td><td>• Cholecalciferol
• Serum levels – provide the following IU/day
• <10ng/ml:5000
• 11-19ng/ml:4000
• 20-29ng/ml:3000</td><td>• α-tocopheryl acetate 10mg/kg/day</td><td>• Give 1mg vitamin K OD and monitor INR</td></tr>
</table></td></tr>
<tr><td>Questions</td><td colspan="2">Decision guide</td><td>Nutrition recommendations</td></tr>
<tr><td>Is galactosaemia suspected/ diagnosed?</td><td>• No – and the infants is growing well</td><td>• Continue with breastmilk/ standard infant formula breastmilk (& fortifier for infants <1.8kg at birth) and Dalavit 0.6ml</td></tr>
<tr><td></td><td>• Yes – and the infant is growing well - change to a lactose free infant milk</td><td>• Stop standard infant formula/ breastmilk (& fortifier)
• Change infant feed to Wysoy
• Provide Dalavit 0.6ml</td></tr>
<tr><td></td><td>• Yes – but the infant is growth faltering – provide an MCT containing lactose free feed until the liver dysfunction has resolved</td><td>• Stop standard infant formula/ breastmilk (& fortifier)
• Start Alfamino (43% MCT) (38-40) and start Ketovite Liquid and Ketovite tablets (Both must be used together to provide all vitamin requirements)</td></tr>
<tr><td>For prolonged jaundice is the infant breastfeed?</td><td>• No – and the infant is growing well</td><td>• Continue with usual preterm infant formula or standard term formula & Dalivit 0.6ml, Sytron 1ml</td></tr>
<tr><td></td><td>• No – but the infant is growth faltering provide</td><td>• Change to Pepti Junior -provide Dalivit</td></tr>
</table>

	150-180 ml/kg MCT containing infant milk	0.6ml, Sytron 1ml. Consider vitamin K – if deranged INR. Monitor ALP if preterm
	• **Yes – and** the infant is growing well aim to provide 150-180 ml/kg	• Continue with **breastmilk/ fortified breastmilk** and provide Dalavit 0.6ml/ (Sytron 1ml – if pre-term)
	• **Yes – but the** infant is growth faltering aim to provide 75-100 ml/kg of a MCT containing infant formula and 75-100 ml/kg of breastmilk	• Provide 50% **breastmilk/ fortified breastmilk** and **50% Pepti Junior** & Dalavit 0.6ml/ Sytron 1ml. Monitor ALP if preterm
Is cholestasis resolving?	• **No – but the** infant is growing well- continue current mix of MCT containing feed/ and breastmilk	• Continue with **50% Pepti Junior / 50% breastmilk** at 150 – 180ml/kg or age appropriate infant formula & Dalivit 0.6ml/ Sytron 1ml
	• **No – and** infant is continuing to have growth faltering – change to nutrient energy dense feed	• Change to nutrient energy dense infant feed - **Infatrini Peptisorb** (90 - 100ml/kg) and **breastmilk** (30 – 50ml/kg) breastmilk or infant formula & change to Ketovite Liquid and Ketovite tablets (Both must be used together to provide all vitamin requirements)
	• **Yes** - **resume** all breastmilk/ fortified breastmilk or standard infant formula	• **Breastmilk / preterm/ term formula 150ml – 180ml/kg** & change to Dalavit 0.6ml/ Sytron 1ml
Are blood glucose levels maintained?	• **Yes -** continue with current nutrition plan	• Breastmilk or preterm/standard infant formula
	• **No -** aim to provide a nutrient energy dense MCT containing feed with 10 g per 100 ml of carbohydrate and 10 % energy from protein	• Infatrini/ Infatrini Peptisorb
	• **If this does not** resolve the issue glucose polymers may be used to increase this further in increments of 12 g per 100 ml in infants < 6 months and 15 g per 100 ml in infants > 6 months	• Ensure protein: energy ratio is maintained (10 – 12%) and sufficient fat soluble vitamins/ micronutrients are provided
Is there faltering growth with an adequate feed volume?	• **No -** continue with current nutritional feed plan	• Infant's usual feed
	• **Yes** - measure trace elements & vitamins and supplement where necessary	• Review Zinc, copper, selenium, vitamin A, D, and E levels – supplement if inadequate serum levels

Appendix N: Nutritional Management of metabolic bone disease of prematurity

Considerations for the prevention of metabolic bone disease of prematurity (MBDP)

- MBDP is characterised by under mineralisation of the skeleton of **preterm**
- **Before stating supplementation: review contributory factors to aberrant serum mineral status.**

Infants with raised alkaline phosphatase (ALP) >500, low phosphate(PO_4) <1.8

1. Check whether there any contributory factors for abnormal biochemistry?
   - Drugs (especially diuretics, steroids)- stop or rationalise if able
   - Prolonged PN > 28 days
2. Ensure milk feeds are being tolerated and absorbed
3. Ensure milk feeds are providing adequate amounts of calcium / phosphorus and in the correct ratio and address any under or over- provision before further tests

	Magnesium (mmol)	Calcium (mmol)	Phosphate (mmol)	Ca:PO4 ratio
Requirements				
Preterm VLBW <1500g (enteral)	0.3 – 0.6	3 - 5	1.9 – 4.5	1.5 -1.7 :1
Growing preterm <1500g (PN)	0.2 - 0.3	1.6 – 3.5	1.6 – 3.5	1.3 : 1
Content in feed				
Fully fortified MBM (100ml/kg)	0.3	2.2	1.7	1.9: 1
Nutriprem 1 (100ml/kg)	0.3	2.4	2	1.2:1
Preterm conc PN (100ml/kg)	0.21	0.99	2.49	1 : 2.5

4. **Measure vit D** (25OH-D) **if <50** provide 1,000 IU vit D e.g.Abidec 0.6ml=400iu, cholecalciferol, 1 drop=200iu

If vitamin D level is > 50 & biochemical abnormalities persist measure parathyroid hormone levels (PTH)

PTH raised: likely calcium deficiency

1. Start calcium supplements
   - Calcium carbonate (Calvive 1 tablet =25mmol)
   - Stop PO_4 supplements, review Ca:P ratio
2. Infants at particular risk are those on diuretics / steroids or unfortified breastmilk or term formula will have increased electrolyte losses
   - Fortified with breastmilk fortifier –which will ensure Ca:P ratio is 1.5:1 – 1.7:1
3. If no breastmilk is available (including donor), ensure preterm infant formula is provide in adequate amounts e.g. NP1/Hydrolysed NP1
4. Where infants only on PN optimise Ca:P ratio

PTH normal: likely PO_4 deficiency

1. Start PO_4 supplements
   - Start with 1mmol/kg/day (in divided doses unless surgical patient, in which case add to 24hr feed volume), and titrate according to response
   - Give either Sodium glycerophosphate 1mmol/ml of phosphate and 2mmol/ml of sodium or
   - Phosphate Sandoz – 1 tablet = 16mmol
2. Infants at particular risk are those on diuretics / steroids or unfortified breastmilk or term formula will have increased electrolyte losses
3. Fortified with breastmilk fortifier –which will ensure Ca:P ratio is 1.5:1 – 1.7:1
4. Where infants only on PN optimise Ca:P ratio
5. Provide calcium supplements in ionised calcium levels/ Ca:P ratio is low - continue vitamin D

Re-assess Ca, PO4, ALP and PTH in 2 weeks

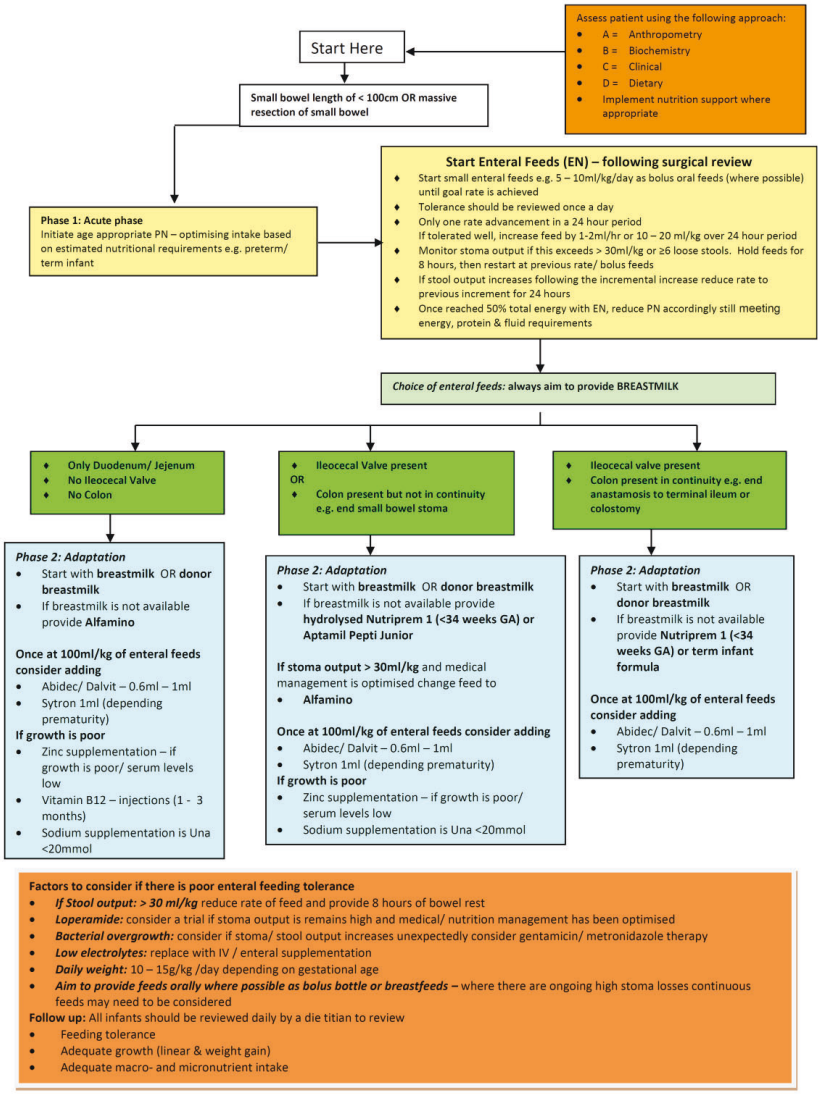
Figure 1: Flow chart of recommended management of metabolic bone disease of prematurity

References: 1.Chinoy A, Mughal MZ, Padidela R. Metabolic bone disease of prematurity: causes, recognition, prevention, treatment and long-term consequences. Archives of disease in childhood Fetal and neonatal edition. 2019. 2. Vuralli D Clinical Approach to Hypocalcemia in Newborn Period and Infancy: Who Should Be Treated? International Journal of Pediatrics 2019. ESPGHAN Parenteral Nutrition 2020, Agostino ESPGHAN 2010, NICE PN 2020

Appendix O – Nutritional management of infants with short bowel syndrome

Short bowel syndrome (SBS) is a global malabsorption syndrome that results from massive resections. In SBS there is loss of absorption function, inability to secrete adequate amounts of gastrointestinal (GI) regulatory peptides, trophic hormones and loss of GI immune function, which is most severe when there, is resection of the ileocecal valve and colon.

The aim of medical and nutritional management is promote adaptation of the remaining bowel. The time period for each infant to achieve adaptation varies depending on numerous factors. A multidisciplinary approach is required to ensure good growth is maintained and the adaptation phase is successfully managed.



Nutritional Care of Infants in the Neonatal Unit Guideline

Version: 3

Document Monitoring Information	
Approval Committee:	Women & Newborn Clinical Governance Steering Group
Date of Approval:	05/02/2021
Ratification Committee:	W&N Governance Steering Group
Date of Ratification:	05/02/2021
Signature of ratifying Committee Group/Chair:	Ash Monga – Chair Women and Newborn Governance Steering Group
Lead Name and Job Title of originator/author or responsible committee/individual:	Dr Mark Johnson, Consultant Neonatologist Dr Freya Pearson, Consultant Neonatologist Dr Luise Marino, Paediatric Dietitian Dr Akshay Batra, Consultant Paediatric Gastroenterologist Mr Nigel Hall, Associate Professor of Paediatric Surgery Suzannah Hibberd, Neonatal Pharmacist
Policy Monitoring (Section 6) Completion Date:	
Policy Monitoring to be presented to responsible committee or PRAMG:	
Target audience:	Staff working on the neonatal unit or caring for neonates.
Key words:	Neonatal, Nutrition
Main areas affected:	Neonatal unit, PICU
Summary of most recent changes if applicable:	Changes to flow charts, feed increments, target nutrient intakes and use of fortifier
Consultation:	Neonatal Consultant group; paediatric dieticians, Paediatric Surgical Team; Paediatric pharmacists.
Equality Impact Assessment completion date:	N/A
Number of pages:	36
Type of document:	Guideline, level 2
Does this document replace or revise an existing document	Revises previous version from 2016
Should this document be made available on the public website?	No
Is this document to be published in any other format?	No

The Trust strives to ensure equality of opportunity for all, both as a major employer and as a provider of health care. This document has therefore been equality impact assessed to

Appendix 6

Protocol for the Growth Assessment of Preterm Infants (GAP) Study, version 8

SOUTHAMPTON Children's Hospital

University Hospital Southampton NHS
NHS Foundation Trust

The Growth Assessment of Preterm Infants (GAP) study

Assessing the growth of preterm infants in Southampton using detailed Anthropometry

Version 7 DATE- 19/07/2019

SPONSOR: University Hospital Southampton NHS Foundation Trust

COORDINATING CENTRE: University Hospital Southampton NHS Foundation Trust- Neonatal Unit

Sponsor's protocol number RHM: CHI0726
Ethics reference no: 14/SC/1275

Protocol authorised by Chief Investigator: Dr R Mark Beattie

Name: **Role:**

Signature: **Date:**

CHIEF INVESTIGATOR AND MEDICAL EXPERT

Contact Details: **Professor R Mark Beattie**
Paediatric Gastroenterology Department
Southampton Children's Hospital
Southampton General Hospital
Southampton SO16 6YD Email: Mark.beattie@uhs.nhs.uk

CO-INVESTIGATORS

Dr Mark Johnson, Prof Sarah Ennis, Dr Freya Pearson, Dr James Ashton, Dr Aneurin Young

PROJECT COORDINATION CENTRE

For general project and clinical queries e.g. participant queries, project supplies, data collection, please contact in the first instance:

Clinical Project Coordinators (joint):

Dr Mark Johnson, Consultant Neonatologist
Neonatal Unit
Princess Anne Hospital
Coxford Road
Southampton
SO16 5YA Email: M.Johnson@soton.ac.uk

Dr Freya Pearson, Consultant Neonatologist,
Neonatal Unit
Princess Anne Hospital
Coxford Road
Southampton
SO16 5YA Email: Freya.Pearson@uhs.nhs.uk

STUDY SITE

Neonatal Unit, Princess Anne Hospital, Coxford Road, SO16 5YA

SPONSOR

University Hospital Southampton NHS Foundation Trust is the research sponsor for this project. For further information regarding sponsorship conditions, please contact the Director of Research and Development at:

Address: R&D Department
University Hospitals NHS Trust
SGH - Level E, Laboratory & Pathology Block,
SCBR - MP 138
Southampton
SO16 6YD
United Kingdom Tel: 023 8079 4989
Fax: 023 8079 8678
Web: www.uhs.nhs.uk

CLINICAL LABORATORIES CONTACT DETAILS:
WELLCOME TRUST CLINCAL RESEARCH FACILITY (WTCRF), SOUTHAMPTON
DUTHIE LABORATORY (NIKKI GRAHAM)- GENETICS LAB

STATISTICIAN CONTACT DETAILS:
HELEN MOYSES

NURSING CONTACT DETAILS:
JENNY POND
Neonatal Unit
Princess Anne Hospital
Coxford Road
Southampton
SO16 5YA

INFRASTRCUTURE SUPPORT DETAILS:
SOUTHAMPTON NIHR BIOMEDICAL RESEARCH CENTRE (NUTRITION THEME)

OTHER MEDICAL AND/OR TECHNICAL DEPARTMENTS AND/OR INSTITUTIONS INVOLVED IN THE TRIAL: (IF APPLICABLE)
NOT APPLICABLE

FUNDER
Funded by NIHR Biomedical Research Centre Southampton (Nutrition theme)

Protocol Information
This protocol describes the above project and provides information about procedures for entering study participants. The protocol should not be used as a guide for the treatment of other patients; every care was taken in its drafting, but corrections or amendments may be necessary. These will be circulated to investigators in the project, but please contact the NIHR Southampton Biomedical Research Centre clinical project coordinator to confirm you have the most recent version.

Compliance
This project will adhere to the principles outlined in the International Conference on Harmonisation Good Clinical Practice (ICH GCP) guidelines. It is subject to the University Hospital Southampton NHS Foundation Trust R&D approvals and will be conducted in compliance with the protocol, the Data Protection Act and all other regulatory requirements, as appropriate.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

A list of the abbreviations, and lists and definitions of specialized or unusual terms or measurement units used in the report should be provided. Abbreviated terms should be spelled out and the abbreviation indicated in parentheses at first appearance in the text

AE	Adverse Event
AR	Adverse Reaction
BRU	Biomedical Research Unit
CRF	Case Report Form
CTA	Clinical Project Authorisation
CTCAE	Common Terminology Criteria for Adverse Events
DMEC	Data Monitoring and Ethics Committee
GA	Gestational Age
GCP	Good Clinical Practice
HSG	Health and Safety Guidance
ICH GCP	International Conference on Harmonisation of Good Clinical Practice
MHRA	Medicines and Healthcare products Regulatory Authority
MRI	Magnetic Resonance Imaging
NHS	National Health Service
PMG	Project Management Group
PSC	Project Steering Committee
REC	Research Ethics Committee
SAE	Serious Adverse Event
SDS	Standard Deviation Scores
SOP	Standard Operating Procedure

List of Definitions

MUAC	Mid upper arm circumference
MTC	Mid thigh circumference
NICU	Neonatal intensive care unit
TBW	Total body water
D2O	Deuterium water

Keywords-

Neonates, Growth, Anthropometry, Nutrition

PROJECT SYNOPSIS

A brief synopsis (usually limited to three pages) that summarises the study should be provided. The synopsis should include numerical data to illustrate results, not just text or p-values.

Full title:	The Growth Assessment of Preterm Infants (GAP) study
Sponsor:	University Hospitals Southampton R&D Department
Sponsor Ref No:	RHM: CHI0726
Chief Investigator	Dr R Mark Beattie
Study phase if not mentioned in title	
Funder:	BRC Nutrition Southampton
Eudract No:	N/A
REC No:	14/SC/1275
Project Type:	Observational Cohort Study
Primary Objective:	Demonstrate the relationship between growth/age and neonatal anthropometry (mid-upper arm circumference (MUAC)/mid-thigh circumference (MTC)) in preterm infants below 32 weeks gestation.
Secondary Objective:	Demonstrate the relationship between nutritional intake and anthropometry in these infants. Understand the clinical, nutritional, genomic and metabolomic factors that influence the growth, body composition and nutritional tolerance in very preterm infants
Rationale:	1. Measures of Mid Upper Arm Circumference (MUAC) and Mid-Thigh Circumference (MTC) in preterm infants will change over time during NICU stay, and these changes will be related to the nutrition these infants receive 2. Changes in MUAC and MTC may occur in the absence of changes in body weight, head circumference and length. 3. Explore whether detailed anthropometry may act as a proxy for body composition measurements 4. Explore whether there is a temporal relationship between the composition of nutrition support, particularly protein energy ratio, and changes in lean body mass accretion, as measured by MUAC and MTC. 5. To inform further study, assessing feasibility and practicality of these methods

	6. Validate serial measurements of MUAC/MTC with serial measurements of body composition in a subgroup of 100 patients 7. Establish the genomic and metabolomic influences on glucose tolerance in early life in preterm infants and subsequent impact on growth. The specific hypothesis is that gestational age (GA) matched preterm infants with differing glucose tolerance during the first 2 weeks of life will have genomic or metabolomic differences relating to glucose metabolism, and different growth outcomes. 8. Establish a clinical, nutritional, genomic and metabolomic data repository for preterm infants 9. Establish the feasibility of an advanced MRI acquisition protocol to collect structural, functional and spectroscopic data in preterm infants to explore relationships between nutritional intake, growth and MRI markers of brain development.
Project Design:	Prospective observational study
Inclusion Criteria:	Preterm infants below 32 weeks gestation (to include up to 31+6)
Exclusion Criteria:	Transfers in with surgical problems. Genetic Syndromes known to impact on growth.
Total No. of Sites:	**1**
Study Duration	8 years (ongoing)
Data collection	Via standardised proformas and weekly measurement of MAUC and MTC, alongside routine anthropometry of length, weight and head circumference. Measurement of body composition by administration of deuterium labelled water and analysis of urine samples by mass spectrometry. Blood will be taken during routine sampling and stored for genetic analysis. Blood and urine samples will be collected at regular intervals from birth and stored for analysis of metabolism (metabolomics) and protein (transcriptomics). MRI brain scans will be carried out on a subgroup infants born <28 weeks upon reaching term equivalent age
Biological samples	Urine for body composition analysis using deuterium dilution (subgroup of patients providing consent) Blood and urine samples for genomics, metabolomics and transcriptomics (subgroup of patients providing consent)
Number of Participants	700

Primary endpoint	Target recruitment reached, including >100 infants with body composition measures using deuterium dilution
Secondary endpoint	Completion of data entry and analysis, publication
Statistical Methods:	Normal plots and the Shapiro-Wilk test Mixed methods linear regression

1. INTRODUCTION

1.1. BACKGROUND INFORMATION

The ex-utero growth of preterm infants is known to fall behind that of their term counterparts, rendering them at risk of further morbidity and mortality[1]. Furthermore the body composition of these infants is significantly altered when compared to those born at term[2]. The distribution of lean and fat mass seen in preterm infants at term equivalent age suggests relative adiposity due to a failure to accrete lean mass during their stay on the neonatal intensive care unit (NICU). The aetiology for this altered body composition, together with the identification of the best way to nutritionally support these infants is an important area of research.

Measurement of weight (together with head circumference and length) is the most commonly used parameter to assess growth on the NICU but as body weight may be influenced by fluid shifts and excess adiposity, current measurement options on the neonatal unit are unable to truly reflect changes in body composition[3]. There is a need to develop measures which better assess the quality of growth in terms of fat and lean mass, in order to better understand how to best deliver nutrition to these vulnerable infants. Several different anthropometric measurements have been used to assess body composition in (mainly term) babies but have not been validated in preterm infants [4-7].

A basic premise of good nutritional support is the understanding of normality, expressed in term infants and children via WHO growth charts [8, 9]. Knowledge of "what we are aiming for" allows tailoring of nutritional support in term infants and children. Without knowing how preterm infants should grow clinical judgements regarding nutritional support (including calorific, fat and protein intake) are difficult to make [10]. Previous studies have used ultrasound to assess body composition

and tissue accretion to focus nutrition but this is not a practical bedside method in most neonatal units [11].

Bedside anthropometry such as Mid Upper Arm Circumference (MUAC) and Mid-Thigh Circumference (MTC) may provide a better way of determining the patterns of growth seen in these infants during NICU stay, and act as an important guide to determining how best to manage their nutrition; promoting a balance of lean mass gain alongside adipose tissue, as is seen in normal term infants[2]. Such anthropometry has been shown to provide information on the relative amounts of fat and lean tissue in term infants and thus may prove useful for preterm infants too. It may also be the case that changes in MUAC and MTC may be seen in the absence of changes in weight, providing insight into growth in terms of its *quality*, rather than *quantity* alone.

Exact measurements of body composition is difficult to conduct in preterm infants. Several methods have been used including bioelectrical impedance (BI), DEXA and air-displacement plethysmography (ADP) with variable results. Besides the limitations of these methods (measurements derived by assumptions) they would be impractical in unwell extreme preterm infants as DEXA and ADP would involve physical movement of the infant. BI has not been shown to be an accurate measure of body composition in previous study of infants. The 'gold standard' of body composition measurement is deuterium-labelled water dilution, with subsequent measurement of the concentration in bodily fluid (plasma, urine etc.). Whilst this method still works from several assumptions it is the most direct measure of total body water (and therefore fat-free mass) available and has a long track record of safety and validation. It would also be a practical approach in extreme preterm infants as it involves administration of a small volume of D20 parenterally or enterally and subsequent plasma/urine sample analysis using mass spectrometry. Deuterium-labelled water is also known as heavy water. This special type of water which we can measure in urine by using a mass spectrometer. It is not

radioactive and it is not toxic. Deuterium-labelled water (D_2O) has been widely used for over 50 years in research including in term babies without any adverse events, protocols for its use in body composition analysis in infants and children have been published by international bodies[12]. Deuterium-water exists in small quantities in normal water from the tap. It also exists normally in small quantities in the human body. When we give it, it is excreted rapidly in urine like normal water and will not build up in the body.

In this study we will validate the novel anthropometric measures of MUAC and MTC against measures of body composition using deuterium dilution.

Furthermore, in early life these infants can have impaired glucose and lipid tolerance, becoming hyperglycaemic and hypertriglyceridaemic despite recommended intakes. It is likely this is multi-factorial, with evidence for continued gluconeogenesis despite adequate glucose delivery and a degree of insulin resistance[13]. This is important because it necessitates the use of insulin or a reduction in PN glucose or lipid content. A need to maintain an energy:protein ratio in PN of 20-27kcal/g protein means that this leads to a subsequent reduction in protein content, which in turn impacts on growth[14]. Reductions in glucose intake, suboptimal energy:protein ratios and insulin use will all have an impact on growth, body composition and subsequent metabolic and developmental outcomes[15]. Whilst some data suggest hyperglycaemia in preterm infants is associated with lower GA, lower birthweight and mode of feeding[15], it is unclear why preterm infants of similar birthweight and GA have differing responses to nutrition, with some exhibiting hyperglycaemia and others not. Variations in GLP-1, insulin receptor β and Akt-1, have been associated with hyperglycaemia in preterm infants and so genomic differences in these proteins may explain the disparity of glucose responses seen[16, 17].

Nutrition and growth are known to influence neurodevelopment in preterm infants. Novel Magentci Resonance Imaging (MRI) techniques, including Magnetic Resonance (MR) spectroscopy, are providing new insights into structural, functional and biochemical abnormalities of the developing

brain. These techniques have been successfully implemented in Southampton for the assessment of term infants suffering from neonatal encephalopathy. We wish to assess whether these approaches can be applied to preterm infants at term equivalent age and to assess for correlations between growth, nutrition and MRI markers of neurodevelopment.

We wish to better understand the clinical, nutritional, genomic and metabolomic factors that influence the growth, body composition and nutritional tolerance in very preterm infants. This will allow us to identify preterm infants at particular risk of nutrition intolerance, poor growth and excess adiposity, and over time enable personalised medicine, including the use of specific nutritional strategies, to help improve their growth and long-term outcomes.

This study will be conducted in compliance with the protocol, GCP and the applicable regulatory requirements(s).

1.2. STUDY SCHEDULE

MONTHS	0-3	4-6	6-12	12-18	18-24	24-36	36-48	48-60	60-72	72-84	84-96
Preparation, Ethics	X	X									
Recruitment, Data Collection and Assessment		X	X	X	X	X	X	X	X	X	X
Analysis of body composition in a subgroup of patients							X	X	X	X	X
Genomics, metabolomics and transcriptomics in a subgroup of patients								X	X	X	X
Data Entry and Analysis			X	X	X	X	X	X	X	X	X

2. STUDY OBJECTIVES AND PURPOSE

2.1. AIM

1. Demonstrate the relationship between growth/age and neonatal anthropometry (mid-upper arm circumference (MUAC)/mid thigh circumference (MTC)) in preterm infants below 32 weeks gestation.
2. Demonstrate the relationship between nutritional intake, anthropometry and changes in body composition in these infants.
3. Demonstrate the relationship between MUAC/MTC and measurement of body composition by D_20 dilution in a subgroup of patients and validate detailed anthropometry (weight, length, head circumference, MUAC and MTC against body composition measures using deuterium dilution (DD) and develop a bedside method for estimating body composition
4. Establish the genomic and metabolomic influences on glucose tolerance in early life in preterm infants and subsequent impact on growth. The specific hypothesis is that GA matched preterm infants with differing glucose tolerance during the first 2 weeks of life may have genomic or metabolomic differences relating to glucose metabolism that impact growth outcomes.
5. To confirm the feasibility of an advanced MRI acquisition protocol to collect structural, functional and spectroscopic data in preterm infants to explore relationships between nutritional intake, growth and MRI markers of brain development.
6. Establish a clinical, nutritional, genomic and metabolomic data repository for preterm infants.

2.2. STUDY DESIGN

This is an observational prospective cohort study. There is no intervention

1. Infants eligible for the study will be identified by the clinical research team. Parents will be given and information leaflet and then consent sought after they have had time to read the information.

2. Recruitment of babies to occur within 1 week of arrival to neonatal unit, at the first available opportunity after arrival. Initial recruitment proforma to be completed, alongside 1st measurement proforma.
3. Anthropometric follow-up measurements will occur weekly thereafter. Where possible this will be timed to occur at 7 day intervals from the date of the first measurement, but as a minimum will occur within each seven day period of an infant's stay.
4. Length of follow-up to be dictated by length of stay on Neonatal Unit. Initial follow-up measurements will be for the duration of admission, though we will seek consent to contact parents when children are older for future follow up studies at a later date (subject to separate funding and approvals)
5. Measurements to be recorded on standardised proforma and entered into an anonymised database. Original proformas containing identifiable information will be held securely on site.
6. Measurements will not be performed if clinical staff deem the infants too unwell to be measured, in the same way as is currently done for routine measures of head circumference, length and weight.
7. Nutrient intake- nutrient intake data will be collected from routine records (casenotes and charts) for the duration of admission. This will be done prospectively where possible, but can also be done retrospectively. Volumes of fluids and feeds received each day will be entered onto a specially designed electronic tool (the Southampton Electronic Neonatal Nutrition Assessment Tool, SENNAT), which then calculates the delivery of most major macro- and micronutrients.
8. For a subgroup of 100 patients who give explicit consent, we will perform deuterium labelled water body composition analysis. These measurements will be serial during the NICU stay and will aim to validate of MUAC/MTC as a measure of body composition. For SOP please see below

9. For patients who give additional consent for blood samples for genetics, and blood and urine for metabolomics and transcriptomics, bloods will be taken during routine sampling and urine will be collected by placing cotton wool balls place in the nappy, with urine extracted using a syringe (see method below). Samples will be processed and stored in the WTCRF until analysis. Blood samples for genetics will be taken once during stay, timed with routine blood sampling. Blood and urine samples for metabolomics and transcriptomics will be taken at regular intervals from birth at opportunistic intervals timed to occur with routine blood sampling (for example, ideally birth, 72 hours, day 7 and day 14 of life, though this will depend on patient stability and appropriate opportunities to collect samples), though not more than twice in any 7 day period.
10. For babies of parents who give additional consent for MRI scanning, an MRI scan will be performed between 38 and 42 weeks corrected gestational age. MRI will be performed during natural sleep at the MRI unit at Southampton General Hospital. The scan will take approximately 45 minutes. If the infant wakes during the scan, imaging will be paused and only resumed if the infant settles back to sleep. During the scan, heart rate and oxygen saturation in the blood will be monitored non-invasively. MR images will be clinically reported by Paediatric Neuroradiologists. Incidental findings will be communicated to families by Dr Vollmer, Consultant Paediatric Neurologist, and the usual clinical pathways will be followed.

Measurements to be taken (Outcome Measures)

- Mid Upper Arm Circumference
- Mid Thigh Circumference
- Head Circumference
- Weight

- Length

Measurements to be conducted in line with the methods formulated and validated by the Southampton Women's Survey (SWS) [18], all measurements to be collected by members of the study team who have had training in the use of these measures in a standardised way, and in accordance with a standard operating procedure (SOP- see methods for details). Every effort will be made to keep the measurer the same throughout the study period with the substitute measurer being used where no choice is available.

- Three measurements to be made (to the nearest millimetre) and recorded; the mean to be used for analysis.
- All measurers will undergo training from the research nurses from the SWS prior to commencement of the study.
- Where possible revalidation of measurers will occur to ensure that there are no significant inter-observer discrepancies.
- All equipment will be validated prior to use.

Subgroup Measures and SOP for body composition by deuterium-labelled water dilution

- Urine D2O concentration (by mass spectrometry)

In order to assess the validity of MUAC/MTC measurements as a proxy for body composition it will be necessary to compare serial MUAC/MTC to a known measure of body composition. In a subgroup of 100 patients we will measure body composition (TBW) by deuterium labelled water dilution, using a protocol based on international recommendations [12].

1. Infants will be recruited into the main study, at the time of recruitment the investigator will discuss recruitment into this extra study group with parents and we will recruit those who express interest. We aim to recruit 100 patients into this study group (including 1 infant from each week of gestation 23-27).

2. Infants will have initial body composition (TBW) measured by deuterium dilution within 7 days of recruitment to the study.
3. Serial measurements of TBW will occur at 3 weekly intervals until discharge (these will be timed to occur at the same times as 3D scanning where infants are consented to both)
4. D20 can be given orally, nasogastrically, orogastrically or intravenously depending on clinician choice (based on infant condition, IV access etc.), though only existing naso/orogastric tube or intravenous access will be used.
5. Baseline infant D20 levels will be measured on the same day of D20 administration, collection of a urine sample will be done at 4 hours (or less) prior to administration of the D2O. This will occur for each separate measurement.
6. The volume of D20 to be administered to the infant will be based on weight (see below). Dilution of small volumes of D2O with normal water or saline will be needed to ensure evaporation loss is minimised.

Patient Weight (kg)	Dose to be administered (35g) (70mg/kg)	Approximate volume to be administered without dilution (based on density of D20 = 1.1056g/ml)
0.5	35	0.0316ml
0.75	52.5	0.0475ml
1.0	70	0.0633ml
1.25	87.5	0.0784ml
1.5	105	0.0949ml

7. A urine sample will need to be collected from T+6 hours to T+ 10 hours after D2O administration. The collection will be via cotton wool balls placed in the nappy in infants with corrected gestational age<32 weeks, or urine collection bags in infants >32weeks corrected gestational age. Any contamination with faeces will result in disposal of that sample.
8. Further urine samples will be collected every morning for the next 7 days (3-4 total body water turnovers).

9. Extraction of urine- Cotton wool balls should be placed into a 50ml syringe and squeezed into a sterile and dry container and sealed. This should then be labelled with patient information, date of sample, time of sample, time after D2O administration and then frozen immediately.
10. Samples will be processed in the Wellcome Trust Clinical Research Facility, Southampton General Hospital, then sent to an approved external contractor or collaborator for analysis.
11. TBW will be calculated by the back extrapolation (intercept) method. Over 7 days all data are plotted onto a graph and the intercept of the curve represents the TBW. The following equation will be used (Nd and TBWd only)-

$$\text{No or Nd (kg)} = ((W \times A/a) \times (\Delta DD/\Delta BW)/ (1000)$$

$$\text{TBWo} = \text{No}/1.007$$

$$\text{TBWd} = \text{Nd}/1.041$$

$$\text{TBW avg} = (\text{TBWd} + \text{TBWo})/2$$

$$\text{TBWav(kg)} = \text{TBWav} \times 1000/18.0153$$

W = amount of water (g) to dilute the dose for IRMS analysis

A = amount of dose (g) given to the subject

a = amount of dose weighed out (g) for the dose dilution

ΔDD = enrichment of diluted dose minus tap water enrichment

ΔBW = enrichment of post-dose sample minus baseline sample enrichment

1000 = transforms TBW in g (mL) to kg (L)

1.007 = correction factor for non-aqueous oxygen exchange

1.041 = correction factor for non-aqueous hydrogen exchange

12. Calculation of FFM and FM can be derived from TBW and weight of patient using the following equations-

$$\text{Weight (kg)} = \text{FM kg} + \text{FFM kg}$$

$$\text{FFM} = \text{TBW}/ \text{Hydration coefficient}$$

$$\text{FM (kg)} = \text{Weight (kg)} - \text{FFM (kg)}$$

$$\text{\% Body fat} = 100 \times \text{FM}/\text{Body weight}$$

The hydration coefficient in infants is significantly different to adults and will be calculated as follows: Hydration coefficient= 0.792 - 0.0028 x age (years)

13. We will then look at these body composition measurements and use these in comparison to changes in MUAC/MTC/other anthropometry.

Subgroup measure and SOP for blood and urine samples for genomics, metabolomics and transcriptomics

Blood

Blood for genomics and transcriptomics (RNA) will be taken opportunistically at regular intervals during routine clinical sampling (for example birth, 72 hours, day 7 and day 14 of life), though no more than two samples will be taken during any 7 day period (genomics needs only be taken once and will be done in place of transcriptomics blood sampling on one occasion). 1ml blood is required for genomics (this only needs to be taken once) and 2ml is required for transcriptomics. Samples will be taken and stored for DNA and RNA analyses.

Urine

Urine samples for metabolomics will be collected regular intervals during stay (for example birth, 72 hours, day 7 and day 14 of life). Urine will be collected via cotton wool balls placed in the nappy in infants with corrected gestational age<32 weeks, or urine collection bags in infants >32weeks corrected gestational age. Any contamination with faeces will result in disposal of that sample. Urine will be extracted from cotton wool balls by placing them into a 50ml syringe and squeezing into a sterile and dry container and sealed. This will then be stored for metabolomic analyses.

Data for NHS pathway participants will be generated and held within the clinical setting in accordance with the standard procedures in place for the clinical diagnostic lab. Anonymised data for these participants (clinical and genomic) will be transferred from the NHS clinical domain to research facilities. For research pathway participants, biological samples taken at recruitment will be stored and processed in appropriately licensed facilities. DNA/RNA will be analysed predominantly using

Next Generation Sequencing (NGS) technologies, or other methods if more appropriate, either locally or with appropriate collaborators/service providers, and the data stored within research facilities. Data and samples will be stored for 10 years following the end of the study. Given that the specific analytical approach will vary dependent upon data available, number of patients/relatives, clinical phenotypes and mode of inheritance, these will not be specifically detailed herein.

Subgroup measures and SOP for MRI Scanning

Parents of infants meeting the additional inclusion criteria specified below will be invited to additionally consent to MRI scanning. Twenty-four infants will be recruited. MRI scan of the brain will be performed during natural sleep at 38-42 weeks corrected gestational age in the MRI unit of Southampton General Hospital. The scan will take approximately 45 minutes. The normal arrangements for clinical MRI scanning will be made, namely:

- Transfer to and return from the MRI scanner will be performed by clinical staff from the neonatal unit
- Natural sleep will be induced by feeding and swaddling (the 'feed and wrap' procedure)
- Non-invasive continuous monitoring of ECG and oxygen saturation will be performed
- Appropriate wrapping will be performed to counteract the cool environment of the MRI room, and ear protection putty will be applied
- Continuous monitoring of the infant shall be performed from the control room of the MRI scanner by a qualified doctor and/or registered nurse by means of audio-visual relay from the scanning room and by monitoring of ECG and saturations
- MRI will be stopped immediately upon the assessment of the attending clinical team in the event of clinical instability.
- Imaging will be paused if the infant awakens and can be resumed once the infant is settled

3T data acquisition will include conventional structural, volumetric (3D) T1-weighted and T2-weighted images for assessment of brain macrostructure/injury and for image registration and segmentation, diffusion-weighted MRI (optimised for high-angular resolution diffusion imaging) and resting-state functional MRI as well as single-voxel 1H MR spectroscopy. MR spectroscopy will examine brain metabolites and associations with nutrition. The imaging protocol is already established at our centre for imaging of term born infants with neonatal hypoxic-ischaemic encephalopathy. Connectivity measures will be derived to enable correlation analyses with the nutrition data of interest.

MR images will be clinically reported by Paediatric Neuroradiologists. Incidental findings will be communicated to families by Dr Vollmer, Consultant Paediatric Neurologist, and the usual clinical pathways will be followed.

Nutritional data to be collected

Table 1 lists the nutrients that are calculated by SENNAT which will be used in this study. The main focus of nutrition data in this study will be around Energy and Protein intake.

Nutrient	Unit	Nutrient	Unit	Nutrient	Unit
Energy	kcal/kg	Zinc	µmol/kg	Vitamin B6	µg/kg
Protein	g/kg	Copper	µmol/kg	Folate	µg/kg
Carbohydrate	g/kg	Selenium	nmol/kg	Vitamin B12	µg/kg
Fat	g/kg	Iodine	nmol/kg	Biotin	µg/kg
Sodium	mmol/kg	Manganese	nmol/kg	Pantothenic Acid	mg/kg
Chloride	mmol/kg	Vitamin A	IU/kg	Niacin	mg/kg
Potassium	mmol/kg	Vitamin D	IU/kg	Vitamin C	mg/kg
Calcium	mmol/kg	Vitamin E	IU/kg	Taurine	mg/kg
Phosphorous	mmol/kg	Vitamin K	µg/kg	Choline	mg/kg
Magnesium	mmol/kg	Thiamin	µg/kg	Carnitine	mg/kg
Iron	µmol/kg	Riboflavin	µg/kg	Inositol	mg/kg

Table 1: The 33 different nutrient intakes calculated and collected by SENNAT

Other data to be collected

- Infants gestational age(GA), sex and birth weight
- Infants CRIB II score (requires GA, sex, birthweight, admission temperature, worse base excess in first 24 hours)[19]
- Days of ventilation
- Presence of Chronic Lung Disease (defined as need for respiratory support or oxygen at 28 days of age)
- Presence of Necrotising enterocolitis (NEC)
- Presence and grade of Intraventricular haemorrhage (IVH)
- Presence of Patent Ductus Arteriousus (PDA) (diagnosed clinically) and any subsequent treatment
- Days of parenteral nutrition
- Maternal smoking and alcohol history
- Maternal pre-pregnancy BMI
- Daily clinical data including minute-to-minute data vital signs, fluids, feeds, and nutrition delivered to the infant each day, routine clinical blood tests (liver function, bone profile and micronutrient status) and all blood gas data, including routine blood sugars will be extracted from the neonatal unit clinical information system (MetaVision).

Equipment required

- Blank Tape Measures
- Metal Rulers
- Neonatal Incubator Measure (measures up to 45cm)
- Infantometer/Infant length measurer (measures infants >45cm)
- Electronic Scales

- Sterile urine bottle
- 50 ml syringe
- Centrifuge (SCBR)
- Freezer -80° (WTCRF)

2.3. DEFINITION OF END OF PROJECT

The primary data collection will be complete when the final participant is discharged from hospital.

3. SELECTION AND WITHDRAWAL OF PARTICIPANTS

3.1. SUBJECT SELECTION

Sites of patient recruitment- Southampton- PAH

- Pre-term infants across both sites
- Infants below 32 weeks gestation (to include up to 31+6)

3.2. SUBJECT INCLUSION CRITERIA

Infants below 32 weeks gestation (to include up to 31+6) presenting to PAH neonatal unit.

Subgroup Inclusion Criteria for MRI scanning:

Infant born below 28 weeks gestation and whose parents are local to the Southampton area (to facilitate scanning at term equivalent age when non-local babies would have been transferred closer to home)

3.3. SUBJECT EXCLUSION CRITERIA

- Transfers in with surgical problems.
- Genetic Syndromes known to impact on growth.

Subgroup Exclusion Criteria for MRI scanning:

Focal ischaemic or haemorrhagic preterm brain injury on routine clinical neonatal cranial ultrasound.

3.4. SUBJECT RECRUITMENT

- Eligible infants will be identified by members of the study team (all of whom are also members of the direct clinical care team)
- Recruitment of babies to occur within 1 week of arrival to neonatal unit; at the first available opportunity after arrival on the unit.
- Consecutive recruitment of babies under the age of 32 weeks gestation.
- No payment to participants
- Parents have the option to opt out of recruitment at any stage
- Patient information given and consent obtained by a member of the study team on arrival to unit or at nearest point within 7 days after arrival.
- Initial recruitment proforma to be completed, alongside 1st measurement proforma.
- All patients approached will be asked to take part in all elements of the study, but will be given the option to opt out of any or all of the 3D Scanning element, the deuterium dilution element, and the collection of blood and urine for genomics, metabolomics and transcriptomics.

3.5. SUBJECT COMPLIANCE

Not applicable for this study

3.6. WITHDRAWAL OF SUBJECTS

Parents can withdraw their child at any point in the study. At this point the anthropometry will cease although routine clinical measurements will continue. Parents will be able to opt out of these measurements being used in the study, along with previously collected measurements being used. If parents wish all data will be destroyed for their child.

4. STATISTICS AND DATA

- At recruitment an initial proforma will be completed for each patient including demographics and basic details (see appendix 1). First measurement proforma to be completed at this point (see appendix 2).
- Measurements to be recorded on standardised measurement proforma (see appendix 2).and entered into two anonymised databases (one for each site).
- Original proformas can identify patients, to be held securely on site.
- There will be an interim data analysis, full data will be analysed after all collection has occurred.
- Descriptive statistics will be used to summarise the demographic and outcome variables of the infants in all the study periods.
- Variables of interest (listed as outcome measures above) will be tested for normality in order to help determine the nature of the analysis methods used. Normal plots and the Shapiro-Wilk test will be used, with a cut off value of $p<0.05$ accepted as evidence of a non-normal distribution. It is anticipated that nutrient intakes are likely to be negatively skewed with a non-normal distribution, whilst growth parameters are likely to take on a normal distribution.
- For continuous variables, if the data are normally distributed, the mean and standard deviation will be calculated. If the data are not normally distributed, the median and interquartile range will be calculated. For categorical or binary variables, these will be summarised as frequency and percentage of total.
- Growth parameters will be plotted over time, and a mixed methods linear regression approach, taking into account the repeated measures nature of the data, will be used to model their changes over time, and in relation to nutrient intakes and other factors including GA, sex, birth weight, CRIB II score and the presence of CLD.

- Where possible anthropometric measures will be converted into Z-scores. A cut off of <-2 is indicative of moderate malnutrition, as defined by the World Health Organisation.
- Body composition measurements will be analysed using the equations described above. A multiple regression analysis will be performed to ascertain the relationship between our simple anthropometric measurements and the body composition by TBW.
- Integrated clinical and research data (including high density longitudinal digital clinical data, growth and body composition data alongside genomic (DNA), transcriptomic (RNA) and metabolomics data) will be stored on contemporary encrypted SQL database platforms (e.g. tranSMART, BC|GENOME) and hosted on a secure servers within a Microsoft Data Centre in South United Kingdom. Data held in the servers are backed-up weekly using the snapshot technology provided by Microsoft Azure with a retention period of two years. The web servers of these platforms are only accessible via encrypted HTTPS protocol with username/password based authentication and are only accessible from UK-based IP ranges. Access to study data is enabled only for relevant study team members that have signed the delegation logs.
- Brain connectivity measures derived from the resting state and diffusion MRI will be derived and correlation analyses with the nutrition data of interest will be performed. Similarly, quantitative measures from MR spectroscopy (brain metabolite measures) will be correlated with the nutrition data.
- Complex statistical and related analytical techniques will be used to understand the relationships between growth, nutrition, body composition, genomics, transcriptomic and metabolism. These will include, but are not limited to, multilevel regression modelling, mixed methods linear regression (taking into account the repeated measures nature of the data) and machine learning techniques (these will enabling complex modelling of the interactions

between nutrition, genome, metabolome, growth and body composition, determining the most favourable conditions for optimal growth and outcome).

5. SAFETY ASSESSMENTS

We do not predict there to be any adverse events associated with anthropometric measurement of these infants.

In the subgroup of infants entered in the body composition element of the study there is the requirement to administer D2O, which has a track record of over 50 years of safe administration. Where administered IV, we will ensure this is compliant with full 'Aseptic Non-Touch Technique' (ANTT). Any extravasation injury will be dealt with in line with local policy. There is no increased risk of injury associated with D2O above that associated with water. Any adverse events associated with urine sampling will be dealt with in line with local policy.

In the event that any harm, or potential harm, comes to an infant, (due to measurement or any part of the study) this will be dealt with immediately by the clinical team on site.

Parents will be informed as soon as possible of any adverse events related to this study.

6. STOPPING / DISCONTINUATION RULES

If the study has to be discontinued then the data collected will continue to be stored in line with data handling as outlined below. We will destroy all data after the allotted time period has expired.

All participants (parents) will be informed if the study has to be discontinued.

7. STUDY GOVERNANCE

7.1. ETHICAL CONSIDERATION

This study will be submitted to an NHS ethic committee via Proportionate Review Service (PRS) via the integrated research application system (IRAS). As measures as non-invasive and minimally disruptive, and can be done alongside routine care, we plan to include all eligible infants born during

the study period, until the recruitment target is reached. Consent will be obtained within 1 week of admission to the neonatal unit, prior to measurements being taken. Nutritional information will be collected from routine records and anonymised prior to analysis. Measurement data will also be link-anonymised. Measurement of MUAC and MTC are less invasive and disruptive for infants than the routine measures of length, weight and head circumference. Weight requires the infant to be undressed and lifted onto scales, length requires the infant to be undressed and held in a stretched out position, and head circumference require any hat to be removed, which can be disruptive given that many infants have their respiratory support such as CPAP or endotracheal tube for ventilation secured using special hats. Conversely, MUAC and MTC can be measured simply by exposing the relevant body part, most of which are often exposed in preterm infants in incubators anyway. Blood will be taken during routine sampling to minimise disruption and samples stored in accordance with appropriate regulations. The project will be conducted in accordance with the recommendations for physicians involved in research on human participants adopted by the 18th World Medical Assembly, Helsinki 1964 as revised and recognised by governing laws and EU Directives; and the principles of GCP and in accordance with all applicable regulatory requirements including but not limited to the Research Governance Framework and the Medicines for Human Use (Clinical Trial) Regulations 2004, as amended in 2006 and any subsequent amendments. The Chief Investigator will submit a final report at conclusion of the trial to the REC within the timelines defined in the Regulations.

7.2. INSURANCE AND INDEMNITY

The sponsor of the project is University Hospital Southampton NHS Foundation Trust. For NHS sponsored research HSG (96) 48 reference no.2 refers. If there is negligent harm during the clinical project when the NHS body owes a duty of care to the person harmed, NHS Indemnity covers NHS staff, medical academic staff with honorary contracts, and those conducting the project. NHS Indemnity does not offer no-fault compensation and is unable to agree in advance to pay compensation for non-negligent harm. Ex-gratia payments may be considered in the case of a claim.

7.3. SPONSOR

University Hospital Southampton NHS Foundation Trust is acting as Sponsor for this project. We will ensure that the study is compliant with submissions to regulatory authorities and GCP. Other delegated duties will be assigned to the NHS Trusts or others taking part in this project by means of the site clinical project agreement where appropriate.

7.4. FUNDING

Internal 2 year grant from the biomedical research centre for nutrition (BRC nutrition) at Southampton University Hospitals, with subsequent 2 year additional project funding and a 3 year clinical doctoral research fellowship for Dr Aneurin Young.

7.5. MONITORING, AUDITS AND INSPECTIONS

This study will be monitored and may be participant to inspection and audit by University Hospital Southampton NHS Foundation Trust, under their remit as sponsor, the BRU as the Sponsor's delegate and other regulatory bodies to ensure adherence to ICH GCP, Research Governance Framework for Health and Social Care, applicable contracts/agreements and national regulations. All study related documents will be made available on request for monitoring and audit by UHS, the relevant REC or other licensing bodies.

8. DATA HANDLING AND RECORD KEEPING

Study documents (paper and electronic) will be collected and retained in accordance with the Data Protection Act 1998 in a secure location during and after the trial has finished. All essential documents including source documents will be retained for a minimum period of 5 years following the end of the study. It is expected that the BRC (Southampton) will aid with said management.

9. PUBLICATION POLICY

9.1. BACKGROUND

When submitting items for publication it is important for authors to acknowledge the organisations involved in the research. This acknowledgement assists with building the reputation of organisations making it more attractive to external bodies that are in a position to place future business and research funding with departments and individuals within Southampton.

The policy is provided to ensure authors identify the University of Southampton and University Hospital Southampton NHS Foundation Trust in a way that will increase the ability of search engines, routinely used by funders of research (e.g. by the RAND association and National Institute of Health Research), to identify publications from the Southampton partnership.

9.2. SCOPE

The policy applies to all staff and students whose research outputs from pre-clinical and clinical research derive from their employment by the University and/or Trust, from research grants awarded to the University and/or Trust or otherwise from the use of University and/or Trust resources and facilities. The policy applies to all authors of publications, and not simply to principal authors or reprint authors. Citing both organisations on all papers covered by this policy acknowledges the success of each organisation resulting from working in partnership.

9.3. POLICY

When submitting items for publication the Trust and University of Southampton require that affiliations and acknowledgements are included in the formats provided below.

Where any author has a substantive or honorary contract with the University of Southampton and/or the University Hospital Southampton NHS Foundation Trust both organisations must appear as affiliations in ALL submissions for publication.

9.4. REQUIRED FORMAT FOR AFFLIATIONS

9.4.1. AUTHOR'S AFFILIATION

University of Southampton and University Hospital Southampton NHS Foundation Trust

9.4.2. ADDITIONAL AFFILIATIONS

The University Faculty should be listed as the second affiliation. Further affiliations of all authors should be listed using the appropriate form of words as given below:

- University of Southampton Faculty of Medicine (change Faculty as appropriate)
- NIHR Southampton Biomedical Research Centre (BRC)
- Southampton Centre for Biomedical Research (SCBR)
- NIHR Southampton Experimental Cancer Medicine Centre
- Southampton University Cancer Research UK Centre

9.4.3. ACKNOWLEDGMENTS

We will ensure acknowledgement of:

- Any sponsorship you have received for the research you are publishing. For grants, include the grant number and source.
- Where resources (staff, space or equipment) of a particular unit or department have been used e.g. NIHR Welcome Trust Clinical Research Facility, NIHR Southampton Biomedical Research Centre, NIHR Respiratory Biomedical Research Unit, is used in the course of the project, the unit / department must appear in the acknowledgements section. For example:
  - 'This study was supported by the Southampton NIHR Wellcome Trust Clinical Research Facility'.
- Names of individual staff that have supported the project that led to the publication as well as the relevant department using an appropriate title.
  - Southampton Faculty of Medicine Bioimaging Unit
  - UHS FT Department of Infection
- NIHR academic training programmes.
  - Supported AF2s, ACFs and CLs should have funding acknowledged as "PERSON X (or initials) was supported by the University of Southampton National Institute of Health

Research Academic Foundation Programme/Academic Clinical Fellowship Scheme

(delete as applicable) or NIHR Clinical Lectureship"

9.5. OPEN ACCESS

The Department of Health and National Institute for Health Research has an Open Access Policy which affects all researchers who are funded (partly or wholly) by them. Resulting research papers that have been accepted for publication in a peer-reviewed journal should be deposited to UK PubMed Central (UKPMC) within six months. For projects that are funded by DH/NIHR, Open Access costs should be budgeted for by the researcher. The Wellcome Trust, however, provides grant holders with additional funding to cover open access charges.

10. SUPPLEMENTS

Please see below for the following documents:

1. Neonatal Nutrition Data Collection Proforma- Initial Recruitment Proforma
2. Neonatal Nutrition Data Collection Proforma- Follow-up measurements

11. PROJECT MANAGEMENT

Dr M Johnson and Dr A Young are responsible for the day-to-day running and progress of this study. Dr RM Beattie will oversee the progress and will act as the medical expert for this study. Dr F Pearson will be on site in Southampton respectively to ensure day to day running of the project is in line with the protocol.

12. REFERENCES

1. Wood, N.S., et al., *The EPICure study: growth and associated problems in children born at 25 weeks of gestational age or less.* Arch Dis Child Fetal Neonatal Ed, 2003. **88**(6): p. F492-500.
2. Johnson, M.J., et al., *Preterm birth and body composition at term equivalent age: a systematic review and meta-analysis.* Pediatrics, 2012. **130**(3): p. e640-9.
3. Yajnik, C.S., et al., *Neonatal anthropometry: the thin-fat Indian baby. The Pune Maternal Nutrition Study.* Int J Obes Relat Metab Disord, 2003. **27**(2): p. 173-80.
4. Catalano, P.M., et al., *Anthropometric estimation of neonatal body composition.* Am J Obstet Gynecol, 1995. **173**(4): p. 1176-81.
5. Ahmed, F.U., E. Karim, and S.N. Bhuiyan, *Mid-arm circumference at birth as predictor of low birth weight and neonatal mortality.* J Biosoc Sci, 2000. **32**(4): p. 487-93.
6. Sharma, J.N., et al., *Mid-arm circumference at birth as a predictor of low birth weight babies and early neonatal mortality.* Indian Pediatr, 1986. **23**(11): p. 915-9.

7. Ramaiya, C., et al., *Newborn's arm circumference as a screening tool of low birth weight in Temeke District, Dar es Salaam, Tanzania.* Trop Geogr Med, 1994. **46**(5): p. 318-21.
8. WHO. http://www.who.int/childgrowth/software/en/. 2014 [cited 2014 14/1/2014].
9. de Onis, M., et al., *[WHO growth standards for infants and young children].* Arch Pediatr, 2009. **16**(1): p. 47-53.
10. Moyses, H.E., et al., *Early parenteral nutrition and growth outcomes in preterm infants: a systematic review and meta-analysis.* Am J Clin Nutr, 2013. **97**(4): p. 816-26.
11. McLeod, G., et al., *Feasibility of using ultrasound to measure preterm body composition and to assess macronutrient influences on tissue accretion rates.* Early Hum Dev, 2013. **89**(8): p. 577-82.
12. *Assessment of body composition and total energy expenditure in humans using stable isotope techniques*. 2009, Vienna: International Atomic Energy Agency.
13. Chacko, S.K., et al., *Gluconeogenesis is not regulated by either glucose or insulin in extremely low birth weight infants receiving total parenteral nutrition.* J Pediatr, 2011. **158**(6): p. 891-6.
14. Koletzko, B., B. Poindexter, and R. Uauy, *Nutritional care of preterm infants : scientific basis and practical guidelines*. World Review of Nutrition and Dietetics, ed. B. Koletzko, B. Poindexter, and R. Uauy. Vol. 110. 2014, Basel: S. Karger AG.
15. Ahmad, A., et al., *Indices of Glucose Homeostasis in Cord Blood in Term and Preterm Newborns.* J Clin Res Pediatr Endocrinol, 2016. **8**(3): p. 270-5.
16. Blanco, C.L., et al., *The ontogeny of insulin signaling in the preterm baboon model.* Endocrinology, 2010. **151**(5): p. 1990-7.
17. Salis, E.R., et al., *Insulin resistance, glucagon-like peptide-1 and factors influencing glucose homeostasis in neonates.* Arch Dis Child Fetal Neonatal Ed, 2017. **102**(2): p. F162-F166.
18. Godfrey, K.M., et al., *Relation of cord plasma concentrations of proinsulin, 32-33 split proinsulin, insulin and C-peptide to placental weight and the baby's size and proportions at birth.* Early Hum Dev, 1996. **46**(1-2): p. 129-40.
19. Parry, G., J. Tucker, and W. Tarnow-Mordi, *CRIB II: an update of the clinical risk index for babies score.* Lancet, 2003. **361**(9371): p. 1789-91.

Neonatal Nutrition Data Collection Proforma- Initial Recruitment Proforma

Patient details (Label)

Date of recruitment	
Study Number	
Time of Recruitment	
Recruited by	
Date of Birth	
Time of Delivery	
Gestational Age	
Mode of delivery	
Ethnic background	
Resuscitation at birth? What?	
Mother's Height and Weight	**Height.....cm** **Weight.....kg**
Maternal Smoking (circle)	**None _____cigarettes/day**
Maternal Alcohol (circle)	**None _____units/day**
Mother's age	
Father's Height and Weight	**Height.....cm** **Weight.....kg**
Father's age	

Neonatal Nutrition Data Collection Proforma- Follow-up measurements

Patient Study Number.....

Date of Collection.....

Time of Collection.....:.....

Patient details (Label)

Measurement Sequence Number.....

Measurement	Result 1	Result 2	Result 3	Mean	Presence of IV line?
Weight					
Length					
Head Circumference					
LEFT Mid Upper Arm Circumference					
RIGHT Mid Upper Arm Circumference					
LEFT Thigh Circumference					
RIGHT Thigh Circumference					

Appendix 7

Guide to the extraction, processing and formating of data for inclusion on the Southampton Preterm Nutrition Database

Database Update Guide

1. Anonymisation Phase

This phase ingests data from BadgerNet, SENNAT and the GAP database. It then assigns study numbers and processes the data to remove all identifiable information and to remove unrequired fields.

Script Language	R – created using R v4.1.1 in RStudio v1.1.456 in Anaconda Navigator v2.1.0
Main script:	1_anonymisation.R
Packages required:	RODBC, dplyr, readr, lubridate, cellranger, readxl, openxlsx
Data source files:	"Full download" of BadgerNet into an MS Access file, renamed to 'badger.mdb' – script can handle more than one file from BadgerNet as long as the string 'badger' is in the filename. SENNAT Excel file copy renamed to 'sennat.xlsb' and with password protection removed – script can handle more than one SENNAT as long as the string 'sennat' is in the filename. GAP MS Access database copy renamed to 'gap.accdb' (only one file is allowed) GAP_Blood_gases.xlsx as supplied by the data science team – script can handle more than one blood gas file as long as the string 'gas' is in the filename. GAP_Pathology.xlsx as supplied by the data science team – script can handle more than one pathology file as long as the string 'path' is in the filename.
Reference files:	'MASTER.csv' – the file to match unique study number to hospital ID and NHS number 'pcd_lookup.rda' – compressed R file matching postcodes to Lower Layer Super Output Areas 'lsoarank.csv' – CSV file which ranks each LSOA by index of multiple deprivation (IMD) along with the decile 'postcode_corrector.csv' – CSV file which lists faulty postcodes with their correction – can be modified.
Output files:	anon_data.rda – a compressed R file which is a list of lists of data frames containing the data

Notes:

- This phase must be run on UHSFT PCs for data security, as per the study protocol and REC approval. R can be directly installed onto Trust PCs by downloading from the R website (https://www.r-project.org/, no administrator privileges required to install). Anaconda Navigator (https://www.anaconda.com/products/individual) can similarly be downloaded and installed. Anaconda allows the user to run R scripts within RStudio, making it easier to run interactively and see the environment – this may be useful to track down problems if errors arise. RStudio cannot be directly installed on Trust PCs without administrator privileges.

- All data source files and reference files need to be placed in the R working directory.
- The R script will install any required packages if they are not already installed.
- Data for blood tests is extracted from Charts by members of the data science team – I contact Florina Borca – she requires a list of hospital numbers and the fields required. I generally send this to her as an xlsx with a tab of hospital numbers and a tab of required items.
- Postcodes:
    - Postcodes are frequently entered incorrectly into BadgerNet and there is no meaningful internal error checking.
    - BadgerNet tries to look up the Lower layer Super Output Area (LSOA) assigned to each postcode but it often fails, even when the postcode is entered correctly. This script will independently look up the LSOA by the postcode and is much more successful at doing so than the internal BadgerNet process.
    - The R script will output a table to the console showing any postcodes which could not be matched to an LSOA. It will also issue a warning if any invalid postcodes are found.
    - These postcodes can then be corrected by updating the 'postcode_corrector.csv' file.
    - If postcodes are corrected in 'postcode_corrector.csv' file, it should be safe to simply rerun the R script. MASTER.csv is automatically updated by the script and it will not create duplicate study numbers if the same infants are run through the script multiple times.
- There are separate anonymisation and field selection steps. This allows further non-patient identifiable fields to be added in the future if required, whilst keeping anonymisation secure.
- Preparing SENNAT for integration: The SENNAT binary Excel workbook is password protected but there is no facility for R to access password-protected files. Therefore, a copy of SENNAT should be made and placed in the working directory. The password protection should then be removed and this file should then be destroyed after processing in the anonymisation step.

2. Consolidation phase

This phase processes the data from Phase 1 into two data frames containing baby-level and daily data from all sources.

Script Language	R – created using R v4.1.1 in RStudio v1.4.1717
Main script:	2_consolidation.R
Packages required:	Dplyr, doParallel, foreach, tidyr
Data source files:	anon_data.rda from Phase 1
Reference files:	encoder.csv – converts field names to standard BC|GENOME fields fluicodes.csv – converts from untidy fluid names to tidier ones for BC|GENOME
Output files:	consolidated_data.rda – a compressed list of data tables

Notes:

- fluidocdes.csv may need to be updated if new fluids are introduced into the database.
- Nutrient intakes are calculated from fluid volumes. This uses *doParallel* and *foreach* to run these calculations in parallel (by default across n-1 cores, where n is the number of processor cores available). Running across seven cores, this process takes approximately one minute for every hundred infants. Therefore, it may take quite a long time, especially if fewer cores are used as it is very sensitive to parallelisation. Using one thread, it will take around seven minutes for every hundred infants, i.e. >30 minutes for the whole dataset.

3. Encoding

This phase processes the data from Phase 2 into an Excel file with tabs for each of the BC|GENOME tables.

Script Language	R – created using R v4.1.1 in RStudio v1.4.1717
Main script:	3_encoding.R
Packages required:	dplyr, tidyr, stringr, openxlsx
Data source files:	consolidated_data.rda from Phase 2
Reference files:	encoder.csv – converts from SENNAT field names to BC|GENOME field names separator.csv – selects data fields for each of the Excel tabs placecodes.csv – contains a key to convert NHS hospital codes to the level of neonatal unit in that hospital pathcodes.csv – contains reference to convert pathology error text into codes
Output files:	GAP_Database_Update_[today's date].xlsx – an Excel file with tabs for each of the BC|GENOME tables

Notes:

- placecodes.csv may need to be updated if infants are transferred from unusual out-of-network locations.

Appendix 8

Approved protocol for the Re-Growth study

Re-Growth

Reassessing the growth of infants born below 32 weeks' gestation in the UK, 2014-2018

Version 2 DATE- 12/4/2019

SPONSOR: University Hospital Southampton NHS Foundation Trust

COORDINATING CENTRE: University Hospital Southampton NHS Foundation Trust- Neonatal Unit

Sponsor's protocol number
Ethics reference no:

Protocol authorised by Chief Investigator:

Name: Mark Johnson **Role:**

Signature: **Date:**

CHIEF INVESTIGATOR AND MEDICAL EXPERT

Contact Details:
Dr Mark Johnson
Neonatal Unit
Princess Anne Hospital
Coxford Road
Southampton
SO16 5YA

Email: M.Johnson@soton.ac.uk

PROJECT COORDINATION CENTRE

For general project and clinical queries e.g. participant queries, project supplies, data collection, please contact in the first instance:

Clinical Project Coordinator:

Dr Aneurin Young
Neonatal Unit
Princess Anne Hospital
Coxford Road
Southampton
SO16 5YA Email: a.young@soton.ac.uk

STUDY SITES

Contact Details:
Neonatal Unit
Princess Anne Hospital
Coxford Road
SO16 5YA

SPONSOR

University Hospital Southampton NHS Foundation Trust is the research sponsor for this project. For further information regarding sponsorship conditions, please contact the Director of Research and Development at:

Address: R&D Department
University Hospitals NHS Trust
SGH - Level E, Laboratory & Pathology Block,
SCBR - MP 138
Southampton
SO16 6YD
United Kingdom Tel: 023 8079 4989
Fax: 023 8079 8678
Web: www.uhs.nhs.uk

CLINICAL LABORATORIES CONTACT DETAILS:

NOT APPLICABLE

STATISTICIAN CONTACT DETAILS:

HELEN MOYSES, NIHR SOUTHAMPTON BIOMEDICAL RESEARCH CENTRE

NURSING CONTACT DETAILS:

NOT APPLICABLE

INFRASTRUCTURE SUPPORT DETAILS:

NIHR BIOMEDICAL RESEARCH CENTRE SOUTHAMPTON, NUTRITION THEME
SOUTHAMPTON CENTRE FOR BIOMEDICAL RESEARCH
MP 218
TREMONA ROAD
SOUTHAMPTON GENERAL HOSPITAL
SO16 6YD
TELEPHONE: 023 8120 8548
EMAIL: BRC@UHS.NHS.UK

OTHER MEDICAL AND/OR TECHNICAL DEPARTMENTS AND/OR INSTITUTIONS INVOLVED IN THE TRIAL: (IF APPLICABLE)

NEONATAL DATA ANALYSIS UNIT
IMPERIAL COLLEGE LONDON
CHELSEA AND WESTMINSTER HOSPITAL CAMPUS
369 FULHAM ROAD
CHELSEA, LONDON
SW10 9NH
TEL: 020 3315 5841
EMAIL: NDAU@IMPERIAL.AC.UK

FUNDER
NIHR BIOMEDICAL RESEARCH CENTRE SOUTHAMPTON, NUTRITION THEME
(DETAILS AS ABOVE)

Protocol Information

This protocol describes the above project and provides information about procedures for obtaining and analysing data. Corrections or amendments may be necessary. These will be circulated to investigators in the project, but please contact the NIHR Southampton Biomedical Research Centre clinical project coordinator to confirm you have the most recent version.

Compliance

This project will adhere to the principles outlined in the International Conference on Harmonisation Good Clinical Practice (ICH GCP) guidelines. It is subject to research ethics committee approvals and will be conducted in compliance with the protocol, the Data Protection Act and all other regulatory requirements, as appropriate.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

A list of the abbreviations, and lists and definitions of specialized or unusual terms or measurement units used in the report should be provided. Abbreviated terms should be spelled out and the abbreviation indicated in parentheses at first appearance in the text

BRC	Biomedical Research Centre
CGA	Corrected Gestational Age
GA	Gestational Age
LNU	Local Neonatal Unit
NICU	Neonatal Intensive Care Unit
NNRD	National Neonatal Research Database
NDAU	Neonatal Data Analysis Unit
NHS	National Health Service
NHSFT	National Health Service Foundation Trust
NIHR	National Institute for Health Research
RCPCH	Royal College of Paediatrics and Child Health
REC	Research Ethics Committee
SCBU	Special Care Baby Unit
SITAR	SuperImposition by Translation and Rotation

Keywords-

Preterm, Growth, Weight, Head circumference, Length, Nutrition

PROJECT SYNOPSIS

A brief synopsis (usually limited to three pages) that summarises the study should be provided. The synopsis should include numerical data to illustrate results, not just text or p-values.

Full title:	Reassessing the growth of infants born below 32 weeks' gestation in the UK, 2014-2018
Sponsor:	University Hospital Southampton NHSFT
Sponsor Ref No:	
Chief Investigator	Mark Johnson
Study phase if not mentioned in title	
Funder:	NIHR BRC Southampton
Eudract No:	N/A
REC No:	
Project Type:	Study limited to working with data (as defined by IRAS)
Primary Objective:	To reassess the early growth pattern of very preterm infants born in the UK.
Rationale:	1. Early growth of infants born below 32 weeks' gestation has been previously described and has demonstrated that preterm infants routinely lose weight during their first week of life and do not regain their birth centile (or z-score) by the time they reach term corrected age [1]. 2. Recent data suggests that the pattern of growth described in this study is not inevitable and that improved growth can be achieved [2]. 3. A dataset of all preterm births in the UK is recorded on the National Neonatal Research Database (NNRD) held by the Neonatal Data Analysis Unit (NDAU) (managed by Imperial College London for the Royal Society of Paediatrics and Child Health)[3]. This database records data covering many aspects of neonatal care, including a comprehensive record of patient demographics/characteristics, growth, nutritional inputs and longer term outcomes 4. This study will use the NNRD to describe any changes in the pattern of early growth of very preterm infants using similar methods to those described by *Cole et al*.[1] 5. This study will integrate data regarding nutritional care from the NNRD to assess associations between nutrition and growth, and two-year follow-up data to assess associations between early growth and subsequent growth and development.
Project Design:	Database review project
Inclusion Criteria:	1. Infant born in England below 32^{+0} weeks gestational age.

	2. Infant admitted to a neonatal unit in England reporting to the National Neonatal Research Database.
Exclusion Criteria:	1. Consent for inclusion not obtained from reporting neonatal unit.
Total No. of Sites:	1
Study Duration	12 months
Data collection	Data to be supplied by NDAU according to their application processes. Application supplied as Appendix 2.
Biological samples	None
Number of Participants	
Primary endpoint	Data analysis completed.
Secondary endpoint	Publication of data in a peer-reviewed journal.
Statistical Methods:	Core data to be analysed using the superimposition by translation and rotation (SITAR) model as previously described[4]. Additional data analysis will be performed using linear regression with mixed modelling.

1. INTRODUCTION

1.1. BACKGROUND INFORMATION

Preterm infants, particularly those born below 32 weeks gestation, are at risk of early growth failure[5]. This is important because there is evidence that patterns of growth in early life have an impact on the risk of non-communicable disease in later life[6 7], and also some evidence that growth in the neonatal period influences subsequent development in childhood[8]. Their growth pattern has been characterised by significant weight loss within the first week of life, followed by weight gain which fails to result in growth sufficient for them to attain the weight, head circumference and length upon reaching term corrected gestational age (CGA) that they would have achieved if they had continued to follow the growth trajectory on which they were on at the time of their premature birth. This means that, when potted on a UK-Newborn Infant Close Monitoring (NICM) growth chart, they fall down several marked centile lines and then follow a pattern that is slightly divergent from that of the marked centiles lines. Mathematically this is represented by a fall in their standard deviation score (SDS, also called a z-score) for weight, length and head circumference between birth and term CGA compared to reference population data. This pattern was described in a seminal paper more than five years ago[1]. This group used SITAR, a novel statistical method, to assess and describe growth curves which they had developed[4].

There is however continuing uncertainty about the optimal growth pattern of infants born very preterm as it relates to this later growth and development. Since 2011 (the latest year included in Cole's study) there has been ongoing interest in increasing the quality and quantity of nutrition which is received by very preterm infants. Recent work has shown that the previously described fall in weight z-score is not inevitable, using data from one neonatal unit[2]. It is not known whether this is due to specific nutritional practices in that particular neonatal unit, or whether this is the result of

more widespread changes in nutritional care and the resulting growth patterns of very preterm infants across the country since Cole's study. Therefore, it is not known whether this change will be confirmed in other settings

The National Neonatal Research Database (NNRD) comprises a large dataset concerning all infants admitted to neonatal units in the UK. It is held by the Neonatal Data Analysis Unit (NDAU) (part of Imperial College London) on behalf of the Royal College of Paediatrics and Child Health (RCPCH)[3]. This database is approved by a REC and is an approved NHS Information Standard (ISB1575). Access to the data for research purposes requires specific REC approval and consent from each contributing neonatal unit. Among its data fields[9], the NNRD contains data defining the baseline characteristics of infants (e.g. gestation, weight and head circumference at birth) along with longitudinal data describing their stay on the neonatal unit (e.g. weight, length and head circumference measured throughout their admission). It also records data describing their nutritional intake during each day of their admission and their weight and development at two years CGA.

1.2. STUDY SCHEDULE

MONTHS	0-6	6-9	9-12
Ethical approval, approval from all NDAU sites, acquiring data from NNRD	X		
Data Analysis		X	X
Writing Up			X

2. STUDY OBJECTIVES AND PURPOSE

2.1. AIM

1. To characterise the growth patterns of very preterm and extremely preterm infants born in the UK between 2014 and 2018, and to compare this with earlier data which has been previously published[1].
2. To look at regional variation
3. To assess for associations between nutritional practices, type of neonatal unit (NICU/SCU/LNU[10]) and growth and in the neonatal period.
4. To assess for associations between neonatal growth and subsequent growth and development as measured at a CGA of two years.

2.2. STUDY DESIGN

- A required dataset has been drawn up using the list (data dictionary) of all NNRD data fields [11] and is included as Appendix 1. In summary, this will provide anonymised data describing the infant's baseline characteristics, the type of neonatal unit providing care, the region of the country, growth, nutritional intake and growth and developmental status at two years CGA.
- A request for this information will be submitted to NDAU using their application form (Appendix 2).
- Local REC approval is required.
- Each contributing neonatal unit will need to consent to our using their data and this will be facilitated by NDAU (as per item 5 in [12]).

Primary Data

- The primary data analysis will be a description of the growth pattern of very preterm and extremely preterm infants using the SITAR method. This will require baseline characteristics of sex and GA at birth, and longitudinal measurements of weight during the neonatal care episode. Longitudinal growth curves for each gestational age group (by weeks of completed gestation age at birth) will be generated and overlaid on the existing NICM chart for comparison. These analyses will be compared with those described for 2006-2011[1].

Secondary Data

- Regression analysis will be performed to explore associations between nutritional care, type of neonatal unit and growth pattern.
- Regression analysis will be performed to assess for associations between neonatal growth and later growth and development as measured at two years CGA.

2.3. DEFINITION OF END OF PROJECT

The primary data collection will be complete when statistical analysis has been completed.

3. SELECTION AND WITHDRAWAL OF PARTICIPANTS

SUBJECT SELECTION

Subjects will consist of all infants delivered below 32 weeks GA and cared for on a neonatal unit in England in the period 2014-2018 (inclusive, 5 years total). Subjects will be excluded if the relevant neonatal unit does not consent to their inclusion. We anticipate approximately 5000 subjects.

3. STATISTICS AND DATA

- The SITAR method has been previously described[4] and used in this context[1]. It will be used to describe the growth patterns of infants stratified by GA at birth. This analysis is carried out in the R programming language.
- Standard mixed linear regression models will be applied to assess for associations (either using R or the Stata statistical analysis package).

4. SAFETY ASSESSMENTS

This study will use anonymised retrospective descriptive data and will not pose a safety risk to subjects.

5. STOPPING / DISCONTINUATION RULES

It is not anticipated that stopping or discontinuation will be required for this descriptive study which uses historical prospectively-collected data.

6. STUDY GOVERNANCE

6.1. ETHICAL CONSIDERATION

This study will not include any intervention. It uses anonymised prospectively collected descriptive data. NDAU will supply data in an anonymised form. In order to assess the type of neonatal unit providing care to the infant, we will be supplied with data which identifies the neonatal unit (NNRD/NDAU does not collect this information directly but it can be derived by cross-referencing with other sources). Data will be analysed stratified by type of neonatal unit and region but not by individual neonatal unit and individual units will not be identified in publications arising from this study. None of

the information supplied will constitute patient-identifiable information as defined by the Caldicott Report[13].

The Chief Investigator will submit a final report at conclusion of the trial to the REC within the timelines defined in the Regulations.

6.2. INSURANCE AND INDEMNITY

The sponsor of the project is University Hospital Southampton NHS Foundation Trust (UHS). For NHS sponsored research HSG (96) 48 reference no.2 refers. If there is negligent harm during the clinical project when the NHS body owes a duty of care to the person harmed, NHS Indemnity covers NHS staff, medical academic staff with honorary contracts, and those conducting the project. NHS Indemnity does not offer no-fault compensation and is unable to agree in advance to pay compensation for non-negligent harm. Ex-gratia payments may be considered in the case of a claim.

6.3. SPONSOR

University Hospital Southampton NHS Foundation Trust is acting as Sponsor for this project. We will ensure that the study is compliant with submissions to regulatory authorities and GCP. Other delegated duties will be assigned to the NHS Trusts or others taking part in this project by means of the site clinical project agreement where appropriate.

6.4. FUNDING

No dedicated source of funding is sought for this project. Aneurin Young is funded as a research fellow by NIHR BRC Southampton.

6.5. MONITORING, AUDITS AND INSPECTIONS

This study will be monitored and may be participant to inspection and audit by UHS, under their remit as sponsor, the BRU as the Sponsor's delegate and other regulatory bodies to ensure adherence to ICH GCP, Research Governance Framework for Health and Social Care, applicable contracts/agreements and national regulations. All study related documents will be made available on request for monitoring and audit by UHS, the relevant REC or other licensing bodies.

7. DATA HANDLING AND RECORD KEEPING

Study documents (paper and electronic) will be collected and retained in accordance with the Data Protection Act 1998 in a secure location during and after the trial has finished. All essential documents will be retained for a minimum period of 5 years following the end of the study.

8. PUBLICATION POLICY

8.1. BACKGROUND

When submitting items for publication it is important for authors to acknowledge the organisations involved in the research. This acknowledgement assists with building the reputation of organisations making it more attractive to external bodies that are in a position to place future business and research funding with departments and individuals within Southampton.

The policy is provided to ensure authors identify the University of Southampton and University Hospital Southampton NHS Foundation Trust in a way that will increase the ability of search engines, routinely used by funders of research (e.g. by the RAND association and National Institute of Health Research), to identify publications from the Southampton partnership.

8.2. SCOPE

The policy applies to all staff and students whose research outputs from pre-clinical and clinical research derive from their employment by the University and/or Trust, from research grants awarded to the University and/or Trust or otherwise from the use of University and/or Trust resources and facilities. The policy applies to all authors of publications, and not simply to principal authors or reprint authors. Citing both organisations on all papers covered by this policy acknowledges the success of each organisation resulting from working in partnership.

8.3. POLICY

When submitting items for publication the Trust and University of Southampton require that affiliations and acknowledgements are included in the formats provided below.

Where any author has a substantive or honorary contract with the University of Southampton and/or the University Hospital Southampton NHS Foundation Trust both organisations must appear as affiliations in ALL submissions for publication.

8.4. REQUIRED FORMAT FOR AFFLIATIONS

8.4.1. AUTHOR'S AFFILIATION

National Institute for Health Research Biomedical Research Centre Southampton, University Hospital Southampton NHS Foundation Trust and University of Southampton

8.4.2. ADDITIONAL AFFILIATIONS

The University Faculty should be listed as the second affiliation where appropriate. Further affiliations of all authors should be listed using the appropriate form of words as given below:

- University of Southampton Faculty of Medicine (change Faculty as appropriate)

- Southampton Centre for Biomedical Research (SCBR)

8.4.3. ACKNOWLEDGMENTS

We will ensure acknowledgement of:

- Any sponsorship you have received for the research you are publishing. For grants, include the grant number and source.
- Names of individual staff that have supported the project that led to the publication as well as the relevant department using an appropriate title.
- The Neonatal Data Analysis Unit

8.5. OPEN ACCESS

The Department of Health and National Institute for Health Research has an Open Access Policy which affects all researchers who are funded (partly or wholly) by them. Resulting research papers that have been accepted for publication in a peer-reviewed journal should be deposited to UK PubMed Central (UKPMC) within six months.

For projects that are funded by DH/NIHR, Open Access costs should be budgeted for by the researcher. The Wellcome Trust, however, provides grant holders with additional funding to cover open access charges.

9. SUPPLEMENTS

Appendix 1 – Required items from the NNRD

Appendix 2 – Request form for NDAU for access to NNRD data

10. PROJECT MANAGEMENT

Dr M Johnson and Dr A Young are responsible for the day to day running and progress of this study. Prof RM Beattie will oversee progress and will act as the medical expert.

11. REFERENCES

1. Cole TJ, Statnikov Y, Santhakumaran S, et al. Birth weight and longitudinal growth in infants born below 32 weeks' gestation: a UK population study. *Arch Dis Child Fetal Neonatal Ed* 2014;99(1):F34-40. doi: 10.1136/archdischild-2012-303536 [published Online First: 2013/08/13]
2. Andrews ET, Ashton JJ, Pearson F, et al. Early postnatal growth failure in preterm infants is not inevitable. *Arch Dis Child Fetal Neonatal Ed* 2018 doi: 10.1136/archdischild-2018-315082 [published Online First: 2018/08/24]
3. Neonatal Data Analysis Unit. Neonatal Data Analysis Unit: Imperial College London; 2019 [Available from: https://www.imperial.ac.uk/neonatal-data-analysis-unit/about-us/ accessed January 2019.
4. Cole TJ, Donaldson MD, Ben-Shlomo Y. SITAR--a useful instrument for growth curve analysis. *Int J Epidemiol* 2010;39(6):1558-66. doi: 10.1093/ije/dyq115 [published Online First: 2010/07/22]
5. Lee SM, Kim N, Namgung R, et al. Prediction of Postnatal Growth Failure among Very Low Birth Weight Infants. *Sci Rep* 2018;8(1):3729. doi: 10.1038/s41598-018-21647-9 [published Online First: 2018/03/01]
6. Barker DJ. The malnourished baby and infant. *British medical bulletin* 2001;60:69-88. [published Online First: 2002/01/26]
7. Barker DJ, Osmond C. Infant mortality, childhood nutrition, and ischaemic heart disease in England and Wales. *Lancet (London, England)* 1986;1(8489):1077-81. [published Online First: 1986/05/10]
8. Shah PS, Wong KY, Merko S, et al. Postnatal growth failure in preterm infants: ascertainment and relation to long-term outcome. *J Perinat Med* 2006;34(6):484-9. doi: 10.1515/JPM.2006.094 [published Online First: 2006/12/05]
9. Neonatal Data Analysis Unit. National Neonatal Dataset. In: London IC, ed. London, 2018.
10. Department of Health. Toolkit for High-Quality Neonatal Services, Section 3.8. London: Department of Health, 2009.
11. Utilising the NNRD. 2018. https://www.datadictionary.nhs.uk/data_dictionary/messages/clinical_data_sets/data_sets/national_neonatal_data_set/national_neonatal_data_set_-_episodic_and_daily_care_fr.asp?shownav=0 (accessed December 19).
12. Neonatal Data Analysis Unit. Utilising the NNRD. 2019
13. The Caldicott Committee. Report on the Review of Patient-Identifiable Information London: Department of Health, 1997:149.

Appendix 1: Required Data from the NNRD

Appendix 1

Appendix 2: Request form for NDAU for access to NNRD data

Appendix 2

Appendix 1: Required Data from the NNRD

Buy? | Total from neonatal period[1]: 25 Total from follow-up[2]: 12 GRAND TOTAL: 37	Category	Category detail	Item Name	Field ID	Data dictionary Item Status	NHS Data Dictionary Item Name	Format/Length	Coding/allowed Values	Description	Purpose	Filter
1	NNUEpisodes	Demographics and Birth Information (Baby)	Unique system identification	AnonPatientID	M	BABY LOCAL PATIENT IDENTIFIER (NATIONAL NEONATAL DATA SET)	an50	ID utilises capital letters of the English alphabet only.	A unique ID that will only identify the baby if used by a user with permission to see the record of that baby.	Used to identify infant in locations of care for the purpose of communication where patient identifiers are not included.	general/Two Year Follow-up
1	NNUEpisodes	Demographics and Birth Information (Baby)	Month of birth	DateTimeofBirthMonth	M	YEAR AND MONTH OF BIRTH (BABY)	n2	Derived month 1 - 12	Calendar month derived from the date and time variable. Date and time variable is identified by prefix to month in the Field ID.	Used in instances where identifiable data is not sent.	anon/general/Two Year Follow-up
1	NNUEpisodes	Demographics and Birth Information (Baby)	Year of birth	DateTimeofBirthYear	M	YEAR AND MONTH OF BIRTH (BABY)	n4	Derived year	Derived from the date and time variable. Date and time variable is identified by prefix to year in the Field ID.	Used in instances where identifiable data is not sent.	anon/general/Two Year Follow-up
1	NNUEpisodes	Demographics and Birth Information (Baby)	Place of Birth NHS Code (Location of baby's birth)	PlaceofBirthNHSCode	M	SITE CODE (OF ACTUAL PLACE OF DELIVERY) or ORGANISATION CODE (OF ACTUAL PLACE OF DELIVERY)	an20	Use organisation code & site code ZZ201 - Not applicable (intended to deliver at home) ZZ888 - Not applicable (intended to deliver at non-NHS organisation) ZZ203 - Not known (intended place of delivery not known)	Place at which the birth took place as recorded.	Used to conduct data analysis on the organisation.	general/Two Year Follow-up
1	NNUEpisodes	Demographics and Birth Information (Baby)	Birth weight (g)	Birthweight	R	BIRTH WEIGHT	n4	Accepted range is between 001-9998 grams.. 9999 Unknown	Birth weight at the time of delivery in grams	Used to identify risk factor on admission to neonatal care.	
1	NNUEpisodes	Demographics and Birth Information (Baby)	Birth length (cm)	BirthLength	O	BIRTH LENGTH	nn.n	99.9 Unknown	Length measured just after birth.	Used to assess growth development.	
1	NNUEpisodes	Demographics and Birth Information (Baby)	Birth head circumference	BirthHeadCircumference	O	BIRTH HEAD CIRCUMFERENCE	nn.n	99.9 Unknown	Occipito-frontal circumference (OFC) measured at birth. In cms to one decimal place.	Used to monitor outcomes according to head circumference	
1	NNUEpisodes	Demographics and Birth Information (Baby)	Gestation age in weeks	GestationWeeks	R	GESTATION LENGTH (AT DELIVERY)	n2	10 to 49	The best obstetric estimate at the time of delivery in weeks. This will normally be based on the postmenstrual age but, if appropriate, may be modified on the basis of antenatal ultrasound. Where gestation at delivery is not known, this is based on the postnatal estimate at results.	Used to identify risk factor on admission to neonatal care.	general/Two Year Follow-up
1	NNUEpisodes	Demographics and Birth Information (Baby)	Gestation age days	GestationDays	R	GESTATION LENGTH (REMAINING DAYS AT DELIVERY)	n	0-6 9 Unknown	Specify, if known, the number of days between whole weeks in the gestation period.	Used to identify risk factor on admission to neonatal care.	general/Two Year Follow-up
1	NNUEpisodes	Demographics and Birth Information (Baby)	Sex of the baby (phenotypic)	SexPhenotype	R	PERSON PHENTOTYPIC SEX	n1	1 Male 2 Female 9 Indeterminate/Intersex	The sex of the baby. 'Not known' is an option if information is missing or not recorded. 'Not specified' is an option for instances where the sex cannot be determined at birth. This option can be changed later if the chromosomal sex of the baby has been determined as follows: Male (XY) or Female (XX) or remain 'Not specified if the genotypic sex is not defined as XX/XY or is still not known	Used to aggregate by sex.	general/Two Year Follow-up
1	NNUEpisodes	Demographics and Birth Information (Baby)	Sex of the baby (genotypic)	SexGenotype	P	PERSON GENOTYPIC SEX (NATIONAL NEONATAL DATA SET)	n1	1 Male 2 Female 9 Indeterminate/Intersex X Genotypic sex unknown	Specify the genotypic sex of the infant when the phenotypic sex of the baby is indeterminate and requires genetic testing. Male (XY) or Female (XX) or remain 'Not specified if the genotypic sex is not defined as XX/XY or is still not known	Used to aggregate by sex.	general/Two Year Follow-up
1	NNUEpisodes	Admission details	Episode of care	CriticalCareIdentifier	R	EPISODE NUMBER (NEONATAL CRITICAL CARE SPELL)	n2	The number of this episode of care for this baby.	The EPISODE NUMBER (NEONATAL CRITICAL CARE SPELL) is used to sequentially identify each CRITICAL CARE PERIOD within a Neonatal Critical Care Spell. The first CRITICAL CARE PERIOD identifier commences at 1; subsequent CRITICAL CARE PERIODS during the same period of care (within the same or different Health Care Providers) are then incremented by 1. For example, a Neonate is admitted to the Neonatal Intensive Care Unit at Trust A, starting a CRITICAL CARE PERIOD and generating EPISODE NUMBER (NEONATAL CRITICAL CARE SPELL) 1. The Neonate is then transferred to a different Health Care Provider, Trust B (ending the CRITICAL CARE PERIOD at Trust A), which generates EPISODE NUMBER (NEONATAL CRITICAL CARE SPELL) 2. The Neonate may then return to Trust A (ending the CRITICAL CARE PERIOD at Trust B), generating EPISODE NUMBER (NEONATAL CRITICAL CARE SPELL) 3.	Used to ascribe outcomes	MODIFIED
1	NNUEpisodes	Admission details	Hospital baby admitted to	ProviderNHSCode	M	SITE CODE (OF ADMITTING NEONATAL UNIT) or ORGANISATION CODE (OF ADMITTING NEONATAL UNIT)	an20	Use organisation code & site code ZZ201 - Not applicable (intended to deliver at home) ZZ888 - Not applicable (intended to deliver at non-NHS organisation) ZZ203 - Not known (intended place of delivery not known)	This is the code for the hospital recording information on this patient. It is a code which identifies an organisation uniquely. For NHS organisations it is a code that is managed by the Corporate Data Administration section of the Department of Health to identify most organisations that exchange information within the NHS or return information to the Centre. Examples of organisations that can be identified this way are NHS Trusts and Health Authorities.	Used to ascribe outcomes	

1	NNUAdhoc	Discharge details	Discharge date and time - anonymised	DischDateTimeAnonDate	R	NUMBER OF MINUTES (BIRTH TO EVENT)	an10	Derived (minutes)	The anonymised version of the data and time variable derived by calculating the time difference, in minutes, between this date and time and the date and time of birth.	Used to measure length of stay	anon
1	NNUEpisodes	Discharge details	Date of death and time -anonymised	DateofDeathAnonDate	R	NUMBER OF MINUTES (BIRTH TO EVENT)	an10	Derived (minutes)	The anonymised version of the data and time variable derived by calculating the time difference, in minutes, between this date and time and the date and time of birth.	Used in survival analyses and calculate anonymised dates	anon
1	Daily	General information	General Information: Date of day of care anonymised	ActiveDateAnonDate	O	NUMBER OF MINUTES (BIRTH TO EVENT)	an10	Derived (minutes)	The anonymised version of the data and time variable derived by calculating the time difference, in minutes, between this date and time and the date and time of birth.	Used in commissioning and to determine level of care	anon
1	Daily	General information	General Information: Day provider NHS code	DayProviderNHSCode	M	SITE CODE (OF ADMITTING NEONATAL UNIT) or ORGANISATION CODE (OF ADMITTING NEONATAL UNIT)	an20	Use organisation code & site code ZZ201 - Not applicable (intended to deliver at home) ZZ888 - Not applicable (intended to deliver at non-NHS organisation) ZZ203 - Not known (intended place of delivery not known)	The code for the hospital recording information on this day and episode for the patient. Can be derived by system provider.	Used in commissioning to ascribe daily activity	
1	Daily	General information	General Information: weight today	DayWorkingWeight	R	PERSON WEIGHT IN GRAMS	n4	001-9998 g 9999 Unknown	The weight, in grams, of the baby on this day of care.	Used to monitor developmental of infant	
1	Daily	General information	General Information: Head circumference today	DayHeadCirc	O	PERSON HEAD CIRCUMFERENCE IN CENTIMETRES	nn.n	99.9 Unknown	Head circumference is in centimetres to one decimal.	Used to monitor developmental of infant	
1	Daily	General information	General Information: Length today	DayLength	O	PERSON LENGTH IN CENTIMETRES	nn.n	99.9 Unknown	Length measured, in cm, on this day of care.	Used to monitor developmental of infant	
1	Daily	Fluids and feeding	Fluids and feeding: Parenteral nutrition today (partial or total)	ParenteralNutrition	R	PARENTERAL NUTRITION RECEIVED INDICATOR	an1	N No Y Yes	Specify if a baby has received any parenteral nutrition on this day of care.	Used to determine level of care	
1	Daily	Fluids and feeding	Fluids and feeding: Intravenous glucose and electrolyte solutions	GlucoseElectrolytes	R	INTRAVENOUS INFUSION OF GLUCOSE AND ELECTROLYTE SOLUTION RECEIVED INDICATOR	an1	N No Y Yes	Specify if a baby has received an intravenous infusion of a glucose and electrolyte solution on this day of care.	Used to determine level of care	
1	Daily	Fluids and feeding	Fluids and feeding: Enteral feeding given	DayEnteralFeed	R	ENTERAL FEED TYPE GIVEN	an250	1 Suckling at the breast 2 Mother's fresh expressed breast milk 3 Mother's frozen expressed breast milk 4 Donor expressed breast milk 5 Breast milk fortifier 6 Formula 9 Not applicable (Nil by mouth)	Types of milk given during this 24 hour period (00.00-23.59). Record all the types of milk given during the day. Nil by mouth is single choice and only selected if no enteral feed for the whole of the 24 hour period.	Used in the National Neonatal Audit Programme	
1	Daily	Fluids and feeding	Fluids and feeding: Formula type	DayFormulaType	R	FORMULA MILK TYPE (SNOMED CT DM+D)	max n18	dm+d Virtual Therapeutic Moiety listing	If enteral feeding is formula, specify the name of the formula. List will be maintained by d-dm+d.	Used to monitor developmental of infant	
1	Daily	Fluids and feeding	Fluids and feeding: Measured volume milk	Totalvolume	R	TOTAL VOLUME OF MILK RECEIVED	an4	000.0 - 999.9	The volume of milk a baby has been fed on this day of care in millilitres (ml)	Used to monitor developmental of infant	Place back in
2	NNUAdhoc	Two Year Follow-up	Date of TwoYear assessment anonymised	TwoYearAssessmentAnonDate	R	NUMBER OF MINUTES (BIRTH TO EVENT)	an10	Derived (minutes)	The anonymised version of the data and time variable derived by calculating the time difference, in minutes, between this date and time and the date and time of birth.	Use in assessing two year health outcomes following discharge from neonatal care	anon/Two Year Follow-up
2	NNUAdhoc	Two Year Follow-up	Date of death post discharge anonymised	TwoYearDeathDateAnon	R	NUMBER OF MINUTES (BIRTH TO EVENT)	an10	Derived (minutes)	The anonymised version of the data and time variable derived by calculating the time difference, in minutes, between this date and time and the date and time of birth.	Use in assessing two year health outcomes following discharge from neonatal care	anon/Two Year Follow-up
2	NNUAdhoc	Two Year Follow-up	Growth – weight	GrowthWeight	R	PERSON WEIGHT (AT TWO YEAR NEONATAL OUTCOMES ASSESSMENT)	nn.n	weight kg.	Specify the weight of the child at follow up.	Use in assessing two year health outcomes following discharge from neonatal care	Two Year Follow-up
2	NNUAdhoc	Two Year Follow-up	Growth – weight, date measured anonymised	GrowthWeightDateAnon	R	NUMBER OF MINUTES (BIRTH TO EVENT)	an10	Derived (minutes)	The anonymised version of the data and time variable derived by calculating the time difference, in minutes, between this date and time and the date and time of birth.	Use in assessing two year health outcomes following discharge from neonatal care	anon/Two Year Follow-up
2	NNUAdhoc	Two Year Follow-up	Growth – length	GrowthLength	R	PERSON HEIGHT IN CENTIMETRES (AT TWO YEAR NEONATAL OUTCOMES ASSESMENT)	nn.n	Length cm.	Specify the length of the child at follow up.	Use in assessing two year health outcomes following discharge from neonatal care	anon
2	NNUAdhoc	Two Year Follow-up	Growth – length, date measured anonymised	GrowthLengthDateAnon	R	NUMBER OF MINUTES (BIRTH TO EVENT)	an10	Derived (minutes)	The anonymised version of the data and time variable derived by calculating the time difference, in minutes, between this date and time and the date and time of birth.	Use in assessing two year health outcomes following discharge from neonatal care	anon/Two Year Follow-up
2	NNUAdhoc	Two Year Follow-up	Growth – head circumference	GrowthHeadCirc	R	PERSON HEAD CIRCUMFERENCE IN CENTIMETRES (AT TWO YEAR NEONATAL OUTCOMES ASSESSMENT)	nn.n	Circumference in cm	Specify the head circumference of the child at follow up.	Use in assessing two year health outcomes following discharge from neonatal care	Two Year Follow-up
2	NNUAdhoc	Two Year Follow-up	Growth – head circumference , data measured anonymised	GrowthHeadCircDateAnon	R	NUMBER OF MINUTES (BIRTH TO EVENT)	an10	Derived (minutes)	The anonymised version of the data and time variable derived by calculating the time difference, in minutes, between this date and time and the date and time of birth.	Use in assessing two year health outcomes following discharge from neonatal care	anon/Two Year Follow-up
2	NNUAdhoc	Two Year Follow-up	Development – at 24 month corrected age. Is development normal (<3months delay)?	DevelopmentNormal	R	TPRG-SEND TWO YEAR CORRECTED AGE OUTCOME ASSESSMENT SCORE (DEVELOPMENT ADDITIONAL QUESTION FOR NATIONAL NEONATAL DATA SET)	an1	N No Y Yes U Unknown	Specify if the development at 24 months is normal with less than 3 months delay.	Additional item for two year follow up	Two Year Follow-up
2	NNUAdhoc	Two Year Follow-up	Development – at 24 month corrected age. Is there mild delay (3-6 months delay)?	DevelopmentMildDelay	R	TPRG-SEND TWO YEAR CORRECTED AGE OUTCOME ASSESSMENT SCORE (DEVELOPMENT QUESTION A)	n	0 No 1 Yes 9 Unknown	Specify the child development is correct at 24 months or is mild delay (3-6 months delay?)	Use in assessing two year health outcomes following discharge from neonatal care	Two Year Follow-up
2	NNUAdhoc	Two Year Follow-up	Development – at 24 month corrected age. Is there moderate delay (6-12 month delay)?	DevelopmentModerateDelay	R	TPRG-SEND TWO YEAR CORRECTED AGE OUTCOME ASSESSMENT SCORE (DEVELOPMENT QUESTION B)	n	0 No 1 Yes 9 Unknown	Specify the child developments are correct at 24 months or are moderate delay (6-12 months delay?)	Use in assessing two year health outcomes following discharge from neonatal care	Two Year Follow-up
2	NNUAdhoc	Two Year Follow-up	Development – at 24 month corrected age. Is there severe delay (>12 month delay)?	DevelopmentSevereDelay	R	TPRG-SEND TWO YEAR CORRECTED AGE OUTCOME ASSESSMENT SCORE (DEVELOPMENT QUESTION C)	n	0 No 1 Yes 9 Unknown	Specify if the child's development is correct at 24 months or if there is severe delay (>12 months delay).	Use in assessing two year health outcomes following discharge from neonatal care	Two Year Follow-up

APPENDIX 2

NNRD data extraction request form

The purpose of this form is to allow the Neonatal Data Analysis Unit (NDAU) to better understand the data you require; please enter as much information as you can

Applicant information:

NNRD reference number (if known)	
Applicant name	Aneurin Young
Applicant job title	Neonatal Nutrition Research Fellow
Email address	a.young@soton.ac.uk
Applicant telephone	07828065831
Affiliated organisation	University Hospital Southampton NHSFT University of Southampton

Chief Investigator information (required if different from applicant):

Chief Investigator name	Mark Johnson
Chief Investigator job title	Neonatal Consultant, Honorary Senior Clinical Lecturer
Chief Investigator Email address	m.johnson@soton.ac.uk
Chief Investigator telephone	023 8120 4643
Chief Investigator organisation	University Hospital Southampton NHSFT University of Southampton

General information

Project title	Reassessing growth in infants born below 32 weeks' gestation in the UK, 2014-2018
Date of application	1st February 2019
Is this a commercial or industry project?	YES ☐ NO ☒
Is this project	Research ☒ Quality improvement ☐ Service evaluation ☐ Other ☐ Audit ☐
What is your research question? (no more than 5 lines)	In preterm infants born <32 weeks gestation and cared for in neonatal units in England, how have advances in neonatal medicine and nutritonal care in the 5-year period since 2014

Research projects please use PICOT format if appropriate P: patient I: Intervention C: comparator O: outcome T: timescale	affected growth (weight, length and head circumference) between birth and discharge, and growth and neurodevelopment at 2 years.
REC number (if available)	

Analysis

Brief purpose of analysis for which you need NNRD data	Postnatal growth of infants born below 32 weeks' gestation in England has been previously described (Cole, Statnikov et al. 2014) using NNRD data, showing a pattern of poor growth, with infants falling 2 marked centile lines on a UK NICM growth chart in the first 2 weeks of life. Recent work published by our unit has demonstrated that the pattern of growth failure previously demonstrated is not inevitable, and that growth similar to that seen in utero can be achieved. However, it is not clear whether this is due to specific practices within our neonatal unit, or in fact the result of more widespread changes in neonatal medicine and nutritional care that have occurred since Cole et al's original study. This study will assess the distribution of birthweights and the patterns of weight gain using routinely-gathered data in infants born below 32 weeks' gestation in the 5 year period since 2014. We will also assess the associated changes in length and head circumference. Feeding data will be used to assess associations with the type of neonatal unit and the mode of feeding and related variables on growth. Additionally, we will use two-year follow-up data to assess associations between growth in the neonatal period and subsequent growth and development. This study will provide important information about the current impact of neonatal care on growth and subsequent outcomes. It will also help identify practices that are associated with improved growth. In addition, if it demonstrates the improved growth seen in our unit over the past 5 years is due to specific practices, we can share these more widely to help other units achieve similar growth.

Patient cohort: Please describe the infants that you would like data extracted for

Cohort Please describe the cohort that you are interested in in no more than 5 lines	All infants admitted to a neonatal in England unit following birth below 32 weeks' gestation.

Time period: Please state the start and end date for your cohort and whether these dates refer to admissions, discharges or other	Start date: 01/01/14 End date: 31/12/18 Infants admitted for neonatal care ☒ Infants discharged from neonatal care ☐ Other ☐ (please provide more detail below)
Geographic criteria: Please indicate which geographic areas you would like to receive data from, you can choose a whole nation or individual Operational Delivery Networks ** For Scottish Data an application must be made prior to data sharing to the PBPP	Whole NNRD ☐ National regions: England ☒ Wales* ☐ Scotland **☐ **Welsh data has been collected since 2012* ***Scottish data has been collected since 2015* If you only require data from individual networks within England then please choose from the list below: East of England ☐ Thames Valley & Wessex ☐ South East Coast ☐ South West ☐ Northern ☐ North West ☐ Trent ☐ Yorkshire ☐ SW Midlands ☐ Staffordshire/Shropshire/Black Country ☐ North East & Central London ☐ South London ☐ North West London ☐
Inclusion criteria: *e.g. Infants born at < 27 weeks; infants with a birthweight >1500g; infants with a diagnosis of Hirschprungs disease*	All infants born at <32 weeks' gestation
Exclusion criteria: *e.g. Infants who were admitted to Transitional care; infants with a congenital anomaly (you will need to define this); infants with missing gestational age at birth data*	None

NNRD data items & data format:

Data format: Please select if you will require your dataset to be at a patient level or if you require aggregated data	*If you are requesting patient level data for research, neonatal units will need to be contacted.* *Aggregated data refers to higher level grouped data where the counts or rates have already been calculated for you, usually presented in a table format.* Patient level ☒ Aggregated data ☐
Patient level data items: If you require patient level data please indicate which NNRD data items you will need in your dataset. ALL data items held in the NNRD are found here ALL diagnostic, procedural and treatment codes used in the NNRD are found here	AnonPatientID, DateTimeofBirthMonth, DateTimeofBirthYear, PlaceofBirthNHSCode, Birthweight , Birthlength , BirthHeadCircumference, GestationWeeks, GestationDays, SexPhenotype, SexGenotype, CriticalCareIdentifier, ProviderNHSCode, DischDateTimeAnonDate, DateofDeathAnonDate, ActiveDateAnonDate, DayProviderNHSCode, DayWorkingWeight, DayHeadCirc, DayLength, ParenteralNutrition, GlucoseElectrolytes, DayEnteralFeed, DayFormulaType, Totalvolume, TwoYearAssessmentAnonDate, TwoYearDeathDateAnon, GrowthWeight, GrowthWeightDateAnon, GrowthLength, GrowthLengthDateAnon, GrowthHeadCirc, GrowthHeadCircDateAnon, DevelopmentNormal, DevelopmentMildDelay, DevelopmentModerateDelay, DevelopmentSevereDelay
Patient level derived data items: If you require additional data items to be derived please list and describe them here	None
Aggregated data items: If you require aggregated data please describe the features of the data set you would like describe	None
Method of aggregation: Please indicate how you would like the data aggregated (i.e. by geographical area)	

Denominator data: Do you need denominator data to describe you cohort within a wider context? E.g. total number of neonatal unit admissions over the same period

Denominator data required	None ☒ Total neonatal unit admissions ☐

We cannot provide data on total number of live births, this is available from ONS	Total neonatal unit admissions by gestation ☐ • Gestation range: Total neonatal unit admissions by birth weight ☐ • Birthweight range: Other ☐ • Describe:
Descriptive variables for denominator data These will be provided as standard summary measures (e.g. means and standard deviation)	Gestation at birth ☐ Birthweight ☐ Other

Dissemination: Please specify how you will disseminate the outcome of your project after analysis is complete	The intention is to publish in a peer-reviewed journal and/or present at appropriate conferences

How did you hear about the NNRD?	A colleague/collaborator ☒ Another research article ☐ A talk or presentation ☐ Social media ☐ NNRD website ☐ Other (please describe):

Cole, T. J., et al. (2014). "Birth weight and longitudinal growth in infants born below 32 weeks' gestation: a UK population study." Arch Dis Child Fetal Neonatal Ed **99**(1): F34-40.

Appendix 9

Application to receive data from the National Neonatal Research Database

Imperial College London

NDAU
Neonatal Data Analysis Unit

NNRD data extraction request form

The purpose of this form is to allow the Neonatal Data Analysis Unit (NDAU) to better understand the data you require; please enter as much information as you can

Applicant information:

NNRD reference number (if known)	
Applicant name	Aneurin Young
Applicant job title	Neonatal Nutrition Research Fellow
Email address	a.young@soton.ac.uk
Applicant telephone	07828065831
Affiliated organisation	University Hospital Southampton NHSFT University of Southampton

Chief Investigator information (required if different from applicant):

Chief Investigator name	Mark Johnson
Chief Investigator job title	Neonatal Consultant, Honorary Senior Clinical Lecturer
Chief Investigator Email address	m.johnson@soton.ac.uk
Chief Investigator telephone	023 8120 4643
Chief Investigator organisation	University Hospital Southampton NHSFT University of Southampton

General information

Project title	Reassessing growth in infants born below 32 weeks' gestation in the UK, 2014-2018
Date of application	1st August 2019
Is this a commercial or industry project?	YES ☐ NO ☒
Is this project	Research ☒ Quality improvement ☐ Service evaluation ☐ Other ☐ Audit ☐
What is your research question? (no more than 5 lines)	In preterm infants born <32 weeks gestation and cared for in neonatal units in England, how have advances in neonatal medicine and nutritonal care in the 5-year period since 2014

Research projects please use PICOT format if appropriate P: patient I: Intervention C: comparator O: outcome T: timescale	affected growth (weight, length and head circumference) between birth and discharge, and growth and neurodevelopment at 2 years.
REC number (if available)	

Analysis

Brief purpose of analysis for which you need NNRD data	Postnatal growth of infants born below 32 weeks' gestation in England has been previously described (*Cole, Statnikov et al. 2014*) using NNRD data, showing a pattern of poor growth, with infants falling two marked centile lines on a UK NICM growth chart in the first two weeks of life. Recent work published by our unit has demonstrated that the pattern of growth failure previously demonstrated is not inevitable, and that growth similar to that seen in utero can be achieved. However, it is not clear whether this is due to specific practices within our neonatal unit, or in fact the result of more widespread changes in neonatal medicine and nutritional care that have occurred since *Cole et al*'s original study. This study will assess the distribution of birthweights and the patterns of weight gain using routinely-gathered data in infants born below 32 weeks' gestation in the 5 year period since 2014. Professor Tim Cole (UCL) will act as a collaborator. In particular, we are interested in regional variations in patterns of growth, and those due to the level of care offered by neonatal units. We will also assess the associated changes in length and head circumference. Feeding data will be used to assess associations with the type of neonatal unit and the mode of feeding and related variables on growth. Additionally, we will use two-year follow-up data to assess associations between growth in the neonatal period and subsequent growth and development. This study will provide important information about the current impact of neonatal care on growth and subsequent outcomes. It will also help identify practices that are associated with improved growth. In addition, if it demonstrates the improved growth seen in our unit over the past 5 years is due to specific practices, we can share these more widely to help other units achieve similar growth.

Patient cohort: Please describe the infants that you would like data extracted for

Cohort Please describe the cohort that you are interested in in no more than 5 lines	All infants admitted to a neonatal unit in England following birth below 32 weeks' gestation.
Time period: Please state the start and end date for your cohort and whether these dates refer to admissions, discharges or other	Start date: 01/01/14 End date: 31/12/18 Infants admitted for neonatal care ☒ Infants discharged from neonatal care ☐ Other ☐ (please provide more detail below)
Geographic criteria: Please indicate which geographic areas you would like to receive data from, you can choose a whole nation or individual Operational Delivery Networks ** For Scottish Data an application must be made prior to data sharing to the PBPP	Whole NNRD ☐ National regions: England ☒ Wales* ☐ Scotland **☐ *Welsh data has been collected since 2012* ***Scottish data has been collected since 2015* If you only require data from individual networks within England then please choose from the list below: East of England ☐ Thames Valley & Wessex ☐ South East Coast ☐ South West ☐ Northern ☐ North West ☐ Trent ☐ Yorkshire ☐ SW Midlands ☐ Staffordshire/Shropshire/Black Country ☐ North East & Central London ☐ South London ☐ North West London ☐
Inclusion criteria: *e.g. Infants born at < 27 weeks; infants with a birthweight >1500g; infants with a diagnosis of Hirschprungs disease*	All infants born at <32 weeks' gestation
Exclusion criteria: *e.g. Infants who were admitted to Transitional care; infants with a congenital anomaly (you will need to define this); infants with missing gestational age at birth data*	None

NNRD data items & data format:

Data format: Please select if you will require your dataset to be at a patient level or if you require aggregated data	*If you are requesting patient level data for research, neonatal units will need to be contacted.* *Aggregated data refers to higher level grouped data where the counts or rates have already been calculated for you, usually presented in a table format.* Patient level ☒ Aggregated data ☐
Patient level data items: If you require patient level data please indicate which NNRD data items you will need in your dataset. ALL data items held in the NNRD are found here ALL diagnostic, procedural and treatment codes used in the NNRD are found here	AnonPatientID, DateTimeofBirthMonth, DateTimeofBirthYear, PlaceofBirthNHSCode, Birthweight , Birthlength , BirthHeadCircumference, GestationWeeks, GestationDays, SexPhenotype, SexGenotype, CriticalCareIdentifier, ProviderNHSCode, DischDateTimeAnonDate, DateofDeathAnonDate, ActiveDateAnonDate, DayProviderNHSCode, DayWorkingWeight, DayHeadCirc, DayLength, ParenteralNutrition, GlucoseElectrolytes, DayEnteralFeed, DayFormulaType, Totalvolume, TwoYearAssessmentAnonDate, TwoYearDeathDateAnon, GrowthWeight, GrowthWeightDateAnon, GrowthLength, GrowthLengthDateAnon, GrowthHeadCirc, GrowthHeadCircDateAnon, DevelopmentNormal, DevelopmentMildDelay, DevelopmentModerateDelay, DevelopmentSevereDelay, PostCodeMotherLSOA, FetusTotal, apgar_1min, apgar_5min, apgar_10min, SurgicalProcedure, RespiratorySupport, ModeofRespiratorySupport, PulmonaryVasodilator, OygenPerc, Inotropesgiven, SurgeryforPDA, DayNEC, LinesIn
Patient level derived data items: If you require additional data items to be derived please list and describe them here	None
Aggregated data items: If you require aggregated data please describe the features of the data set you would like describe	None
Method of aggregation: Please indicate how you would like the data	

aggregated (i.e. by geographical area)	

Denominator data: Do you need denominator data to describe you cohort within a wider context? E.g. total number of neonatal unit admissions over the same period

Denominator data required We cannot provide data on total number of live births, this is available from ONS	None ☒ Total neonatal unit admissions ☐ Total neonatal unit admissions by gestation ☐ • Gestation range: Total neonatal unit admissions by birth weight ☐ • Birthweight range: Other ☐ • Describe:
Descriptive variables for denominator data These will be provided as standard summary measures (e.g. means and standard deviation)	Gestation at birth ☐ Birthweight ☐ Other

Dissemination: Please specify how you will disseminate the outcome of your project after analysis is complete	The intention is to publish in a peer-reviewed journal and/or present at appropriate conferences

How did you hear about the NNRD?	A colleague/collaborator ☒ Another research article ☐ A talk or presentation ☐ Social media ☐ NNRD website ☐ Other (please describe):

Appendix 10

Protocol aid for the acquisition of urine samples from infants

Collecting a Urine Sample for the GAP Study

You will need:


Cotton wool balls


Universal container


50ml syringe


A baby

1. **Put a few cotton wool balls in the nappy**
and wait for the wee… …

2. **Take the plunger out of the syringe**

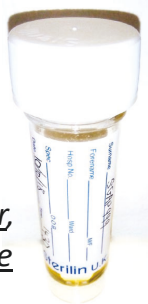
3. **Put the urine-soaked cotton wool in the syringe** (not sterile, minimal poo please!)

4. **Uncap the universal, put syringe tip in universal and replace plunger**

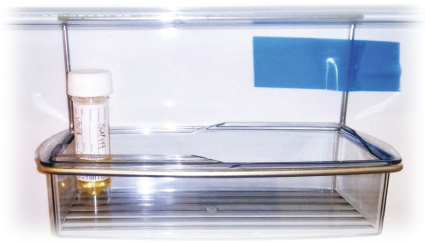
5. ***Squeeee eeeeeze!***

6. **Screw the lid *TIGHT***

Label with *study number, date and time*



7. **Store in Research Fridge in sluice**



8. **CONGRATULATIONS!**

Why not celebrate with a little dance?

Appendix 11

Agreed workflow for blood sample acquisition, storage and DNA extraction

Blood sampling and DNA extraction workflow

University of Southampton | University Hospital Southampton NHS
NHS Foundation Trust

Genomic Samples Acquisition, Storage and DNA Extract

1. Check informed written consent obtained from parents.
2. Draw **1ml whole blood** into **standard neonatal pink EDTA blood tube** at the same time as clinical sampling.
3. **Label** blood sample with study number, date and time (or with hospital number, date and time if no research personnel present).
4. Immediately **store blood sample in research fridge** in sluice on neonatal unit.
5. **Send email** to aneurin.young@uhs.nhs.uk (Aneurin Young, Research Fellow) to inform that sample has been taken.
6. Research fellow to retrieve blood sample from fridge within 18 hours.
7. Research fellow to anonymise sample using study number if required.
8. Research fellow to take sample to Duthie Building Laboratory DA006 and store in -80°C freezer on third shelf down in dedicated rack.
9. Research fellow to email laboratory team (Nikki Graham, njg1@soton.ac.uk) to advise that sample has been deposited.
10. Samples will be stored at -80°C and DNA will be extracted in batches (with the first batch aiming to be performed when 12 samples available).
11. DNA to be extracted using the Canvax HigherPurity™ Blood Genomic DNA Extraction Kit (Cat. No: AN0043) – data sheet included as appendix to this document [but redacted from thesis due to copyright].