Investigating the immune system of extremely preterm infants and the effect of diet

Dr Tom Derek Robert Sproat

This thesis is submitted in part fulfilment of the degree of Doctor of Medicine (MD) at the Translational and Clinical Research Institute, Newcastle University

October 2020

Abstract

Extremely preterm infants are susceptible to life-threatening diseases, specifically late onset sepsis (LOS) and necrotising enterocolitis (NEC). Both these diseases are associated with changes in the gut microbiota and an immature immune response. Mother's own milk (MOM) has been shown to reduce the incidence of both NEC and LOS. When MOM is not available, the alternatives are donor human milk (DHM) or cow's milk formula (CMF). Clinical trials have shown an inconsistent effect on rates of NEC or LOS when DHM is used instead of CMF to make up any shortfall of MOM, although the use of DHM or CMF as sole diet tends to favour DHM. Inconsistencies or lack of effect could be due to differences in the concentration of bioactive components of DHM compared to MOM, as DHM is usually from donors who are longer post-partum, and is usually pasteurised and frozen. The diet of preterm infants affects both their gut microbiota and gut mucosal T cells, which may be instrumental to any impact on LOS and NEC.

This study aimed to identify differences in gut microbial or T cell composition if infants were fed an exclusively human milk (MOM+/-DHM) diet (Intervention) compared to a diet containing bovine products (MOM+/-CMF) (Control).

The hypothesis was that an exclusively human milk diet would be associated with changes in microbial diversity, abundance of *Bifidobacteria*, Regulatory T cells, Mucosa-associated invariant T cells and invariant natural killer T cells.

Infants of less than 30 weeks gestational age (GA) were recruited to a randomised controlled trial comparing the two diets until 34 weeks GA. Stool samples were taken throughout the trial period, which were analysed using 16S rRNA sequencing at 5 time-points, and blood samples taken at 2 time-points were analysed using mass cytometry.

This report provides data from a pilot study of 59 infants. Infants in the intervention group paradoxically received less MOM overall and had decreased rate of growth (weight). There was a significant difference in unweighted microbial beta-diversity at 34 weeks GA and a significantly increased abundance of lactobacillus at 34 weeks GA in the control group. There was no difference in T cell populations between the trial populations, however clear differences were noted when compared to adult control samples.

In conclusion, an exclusively human milk diet did not result in measurable changes in gut bacterial community structure or changes in T cell immunophenotype when compared to a diet containing bovine products. However, the routine use of supplemental probiotics

containing *Bifidobacteria* and *Lactobacillus* in this study population may mask important effects.

Dedication

This thesis is dedicated to my beautiful family; Emma, Samuel and Flora who have (perhaps unwittingly) supported me throughout this process and allowed me to see the lighter side of life at times when I have struggled.

Acknowledgments

There are a huge number of people that have helped me complete this work. Without them I would not have been able to present this thesis and I am truly indebted.

Firstly, I must thank my supervisors; Nick, Janet and Sophie. They have guided me through my ignorance and helped me develop a project that I can be proud of. They have all been thoughtful, considerate and given me many opportunities along the way, as well as assist in the running and recruitment of the study. Nick and Janet designed and set-up the INDIGO study. Sophie kindly gave me access to her laboratory and the resources available there.

This work would not have been possible without the many parents who agreed for their children to participate in the INDIGO study and I thank them profusely. This is together with all the neonatal nurses in the Royal Victoria Infirmary Neonatal Unit who not only cared for the babies in the study, but provided the study intervention and collected the precious stool samples. Julie Groombridge, neonatal research nurse was instrumental in recruiting, sampling and the general running of the study, and became a good friend during my research time. Julie was often assisted by Gail, and I thank them both. Julie (JG), Nick (NDE), Janet (JEB) and Andreas (AK) were the recruiters mentioned later in this thesis.

In the laboratory I owe a lot to a number of people who patiently taught me laboratory skills from the rudimentary knowledge base I possessed. In particular I would like to thank:

Rebecca Payne, who developed the mass cytometry staining protocol and the initial antibody panel, particularly for her wisdom of all things mass cytometry and her ongoing assistance. Jarmila Spegarova, who provided me with her huge knowledge of cytometry and toiled with me through the, at times, hugely frustrating analytical process!

Florian Gothe, another big help with cytometry, and keeping a ready supply of biscuits (preferentially Caramel Hobnobs).

Juan Ojeda humoured me with my many requests for assistance in mass cytometry analysis using his bioinformatic experience, he performed the Nowicka analysis described in this work and displayed in the results section and his work was greatly appreciated.

David McDonald ran the majority of the mass cytometry samples in the Newcastle University flow cytometry core facility.

Greg Young has banked and recorded all the baby poo in Newcastle for a number of years, and gave me access to his expert help in extracting DNA from stools as well as the many challenges microbial analysis brings.

Chris Stewart has always given help whenever asked, given me access to his laboratory for DNA extractions, and providing a colossal amount of help with 16S data analysis. All the 16S analysis described in this work was performed by myself using software developed at the Centre for Metagenomics and Microbiome Research, Baylor University. Chris kindly gave me access to this software.

In the NHS laboratories, Dawn Barge, Ann Bowron and Helen Watson provided help whenever asked. This was together with their expert opinion on my, often silly, queries on amino acid and flow cytometry analysis. Sue Mclellan was ever helpful in retrieving salvaged blood samples for babies taking part in the study and was incredibly responsive to my e-mails

Last, but by no means least, Andreas Kakaroukas and Becky Naples provided moral support through what can be a challenging experience as a clinical research trainee. Thank you for listening to my rants and helping me maintain my sanity (at times).

Funding

Prolacta Biosciences funded the INDIGO study. I received no funding from Prolacta Biosciences to conduct the study.

The mass cytometry analysis was paid for from auxiliary monies held by Professor Nick Embleton, collected through research activities, such as honoraria from speaking at meetings.

Table of Contents

Table of Figures

Table of Tables

List of Abbreviations

16S	-	16S rRNA sequencing
AGA	-	Appropriate for gestational age
APC	-	Antigen presenting cells
AXR	-	Abdominal x-ray
BPD	-	Broncho-pulmonary dysplasia
BMF	-	Bovine milk-based breast milk fortifier
BW	-	Birth weight
C	-	Celsius
CD	-	Cluster of differentiation
CGA	-	Corrected gestational age
CI	-	Confidence interval
CCR	-	Chemokine receptor
CM	-	Central memory
CMF	-	Cow's milk formula
C&W	-	Chelsea and Westminster
DGGE	-	Denaturing Gradient Gel Electrophoresis
DHM	-	Donor human milk
DN	-	Double negative
DOL	-	Day of life
EDTA	-	Ethylenediaminetetraacetic acid
E. Coli	-	Escherichia Coli
EM	-	Effector memory
ESPGHAN		European Society of Paediatric Gastroenterology, Hepatology and Nutrition

FCS	-	Flow cytometry standard format
FDR	-	False discovery rate
GA	-	Gestational age
GP1	-	Goldprem 1 cow's milk formula
HLA	-	Human Leucocyte Antigen
HMF	-	Human milk-based breast milk fortifier
HMO	-	Human milk oligosaccharide
HRA	-	Health Research Authority
IEC	-	Intestinal epithelial cell
IFN-γ	-	Interferon-γ
Ig	-	Immunoglobulin
IL	-	Interleukin
INDIGO		Interactions between the diet, microbes, metabolism and body composition study
iNKT	-	Invariant natural killer T cell
IQR	-	Interquartile range
IRAS	-	Integrated Research Application Service
ISRCTN	-	International Standard Randomised Controlled Trial Number
IVH	-	Intra-ventricular haemorrhage
LBW	-	Low birth weight
LOS	-	Late onset sepsis
LPS	-	Lipopolysaccharide
MAIT	-	Mucosal associated invariant T cell
MAMP	-	Microbe associated molecular pattern
MDS	-	Multi-dimensional scaling
MHC	-	Major histocompatibility complex

MLN	-	Mesenteric lymph nodes
MOM	-	Mother's own milk
MVP	-	Moderate to very preterm
NEC	-	Necrotising Enterocolitis
NHS	-	National Health Service
NICE	-	National Institute of Clinical Excellence
NICU	-	Neonatal Intensive Care Unit
NKT	-	Natural Killer T cell
NLR		Nucleotide-binding oligomerisation domain (NOD) like receptor
NRS	-	Non-redundancy score
OFC	-	Occipito-frontal circumference
OTU	-	Operational Taxonomic Unit
PBS	-	Phosphate Buffer Solution
PGCT	-	Preterm gut community type
PIL	-	Patient information leaflet
PBMC	-	Peripheral blood mononuclear cells
qPCR	-	quantitative polymerase chain reaction
RBC	-	Red blood cell
RCT	-	Randomised clinical trial
RDS	-	Respiratory distress syndrome
REC	-	Research Ethics committee
rEDF	-	Reversed end diastolic flow
ROP	-	Retinopathy of prematurity
RTF 26	-	Prolacta Ready to Feed 26 formula
RR	-	Relative risk

RTE	-	Recent thymic emigrants
RVI	-	Royal Victoria Infirmary
SGA	-	Small for gestational age
SIP	-	Spontaneous intestinal perforation
TCR	-	T cell receptor
T_c	-	Cytotoxic T
T_h	-	T-helper
TLR	-	Toll-like receptor
TNF-α	-	Tumour necrosis factor – α
T_{reg}	-	Regulatory T cell
tSNE	-	t-distributed Stochastic Neighbour Embedding
TP	-	Time-point
T&W	-	Tyne and Wear
Vα 7.2	-	Vα 7.2 – Jα3.3/20/12
WGS	-	Whole genome sequencing
WBC	-	White blood cells

Chapter 1. Introduction

1.1 Diet and extremely preterm infants

1.1.1 *Extremely preterm infants*

Extremely preterm infants are those born prior to 28 completed gestational weeks. Whilst the majority of preterm infants (those born <37 weeks gestation) require intervention to establish oral feeding and maintain a normal body temperature, extremely preterm infants typically require additional support, including respiratory and nutritional support. Without this additional support the mortality rate in this group is high, as seen in low- and middle-income countries (Katz et al., 2013, Gladstone et al., 2015). In high income countries the mortality rate remains around 30%, with an even higher morbidity burden (Stoll et al., 2015). The causes of mortality and morbidity have changed in this population over the last 20 years, with reduced deaths from respiratory causes but increases in the proportion of infants that are dying from other conditions such as necrotising enterocolitis (NEC) and late onset sepsis (LOS) (Berrington et al., 2012).

1.1.2 *Necrotising Enterocolitis*

NEC is a disease characterised by necrotic or inflamed bowel confirmed during post-mortem or surgery. Diagnosis is initially based on clinical findings which can be subjective. NEC can present in a myriad of ways with intestinal symptoms of obstruction, inflammation or perforation, or with cardio-respiratory compromise.

The condition primarily affects infants born before 32 weeks gestational age (GA) (Rose and Patel, 2018). Investigations typically include serum C-reactive protein, full blood count and abdominal radiograph although these investigations lack specificity and sensitivity, which continues to present a major challenge in accurate diagnosis and subsequent management.

There are have been a number of attempts to provide diagnostic or grading schemes for NEC, which is unsurprisingly difficult for a condition that is in reality a pathological description. The first grading criteria were proposed by Bell et al in 1978 which used risk factors, signs, symptoms and x-ray findings to categorise NEC into one of three stages (Bell M. J., 1978). This grading system has been modified to include laboratory findings and remains widely used in clinical trials. Diagnostic criteria widely used in the United Kingdom are those of the National neonatal audit programme (NNAP) which include a combination of clinical and radiological

findings to diagnose NEC (Health, 2017). A weakness of most clinical diagnostic criteria used in NEC is the interpretation of abdominal x-rays. Whilst the diagnosis of free intra-peritoneal air is relatively unambiguous, the pathognomic finding of intra-mural gas (also known as pneumatosis intestinalis, example in highlighted area on Figure 1) is more difficult as impacted stool can resemble this appearance in a well infant.

1.1.2.1 Disease Burden

Between 2-10% of infants born less than 32 weeks corrected GA (CGA) are diagnosed with NEC (Battersby et al., 2018) which has a mortality rate of 20-40% (Battersby et al., 2018). Long-term sequelae are also important and can include intestinal stricture, short bowel syndrome, or renal failure, and the risk of impaired neuro-development is also increased (Hickey et al., 2018). A greater understanding of the underlying mechanisms may improve these outcomes.

1.1.2.2 Pathogenesis

Although the underlying mechanisms for NEC are not known, current data suggest a role for environmental and genetic factors that influence the gut microbiota and hence the integrity of the gut epithelial barrier as illustrated in Figure 1. Impaired barrier function results in bacterial translocation and initiation of an immature immune response, eventually manifesting as a systemic disease (Bode, 2018). An association that has been the focus of research for over 30 years is that NEC is more likely to occur in infants fed cow's milk formula (CMF) based diets rather than their mother's own breast milk (MOM) (Meinzen-Derr et al., 2009, Corpeleijn et al., 2016, Lucas and Cole, 1990, E Corpeleijn et al., 2012). It is not clear if this is due to protective properties of MOM or detrimental properties of CMF. MOM has been the focus of research regarding the prevention and treatment of NEC.

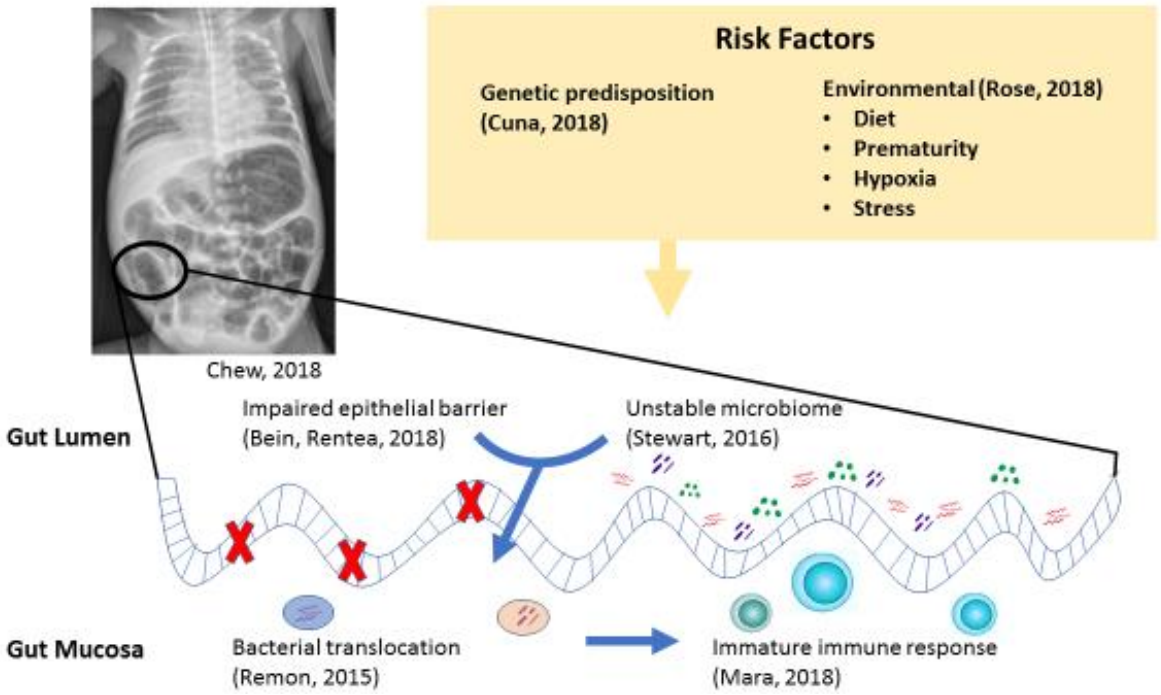


Figure 1 - Current understanding of the risk factors and pathogenesis of Necrotising Enterocolitis (Bein et al., 2018, Rose and Patel, 2018, Stewart et al., 2016, Remon et al., 2015, Mara et al., 2018, Chew et al., 2018)

1.1.3 *Late Onset Sepsis*

LOS is a disease that may present in a similar fashion to NEC. LOS is typically defined as sepsis occurring more than 72 hours after birth, and this thesis will refer only to blood culture-positive LOS unless otherwise stated. LOS is diagnosed in 20-40 % of infants born extremely preterm and contributes to the death of 4-11% of infants born before 32 weeks (Stoll et al., 2015, Berrington et al., 2012). It can be difficult to distinguish LOS from NEC, and these diseases are commonly grouped together when reporting outcomes in clinical research. This is in part due to NEC and LOS sharing similar risk factors, and the lack of robust case definitions.

Organisms associated with LOS are typically commensals of either the skin or the gastro-intestinal tract. Risk factors include prematurity, poor hand hygiene and repeated intravenous catheter insertion. Importantly, prospective studies have shown that the abundance of the LOS pathogen expands in the gastro-intestinal tract of preterm infants before the onset of LOS (Stewart et al., 2017b). This raises the possibility that the gut microbiota is involved in initiation of LOS. Modulation of the gut microbiota through the use of medications that contain (probiotics) or promote (prebiotics) specific bacteria has been an area of intense interest

(AlFaleh and Anabrees, 2014). Similar to NEC, a reduction in LOS has been shown with a diet of MOM (Patel et al., 2013), again suggesting that an understanding of the diet of preterm infants is important in improving outcomes.

1.1.4 *Preterm Nutrition*

Preterm infants have increased energy requirements compared to infants born at term (Joosten et al., 2018), although extremely preterm infants are not able to tolerate large volumes of enteral milk in the first days of life. Parenteral nutrition is therefore used to provide nutrition whilst enteral feeding is established slowly over a period of days (Dutta et al., 2015) and most Neonatal Intensive Care Units (NICU) provide parenteral nutrition immediately from birth. Infants are typically able to feed orally starting from 32-34 weeks' gestational age, so prior to this point, all enteral nutrition is administered via a gastric tube (nasogastric or orogastric). When enteral feed nutrient intakes are tolerated (between 135-200ml/kg/day) parenteral nutrition is stopped (Agostoni et al., 2010). If the enteral nutrition consists of MOM, which has lower protein and energy concentration than CMF (explored in Section 1.1.6), then a breast milk fortifier is commonly added to MOM in order to meet recommended nutrient intakes (Dutta et al., 2015).

1.1.4.1 Breast milk fortifier

Breast milk fortifiers are food supplements than contains protein, energy, minerals and other micronutrients and increases the nutrient density of MOM-based diets. This is typically as a freeze-dried powder derived from cow's milk, which will be referred to as bovine-milk based breast milk fortifier (BMF). More recently, a breast milk fortifier derived from donated human milk has become commercially available, which will be referred to as human-milk based breast milk fortifier (HMF).

1.1.5 *Comparison between human milk and cow's milk formula*

1.1.5.1 Clinical trials

MOM is the optimal diet of preterm infants, primarily due to the lower risk of LOS and NEC (E Corpeleijn et al., 2012, Sisk et al., 2007). Unfortunately, many mothers are unable to or may choose not to meet their infant's requirements with MOM alone (Victora, 2016). As a result, MOM is often supplemented or replaced altogether with either CMF or donor human milk (DHM). The question of whether to supplement MOM with DHM or CMF, if there is a shortfall,

was considered in clinical trials conducted in the 1980's. These trials showed lower rates of NEC but found low birthweight infants had slower growth when DHM was used (Behrman et al., 1983, Lucas et al., 1984). Clinical care (both maternal and neonatal), processing of DHM and CMF have since changed dramatically, making it difficult to draw strong conclusions for current practice from these historic trials for current populations of preterm infants (McGuire and Anthony, 2003).

In the last 20 years, six randomised clinical trials have explored the role of DHM diets. These trials have had different designs and primary outcomes which are summarised in Table 1. Unfortunately, none of these trials were powered solely for NEC or LOS as a primary outcome, however two used a composite outcome involving these diseases. A meta-analysis conducted in 2018 reviewed 5 of the trials in Table 1 as well as including trials performed in the 1980's. The meta-analysis concluded that DHM reduced the incidence of severe NEC (RR 1.87 (95% CI 1.23-2.85)) compared to CMF when there is a shortfall of MOM (Quigley et al., 2018) however there was no difference in mortality or neuro-development which might have been expected to accompany a reduction in severe NEC. Furthermore, as was found in the earlier trials, there was greater weight gain and increase in head circumference in infants fed CMF when there is a shortfall of MOM. Poor weight gain is associated with worse neuro-development so this finding remains a concern for exclusive human milk diets.

1.1.5.1.1 Comparison of fortifiers

Only two trials listed in Table 1 used a HMF, with the others using a BMF. Cristofalo et al. compared a group receiving CMF with a group receiving DHM with HMF and therefore did not compare the use of different fortifiers for MOM (Cristofalo et al., 2013). Sullivan et al. used both a HMF and a BMF in the intervention and comparison arms respectively, finding no significant difference in their primary outcome (duration of parenteral nutrition). However, they did report a high use of MOM (>70%) and a significant reduction in NEC as a secondary outcome (Sullivan et al., 2010).

Differences in clinical outcomes depending on the type of fortifier used have been explored by O'Connor et al. who compared the effect of using either a HMF or a BMF on feeding tolerance and morbidity in a blinded study of low-birth weight (birth weight <1250 grams) infants. Infants (n = 127) were randomised to receive either HMF or BMF when they were receiving 100ml/kg/day of either MOM or DHM. The intervention stopped when the infant was either 84 days old, discharged or established oral feeding. The primary outcome was interruption in feeding for >12 hours or >50% reduction in feeding volume at any time. No significant difference was found in this primary outcome or a secondary outcome using a morbidity and mortality index (O'Connor et al., 2018).

In summary, no adequately powered trials to determine a reduction in NEC or LOS as a primary outcome have been conducted that compare preterm infants fed an exclusive human milk diet to one containing bovine products, there remain concerns regarding growth using an exclusive human milk diet.

(Primary author, Year)	Population	Intervention	Comparison	Primary Outcome	Results	Strengths	Limitations
(Schanler et al., 2005)	23-30 weeks GA at birth. N = 243 (81 DHM, 92 CMF, 70 MOM only)	DHM if shortfall of MOM. Intervention from 4-90 days age (or discharge)	CMF if shortfall of MOM or MOM only BMF used in both groups	Composite outcome of NEC/LOS.	MOM group had lowest rates of LOS +/-NEC. DHM group had lower rate of NEC than CMF group (6 vs 11%) but not significant	Blinded, 3 groups (including MOM only group)	Over recruited to meet MOM group numbers. 17 (21%) of DHM group given CMF due to poor weight gain BMF in both groups
(Sullivan et al., 2010)	BW 500-1250g. N = 207 (HM 100 = 67, HM 40 = 71, CMF =69)	DHM if shortfall of MOM. HMF used as fortifier Intervention until either 91-days, discharge or 4 oral feeds/day	CMF if shortfall of MOM BMF used as fortifier.	Duration of parenteral nutrition (days)	No change in primary outcome. Significant reduction in NEC/surgical NEC in HM groups	Kept to study protocol. HMF and BMF used	Not blinded. Not powered for any clinical outcomes. >70% of all enteral nutrition was MOM (82% in CMF group)
(Cristofalo et al., 2013)	BW 500-1250g. N = 53 (DHM = 29, CMF =24) Eligible if no MOM to be provided	DHM as sole diet. HMF used as fortifier. Intervention until either 91 days, discharge or 4 oral feeds/day	CMF as sole diet.	Duration of parenteral nutrition (days).	**Significant reduction in primary outcome** Significant reduction in NEC (21% vs 3%)	Blinded, powered.	Not powered for clinical outcomes.
(Corpeleijn et al., 2016)	BW < 1500g N = 373 (DHM = 183, CMF = 199)	DHM if shortfall of MOM Intervention for first 10 days of life	CMF if shortfall of MOM	Composite outcome of survival without NEC/LOS (until 60 days)	No change in primary outcome (at 10 or 60 days)	Double blinded, powered.	More than 80% of enteral nutrition was MOM BMF in both groups
(O'Connor et al., 2016)	BW < 1500 g N = 363 (DHM = 181, CMF = 182)	DHM if shortfall of MOM Intervention until 90 days of age or discharge	CMF if shortfall of MOM	Cognitive Score on Bayley Score (3rd Edition) at 18 months CGA	No change in primary outcome. Significant changes were seen in NEC (worse in CMF group) and in severe cognitive impairment (worse in DHM group)	Blinded, powered.	Not powered for NEC outcome BMF in both groups
(Costa et al., 2018)	<32 weeks GA at birth N = 70 (DHM = 35, CMF = 35)	DHM if shortfall of MOM Intervention until primary outcome	CMF if shortfall of MOM	Day of full enteral feeds (150ml/kg/day for 3 consecutive days)	No change in primary outcome	Powered. 90% recruitment rate of eligible participants	Not blinded. Relatively mature cohort (mean GA 30.1 weeks and BW 1350g) No cases of NEC BMF in both groups

Table 1 - Summary of Randomised trials comparing DHM and CMF in preterm infants from the year 2000-2020. BW (birth weight)

1.1.5 Growth of preterm infants

1.1.5.2 Growth references and standards

The optimal growth of preterm infants is not clearly defined. The UK-World Health Organisation (WHO) growth chart is based on birth weights of infants who were part of 3 studies in different areas of the United Kingdom between 1983 and 1994 (Cole, 1998). This growth chart is a reference chart for the birthweights of preterm infants. In a similar fashion, Fenton et al. have published a meta-analysis of birthweights of 34,639 infants from 8 countries who were all born less than 30 weeks GA, providing a reference range of birthweights (Fenton and Kim, 2013). These birthweight references are often used as growth standards, with the extra-uterine growth of a neonate being expected to follow their birth centile, therefore matching intra-uterine growth. This may be unrealistic because of the increased energy requirements of preterm infants (Joosten et al., 2018) and indeed may not be desirable when longer term metabolic outcomes are considered.

The development of growth standards based on the postnatal growth of preterm infants appears a more realistic target. The INTERGROWTH-21st Project is a recent example, where the growth of 201 preterm infants born after 26 weeks GA was followed for the first 5 months of life (Villar et al., 2015) This project was conducted in 8 different countries including the United Kingdom. There are still challenges in the interpretation of these data as infants were excluded if intra-uterine growth restriction or congenital abnormalities were present. Despite optimal growth being difficult to define, poor growth is known to be associated with worse cognitive outcomes in preterm infants.

1.1.5.3 Consequence of abnormal growth

Optimal growth is difficult to define in preterm infants, however poor growth is commonly reported. This is felt to be mainly due to high macronutrient needs and because of the practical difficulties in providing adequate nutrient intakes. Poor growth has been associated with early morbidities and poor neurodevelopmental outcomes, whilst excessive growth in early life has been associated with adverse metabolic outcomes.

1.1.5.3.1 Early morbidities

Regev et al. retrospectively explored a group of 12,992 infants born extremely preterm from 1995 to 2013. They hypothesised that head growth failure, defined as a decrease in z-score of >=2 between birth and discharge, would be associated with common morbidities (NEC,

respiratory distress syndrome (RDS), broncho-pulmonary dysplasia (BPD), sepsis (early or late-onset)). They found that there was an association between all morbidities and head growth failure, however not for infants that were small for gestational age (SGA), based on weight (Regev et al., 2016).

1.1.5.3.2 Neurodevelopmental outcome

Ramel et al. retrospectively explored the association between growth and cognitive outcome at 2 years in very low birth weight (LBW) but appropriate for gestational age (AGA) infants (n=62). They found that shorter length, using z-scores, at 4 months and 24 months once corrected for weight and head circumference was associated with worse 2 year neurodevelopment outcomes (Ramel et al., 2012). Meyers et al. conducted a similar study retrospectively analysing 1227 infants born between 23 and 29 weeks. They defined two groups based on the change in body length between birth and discharge, hypothesising that poor linear growth would be associated with poor neurodevelopmental outcomes at 2 years. They demonstrated a linear relationship between poor growth (length) and reduction in scores in the language and cognitive domains of the Bayley III test (Meyers et al., 2019). Similarly, Pfister et al. prospectively followed up a cohort (n=32) for 4 years hypothesising that poor weight gain would be associated with slower speed of processing, whilst fast weight gain would be associated with hypertension. They demonstrated a linear relationship between weight gain (from term corrected age to 4 months) and hypertension but not speed of processing (as measured by visual evoked potentials) at 4 years (Pfister et al., 2018). However, Belfort et al. demonstrated a positive linear relationship between weight gain from birth to term GA and Bayley III scores, in a group of 613 infants less than 33 weeks GA. They also found a linear correlation in motor scores and weight gain from term to 4 months of age (Belfort et al., 2011)

These studies highlight that appropriate weight gain in the first few months of life appears important for normal neurodevelopment. Research associating growth with outcomes is difficult to conduct therefore studies are often subject to limitations. Namely, they are often retrospective studies, have poor follow-up rates, and in the studies mentioned above there was no correction for false discovery rates (due to repeated statistical tests being used). Nonetheless, the consistent association between poor weight gain in preterm infants and adverse neurodevelopmental outcomes is a concern. An understanding of the variation in nutrient composition of human milk and CMF is therefore vital.

1.1.5.3.3 Metabolic outcomes

The period of early growth that impacts on metabolic outcomes is unclear. Kerkhof et al. investigated 162 adolescents who were born preterm (<36 weeks GA) and a term control

group with a range of metabolic investigations. They demonstrated that rapid weight gain up until 3 months of age was associated with increased body fat percentage and waist circumference at 20 years of age (Kerkhof et al., 2012). A similar study investigating metabolic outcomes in 153 children born at a median GA of 30 weeks did not confirm this. Although, it did demonstrate strong associations between rapid growth after 1 year of age and a range of metabolic outcomes including insulin sensitivity and fat mass. (Embleton et al., 2016).

1.1.6 *Composition of human and cow's milk formula*

Human milk varies in composition depending on the time of day, the length of the feed (fore and hind-milk) and the age of the child. Milk composition is described as going through 3 stages; colostrum, transitional milk (5 days to 2 weeks post-delivery) and mature milk (Ballard and Morrow, 2013). The constituents vary depending on the stage of the milk. This is in distinction to the constituents of CMF which vary little due to being manufactured according to strict guidance (Health, 2013). The contents of milk can be broadly separated into the macronutrients, micronutrients and bioactive components.

1.1.6.1 Macronutrient content

The macronutrient content of milk mainly comprises protein, fat and lactose (major carbohydrate). The mean macronutrient intake from CMF, DHM and MOM that a preterm infant may receive are summarised in Table 2 assuming a fluid intake of 165ml/kg/day and a weight of 1kg. They are shown alongside nutrient intake recommendations for preterm infants from the ESPGHAN (European Society of Paediatric Gastroenterology, Hepatology and Nutrition). The ESPGHAN recommendations are based upon a systematic review of prospective studies and the use of a factorial model based on in-utero accretion of macro-nutrients (Agostoni et al., 2010).

The protein content of human milk varies depending on post-conceptual age, with higher concentrations in preterm milk (Stoltz Sjostrom et al., 2014). Human milk protein content decreases with age after birth and is lowest at approximately 6 months of age (Saarela et al., 2005, Lucas and Hudson, 1984). As DHM is commonly from mothers who are more than 6 months post-partum, this explains the lower protein content of DHM compared to MOM (Table 2).

The fat content of human milk provides the majority of the calories. If considered in a 24 hour period, the amount of fat does not differ greatly between human milk and CMF, however the

amount is highly variable depending on being from hind or fore milk (Ballard and Morrow, 2013).

Lactose is the main carbohydrate of human milk. There is little variability in lactose concentration between preterm or term human milk or between fore and hind milk (Saarela et al., 2005).

Table 2 - Macronutrient content of Preterm CMF, DHM, MOM and Fortified MOM assuming a 1kg infant and a fluid intake of 165ml/kg/day. Mean values are shown (1 decimal place). Values highlighted orange are below recommended levels and values highlighted red are above recommended level.

***Preterm CMF includes SMA® PRO Gold Prem 1, Cow & Gate® Nutriprem 1 Low Birthweight formula and Cow and Gate Hydrolysed Nutriprem (range shown)**

^Fortified MOM represents MOM fortified with Cow and Gate® Nutriprem human milk fortifier (BMF)

Nutrient	Preterm CMF*	DHM (mean values) (Wojcik et al., 2009)	MOM (Preterm) (Stoltz Sjostrom et al., 2014)	Fortified^ MOM (Preterm)	ESPGHAN recommendations (Agostoni et al., 2010)
Energy (kcal)	132	107	114	139	110-135 kcal/day
Protein (g)	4.29-4.78	1.9	2 - 3.6	3.8 - 5.4	3.5-4.5 g/day
Protein: Energy (% of kcal)	13 - 14.5	7.1	7 - 12.6	10.9 - 15.5	12.8-16.4
Carbohydrate (g)	13.4 - 13.9	12.9	10.6	15	11.6-13.2 g/day
Fat (g)	6.5 - 6.6	5.3	5.8	5.8	4.8-6.6 g/day

1.1.6.2 Micronutrient content

Micronutrients including vitamins, minerals and trace elements are found in breast milk. Human milk is known to contain relatively low amounts of certain vitamins, especially vitamin K and vitamin D (Ballard and Morrow, 2013). Trace elements found in human milk are known

to vary between geographical populations with less variance within a population (Bjorklund et al., 2012). Whilst variability has been attributed to diet, it is difficult to draw strong conclusions due to limited sample size and heterogeneity in the design of studies exploring micronutrient content of human milk (Bravi et al., 2016). In order to reduce micronutrient deficiencies, supplementation with vitamins A, C and D is recommended by the National Institute of Clinical Excellence (NICE) to breastfeeding women and children between 6 months and 4 years of age who receive less than 500ml CMF per day. Vitamin D supplementation is recommended to breast fed infants from one month if their mother did not take vitamin D during pregnancy (Excellence, 2014a, Excellence, 2014b). Whilst it may be possible to control the intakes of certain nutrients that infants receive, the same is not possible for bioactive components.

1.1.6.3 Bioactive components

Bioactive components likely drive the health benefits of human milk. These are non-nutritive elements with functional activity (e.g. for immunity) and include human milk oligosaccharides (HMO's), lactoferrin, bacteria and immune cells. These components have all been implicated in the host-microbe interaction, explored further in Section 1.4.

1.1.6.3.1 Human Milk Oligosaccharides

HMO's are carbohydrates that comprise 1-2% of the total carbohydrates of human milk. They are not digestible by the infant but have been shown to stimulate the growth of specific bacteria which are deemed beneficial, as well as being absorbed intact into the vasculature (Underwood et al., 2015). Whilst they are seen as a key component of MOM, the composition and balance of HMOs differs between women possibly explaining some of the variation in the effect of MOM (Azad et al., 2018). The inclusion of HMO's in CMF is now being explored (Puccio et al., 2017).

1.1.6.3.2 Lactoferrin

Lactoferrin is a member of the transferrin family. The concentration of lactoferrin is higher in preterm than term MOM, and is highest in colostrum (Villavicencio et al., 2017). Lactoferrin has bactericidal, bacteriostatic and bacteria-stimulating properties, and is part of the innate response to infection (Legrand, 2016).

1.1.6.3.3 Bacteria

Human milk is known to contain diverse microbes, which vary between individuals and are not solely due to skin bacterial colonisation (Gomez-Gallego et al., 2016). DHM undergoes pasteurisation, largely to destroy any pathogenic bacteria and viruses, however this also eliminates beneficial bacteria (Lima et al., 2017). Whether DHM can be personalised with the

microbiome of MOM in order to convey benefit to the infant is an area of ongoing research (Cacho et al., 2017). The human milk microbiota are explored in more detail in section 1.2.4.3.

1.1.6.3.4 Immune cells

Human milk contains leucocytes (neutrophils, eosinophils, basophils, lymphocytes and monocytes) and stem cells, which are all in greater abundance in early human milk (Trend et al., 2015). Milk derived maternal stem cells have been shown to translocate from the infant gut to distant organs including the brain, kidneys, thymus and liver in a mouse model (Hassiotou et al., 2015). Once there, stem cells have been shown to differentiate (Aydin et al., 2018). If this were shown to be true in humans, this could help to explain the diverse range of benefits of MOM.

Unfortunately, the processing of DHM is known to affect its bioactive components, eliminating viable cells (Table 3). This may explain why it does not have the same effect as MOM in extremely preterm infants.

1.1.7 *Donor human milk production*

A challenge of using DHM is the inter-sample variation in the concentration of nutrients and bioactive components. This may lead to concerns regarding the adequacy of nutritional intake and growth of preterm infants fed DHM. Bulk processing of DHM and subsequent production of a standardised content of DHM can overcome some of the nutrient variation (Medo, 2017). Pasteurisation is still deemed necessary to ensure DHM is microbiologically safe. New pasteurisation techniques are being explored to minimise the effect on the bioactive components of human milk (Escuder-Vieco et al., 2018).

Table 3 - Effect of pasteurisation on bioactive components of human milk

Bioactive Component	Effect of pasteurisation
HMO's	No reduction (Hahn et al., 2017)
Lactoferrin	More than 50% reduction (Daniels et al., 2017)
Bacteria	Almost total loss, *Bacillus cereus* can survive (Lima et al., 2017)
Leucocytes	More than 90% reduction (Contador et al., 2013)

1.1.8 Summary of the diet of extremely preterm infants

All infants less than 32 weeks GA need additional nutritional intakes of calories, protein and vitamins.

NEC is a disease predominantly affecting infants born less than 32 weeks GA that frequently leads to death or long-term morbidity. Diet is a known risk factor with MOM the preferential choice. When there is a shortfall of available MOM, the best alternative is not clear. DHM is an alternative however its nutrient and bioactive content can vary. This variation can be minimised through processing of DHM, as well as use of a fortifier.

The bioactive components of human milk are what likely convey its benefits to infants, they are virtually absent from CMF and reduced in DHM. Due to this reduction in bioactive components, any beneficial effect of DHM is likely reduced compared to MOM. Whilst a large randomised clinical trial (RCT) with clinical outcomes is unlikely to occur due to the size and therefore expense required, mechanistic research may help uncover differences. Diet, NEC and LOS have been shown to be associated with changes in the gut microbiota and immune system of preterm infants. Modulation of the gut microbiota may be the reason for the beneficial effect of MOM, and any effect of DHM or HMF.

1.2 The preterm gut microbiota

1.2.1 *The microbiota*

The microbiota describes the microbial community of a particular site, including bacteria, fungi, viruses and protozoa. This work focuses on the bacterial microbiota, and uses the term microbiota to refer to the bacterial microbiota. The human body is known to contain more bacteria than human cells, with the gut containing approximately 10^{11} bacteria/ml, which are collectively termed the gut microbiota (Sender et al., 2016).

The identification of these bacteria was made possible by culture-independent techniques which identify bacteria by the presence of their genetic material. The most commonly used technique in recent years is 16S rRNA gene sequencing (16S). The 16S region of the ribosomal RNA is highly variable between bacteria but is also slow to evolve making it ideal for identifying different bacteria. Bacterial identification is made by correlation with a known database, with the output referred to as the operational taxonomic unit (OTU). 16S sequencing can reliably identify bacteria to the genus taxonomic level. This may be a limitation, as specific species or strains of bacteria may be important, which cannot be precisely identified using 16S. The microbiome refers to the microbiota and their associated genetic material.

There are other culture-independent techniques including Denaturing Gradient Gel Electrophoresis (DGGE), quantitative polymerase chain reaction (qPCR) and whole genome sequencing (WGS). The limitations of DGGE and qPCR include bias due to the use of primers and lack of resolution between bacterial communities. qPCR does have a useful role in the quantitative measurement of bacteria whereas 16S will only provide relative abundance of OTU's. Whole genome sequencing (WGS) is increasingly being utilised as it becomes cheaper. WGS allows resolution of bacterial communities to a sub-species (strain) level.

1.2.1.1 Why is the microbiome important?

The microbiome has been the subject of intense research over the last 10 years with over 12,000 publications devoted to this topic in a 5 year period up to 2017 (Cani, 2018). This is due to the association of changes in the microbiome with a number of diseases ranging from Inflammatory Bowel Disease to Alzheimer's Disease (Vogt et al., 2017, Gkouskou et al., 2014). The majority of human studies can only demonstrate an association between the gut microbiome and disease, and not any causal relationship. This is because experimentally it is difficult to manipulate the microbiome, hence to demonstrate causality. Understanding how different microbiota may be causing disease is therefore best understood from animal studies.

1.2.1.2 By what mechanism do microbiota affect their host?

Animal models have demonstrated a number of mechanisms by which microbiota may induce disease, two examples below highlight the role of *Klebsiella* in inducing colitis and *Bifidobacterium longus* in protecting from *Escherichia Coli (*E. Coli*)* 0157 infection.

White et al. (2017) investigated the role of Paneth cells in the development of colitis in mice. They demonstrated that colitis only occurred after the intestine matured to contain Paneth cells. Colitis only developed in the presence of live (not dead) bacteria and when the Paneth cells were chemically disrupted. A combination of chemical disruption of Paneth cells and pathogenic bacteria (*Klebsiella*) was shown to result in translocation of intra-intestinal contents through the intestinal epithelial barrier with measurement of increased concentrations of FITC-tagged dextran in the serum compared to controls. Either *Klebsiella* or Paneth cell disruption in isolation did not have the same effect (White et al., 2017).

Bifidobacteria are of intense interest in newborn infants due to their increased presence in the gut of breast-fed infants. In mice, it has been shown that a specific *Bifidobacterium* species (*Bifidobacterium longum*) protects mice from a usually fatal *E. Coli* 0157 infection. Fukuda et al. used metabolomics and transcriptomics to demonstrate that this protection was due to the ability of *Bifidobacterium longum* to secrete acetate, which reduced inflammation and apoptosis in epithelial cells (Fukuda et al., 2012).

Table 4 – Bacterial taxonomy using the example of the species *Escherichia coli*

Domain	Bacteria
Phylum	Proteobacteria
Class	Gammaproteobacteria
Order	Enterobacteriales
Family	*Enterobacteriacae*
Genus	*Escherichia*
Species	*Escherichia coli*
Strain	*Escherichia Coli 0157:H7*

1.2.1.3 How is the microbiome analysed?

Analysis and interpretation of a high-dimensional dataset such as WGS or 16S is a major challenge. During analysis, the microbiome is commonly described in terms of diversity and composition. Alpha-diversity refers to the variation in OTU's within a sample, commonly described in terms of number of different OTU's and total number of OTU's (richness). Beta-diversity is used to describe the variation in OTU's between groups of samples, which can take into account alpha-diversity. Beta-diversity can be used to compare diversity between two groups of samples but cannot be quantified to be higher or lower in one group. Microbiome composition typically describes the relative abundance of different microbials at a particular taxonomic level (Zalewski et al., 2018).

Inferred metabolic profiling can be derived from the bacteria identified, when a database is used with information about specific bacteria's metabolic potential. This can become more specific when OTU's can be identified to the species or strain level.

There is no standard method of reporting microbiome data, however maintaining consistent analysis by minimising variation in sample collection and processing is thought to be vital (Pollock et al., 2018). This includes immediate freezing of samples and use of an appropriate lysis method to ensure breakdown of all bacteria.

1.2.2 *Development of the preterm gut microbiome*

The preterm gut microbiome has been shown to be influenced by a number of demographic (e.g. gestational age) and environmental factors (e.g. diet, disease). This review will focus on

research using 16S or WGS techniques on stool samples from a population including infants born less than 32 weeks GA unless otherwise stated. The descriptor “preterm infants” will refer to infants born less than 32 weeks GA unless stated otherwise.

Microbial colonisation of the preterm gut may start before birth with one study suggesting that amniotic fluid contains bacteria (Urushiyama et al., 2017). Immediately following birth, the faeces show a unique composition dominated by human proteins. It takes approximately 2 weeks for the abundance of faecal bacterial proteins to exceed the faecal human proteins (Brown et al., 2018, Xiong et al., 2017). The preterm gut microbiome has a simpler composition when compared to adults and consists of between 6-7 preterm gut community types (PGCT) (Brown et al., 2018, Stewart et al., 2016). Samples in these studies were divided into PGCT’s based on the abundance of a particular OTU.

Time (gestational age or postnatal age) has a major influence on the preterm gut microbiome (Brown et al., 2018, Zwittink et al., 2017, Korpela et al., 2018). At the genus level, Staphylococci and Enterococci dominate the first weeks of life (Zwittink et al., 2017, Korpela et al., 2018, Stewart et al., 2016). There is less agreement on how the preterm gut microbiome changes after this time. Korpela et al. described four phases of microbial development based on post-menstrual age, with phases dominated by the genera; *Staphylococcus*, *Enterococcus*, *Enterobacteriacae* and *Bifidobacterium* respectively (Figure 2) (Korpela et al., 2018). A similar pattern of development has not been described in such a clear manner elsewhere which may reflect differences in interpretation of the data or be due to true variation in the gut microbiota between different locations. The importance of the local environment as a variable has been highlighted as infants nursed together in the same room have been shown to share similarities in microbiome composition (Brooks et al., 2018). Environmental factors may be indirectly affected by location, through variation in clinical practice, including diet and delivery mode.

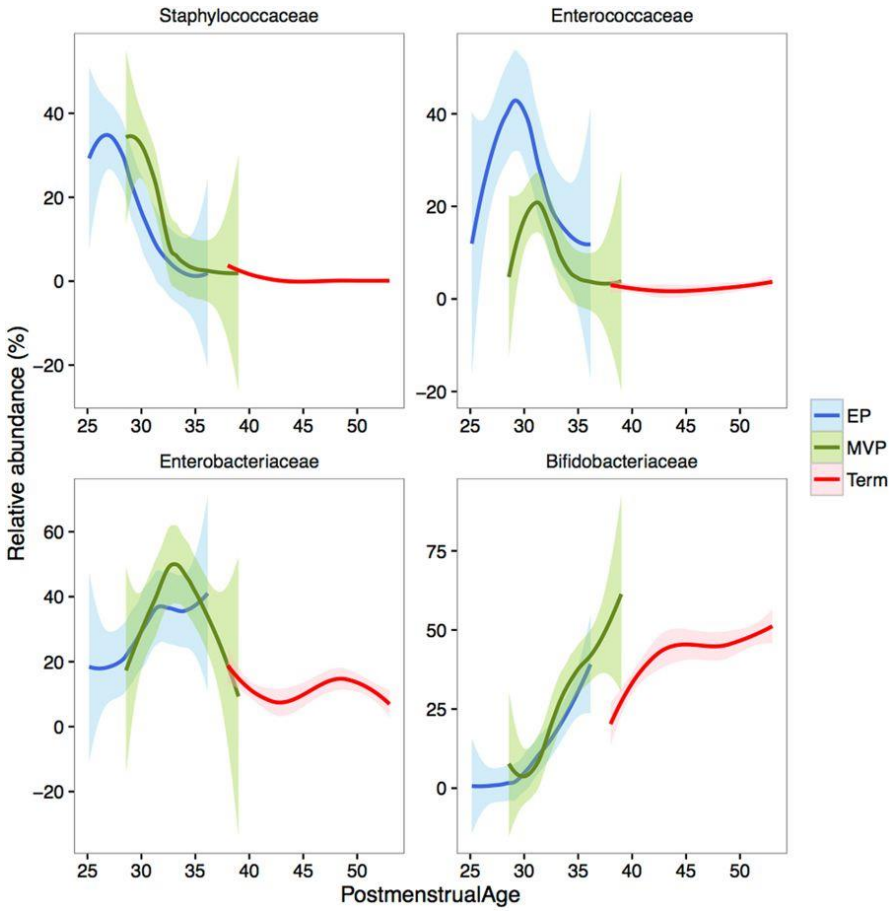


Figure 2 - Development of the most abundant bacterial genera in preterm and term-born infants with respect to postmenstrual age. The trend lines show the best-fit (second polynomial) and the shaded areas represent 95% confidence intervals. Figure taken from (Korpela et al., 2018)

EP (Extremely preterm), MVP (moderate to very preterm, 28-33^{+6} weeks GA)

1.2.3 *Effect of mode of delivery on the preterm gut microbiota*

1.2.3.1 Term infants

Delivery by caesarian section has been associated with an increased risk of childhood obesity and asthma in term infants (Keag et al., 2018). The mechanism by which this occurs is not known however the altered patterns of microbial colonisation may be important (Dogra et al., 2015). In term infants the gut microbiome is influenced by delivery mode to resemble the microbiome of the maternal skin (caesarean section) or vagina (vaginal delivery) (Dominguez-Bello et al., 2010). Delivery mode continues to have some influence on the composition of the gut microbiome throughout the first year of life (Stewart et al., 2018).

1.2.3.2 Preterm infants

Mode of delivery does not appear to have the same effect on the gut microbiome of preterm infants. Ho et al. described an increase in the bacterial class Gammaproteobacteria and phylum Firmicutes in vaginal and caesarean delivered infants respectively, however no difference in microbiota composition was found in larger cohorts (Zwittink et al., 2017, Stewart et al., 2017a, Ho et al., 2018). Interestingly, there has been a report of increased abundance of Staphylococci in vaginally born preterm infants which is contrary to the findings in term infants (Korpela et al., 2018). Most studies suggest any impact of mode of delivery upon the gut microbiota is not sustained beyond the first few weeks of life (Korpela et al., 2018, Stewart et al., 2017a, Ho et al., 2018). The limit of any effect of mode of delivery in preterm infants compared to term infants may be related to variation in environmental exposures, such as high rates of antibiotic exposure in preterm infants, in the first weeks of life.

1.2.4 *Effect of diet on the gut microbiota*

1.2.4.1 Term infants

The gut microbiota of term infants fed MOM contains an increased abundance of *Lactobacillus* and *Bifidobacterium* species compared to infants not fed MOM. MOM feeding has a major influence on the gut microbiota for more than 6 months after birth (Stewart et al., 2018). This finding correlates with the ability of Bifidobacteria to utilise HMO's as a substrate (Bode, 2012). MOM does not seem to have the same effect on the gut microbiota of preterm infants as it does on term infants (Underwood et al., 2017).

1.2.4.2 Preterm infants

In preterm infants the relative abundance of the genus *Bifidobacterium,* is decreased compared to term infants, even when supplemented by probiotics (Underwood et al., 2017), or in an RCT comparing two predominantly MOM diets (Butcher et al., 2018). However, an increase in Bifidobacteria abundance in MOM-fed compared with CMF-fed preterm infants has been demonstrated (Parra-Llorca et al., 2018). Studies by Cong et al. and Gregory et al. have conflicted with regards to the genera *Lactobacillus*, showing an increased relative abundance in CMF-fed or conversely MOM-fed preterm infants respectively, suggesting that any beneficial effects of MOM may not be specifically due to this group of bacteria (Cong et al., 2016, Gregory et al., 2016).

1.2.4.2.1 MOM vs CMF vs DHM

A more striking effect of diet appears to be the rate of change in diversity of bacteria seen when comparing a MOM and CMF diet. A predominantly MOM diet is associated with an increased initial α-diversity before a more gradual increase while α-diversity is initially lower then increases faster in preterm infants fed a CMF diet (Gregory et al., 2016). This suggests that microbial stability or rate of change may be a factor promoting health in MOM fed infants. Exclusive feeding with DHM appears to give a microbiome composition and diversity more similar to a MOM-fed than a CMF-fed preterm infant (Gregory et al., 2016, Parra-Llorca et al., 2018). However, neither of the studies by Gregory et al. or Parra-Llorca et al. reported using fortifier.

As described in Section 1.1.5 the normal situation in a NICU is a predominantly MOM diet with small amounts of DHM or CMF, with the introduction of a fortifier when an infant is fully enterally fed. There is a paucity of evidence exploring this normal situation.

1.2.4.2.2 Exploring the impact of the normal preterm infant diet on the gut microbiota

Butcher et al. have reported gut microbiome data from an RCT exploring a dietary intervention when there was a shortfall of MOM, however only infants who were exclusively MOM-fed were included in the microbiome analysis (Butcher et al., 2018). Cong et al. reported on a group of infants stratified into 6 feeding groups (MOM; MOM + DHM; MOM + CMF; DHM; CMF; DHM+CMF) based on the percentage of feeds the infant received in a 10-day period (Cong et al., 2017). They reported a higher α-diversity with MOM feeding, and that feeding type explained the greatest variance in β-diversity, with changes in relative abundance of bacterial species depending on feeding type (Figure 3). A limitation of the analysis as represented by Figure 3 appears to be the use of repeated measures for each infant as this

figure represents 389 samples from 33 infants. Infants have therefore been represented more than once in some columns, which is a limitation as an infant's sample is more likely to be similar to itself. However, the stratification into 6 feeding groups gives a closer representation of the variation in diet found in a typical NICU. After reviewing the literature, there are no studies assessing the effect of fortifier on the preterm gut microbiome.

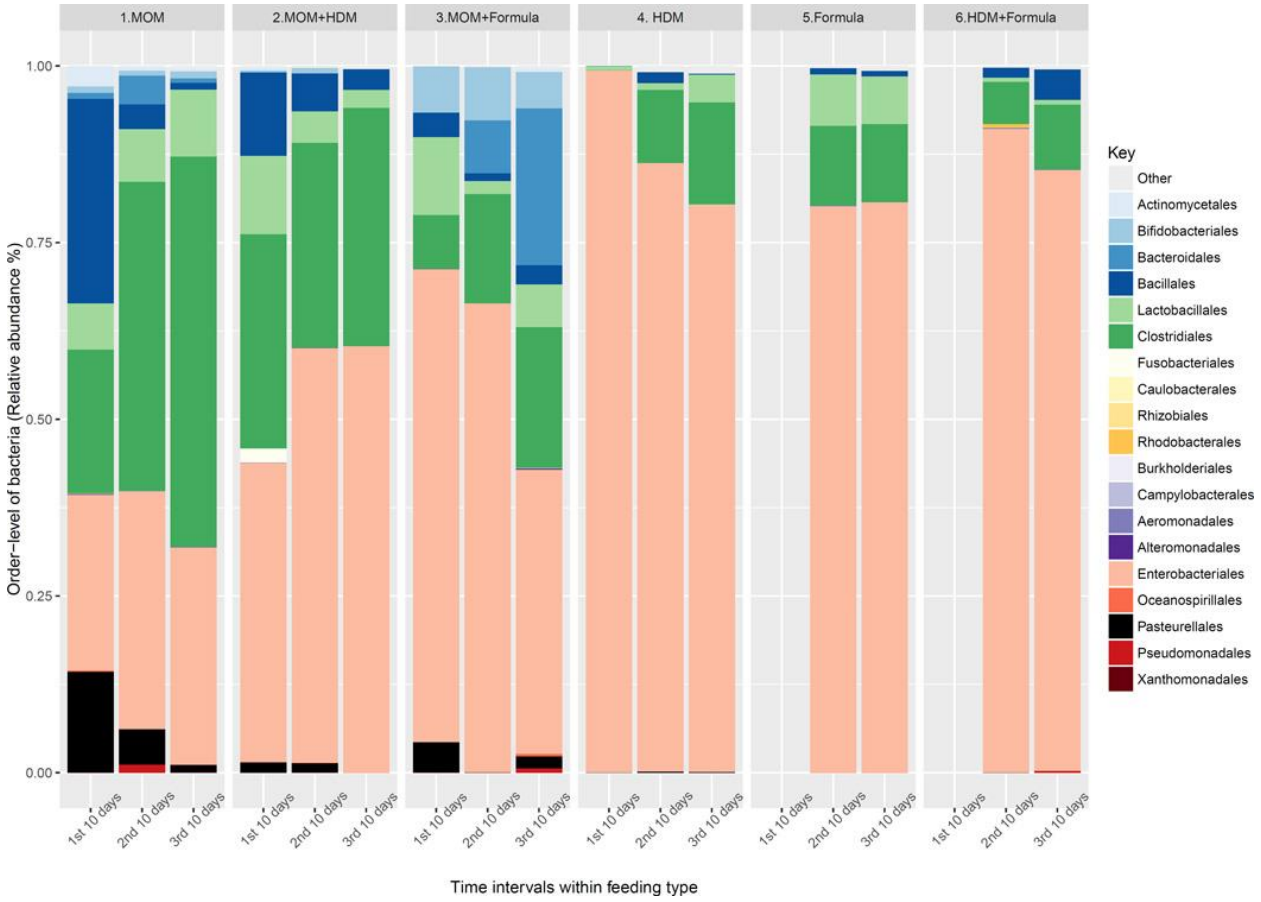


Figure 3 - Comparison of mean relative abundance of [OTUs within the] gut microbiome (order level) in preterm infants in the first 30 days of life. Samples stratified by feeding type in each 10-day period. 389 samples from 33 infants represented.

Figure 3 taken from (Cong et al., 2017)

1.2.4.3 Breast milk microbiome & colonisation of the preterm gut

As briefly mentioned In Section 1.1.6.3 breast milk has been shown to contain bacteria (Damaceno et al., 2017, Cacho et al., 2017) but it is unclear whether bacteria reach the lactating breast by ascent from the skin, the infant's oral cavity or translocation from the maternal intestine. Sample contamination appears unlikely as a microbiome has been identified both using sterile techniques (Sakwinska et al., 2016) and in the mammary glands of

non-lactating women (Urbaniak et al., 2014). In the latter study, breast tissue from planned operations was collected and bacterial DNA extracted sterilely. Following 16S analysis, individual bacteria were isolated and cultured proving their viability.

Studies looking at the breast milk microbiome have predominantly included mothers of term infants. Bifidobacteria have been shown to be the most abundant bacteria shared between the mother and infant in a study examining paired maternal milk and infant gut microbiome (Biagi et al., 2017). Interestingly, one study has shown that the milk microbiome has decreased α and β diversity, together with decreased abundance of a *Bifidobacterium* species, when the infant is indirectly breast fed (pumped or bottled breast milk) (Moossavi et al., 2019). Moossavi et al suggest this may be due in part to a breast pump biofilm.

This is relevant for preterm infants who are almost all exclusively fed by a gastric tube until at least 32 weeks corrected gestational age. This may explain the finding of decreased Bifidobacteria in the preterm gut compared to term infants mentioned in section 1.2.4.2 (Underwood et al., 2017). Feeding by naso-gastric tube may have other effects on the gut microbiota, as naso-gastric tubes removed from preterm infants have been shown to be colonised by large numbers of potentially pathogenic live bacteria even after use for only one day (Petersen et al., 2016).

1.2.5 *Intestinal maturation*

When considering the preterm microbiota, it is important to understand their intestine is not mature, and that bacteria have been demonstrated to aid that maturation.

The microstructure of the fetal intestine resembles the adult by 20 weeks gestation, with Paneth cells, Peyer's patches, goblet cells and all major epithelial cell lines (Colony, 1983, Montgomery et al., 1999). However, the maturation of the intestine, both in terms of structure and function, continues until adulthood (Montgomery et al., 1999). Important deficits in the preterm intestine include a deficiency in producing mucus, poorly regulated tight junctions and increased response to antigens. Immaturity is exacerbated by gut motility being impaired until approximately 36 weeks GA (Humphrey and Caud, 2018).

1.2.5.1 Mucus production

The mucus lining of the intestine functions to limit epithelial exposure to the luminal bacteria. The colonic mucus has two layers, a thicker outer layer, containing the majority of bacteria, and a watery inner layer where bacterial penetration is limited (Hooper et al., 2012). The mucus is produced mainly by goblet cells. Buisine et al. used in situ hybridisation to measure

the messenger RNA (mRNA) expression of eight mucin genes in fetal intestines from 8 to 27 weeks GA and adults. They found that mucin genes were expressed in different locations early in gestation and did not have an adult expression pattern until 23 weeks GA in the small intestine or 27 weeks GA in the colon (Buisine et al., 1998).

The ability of gut bacteria to influence mucus secretion has been demonstrated. In vivo, Mack et al. demonstrated through co-culturing *E. Coli* and a *Lactobacillus* species with intestinal epithelial cells (IEC's) that the *Lactobacilli* inhibited the bind of *E. Coli* to the IEC's. They demonstrated upregulation of mRNA for mucin 2 and 3, hypothesising that the action of *Lactobacillus* was through stimulating mucus production (Mack et al., 1999).

Whilst enhanced mucus production appears to be a desirable effect of bacteria, *Bifidobacterium* species have been demonstrated to break down mucin in an in vitro experiment (Ruas-Madiedo et al., 2008). The breakdown of the mucus barrier has been thought to be detrimental, with pathogenic bacteria such as *E. Coli* having an enhanced ability to do this (Cornick et al., 2015). However, the breakdown of mucin by Bifidobacteria has been shown to release oligosaccharides, which then influence the growth of other bacteria (Bunesova et al., 2018). This highlights that a similar action of different bacterial species in the intestine may have both detrimental and positive effects.

1.2.5.2 Intestinal permeability

Newborns show increased intestinal permeability in the first few days of life, this is thought to be prolonged to around 1 week in preterm infants based on data from a study measuring the urinary excretion of two sugars (mannitol and lactulose) that are not readily absorbed in the intestine, as a surrogate of intestinal permeability in infants of 34 weeks GA (Riezzo et al., 2009).

Bifidobacteria have been shown to modulate several gut proteins (claudin 4, occludin and zonulin) that are known to be important in maintaining tight junction integrity between IEC's in both a mouse model and in vitro (Bergmann et al., 2013, Ling et al., 2016). NEC was inhibited in the presence of Bifidobacteria in mouse models in both studies and was attributed to the effect on tight junction integrity.

1.2.6 *Medical interventions and the gut microbiota*

1.2.6.1 Probiotics

Probiotics have been shown to be beneficial in reducing the risk of NEC in preterm infants and are therefore used in NICU's (Chang et al., 2017, Dermyshi et al., 2017). A meta-analysis by Chang et al. included 7345 infants from 25 RCT's demonstrating a reduction in risk of NEC, with a pooled odds ratio of 0.6 (95% Confidence interval 0.48 - 0.74) (Chang et al., 2017). The effect of probiotics may be due to enhancing tight junctions between IEC's or by modulation of the microbiota as discussed in Section 1.4.1.2. A probiotic containing both Lactobacilli and Bifidobacteria led to an increase only in the relative abundance of the Bifidobacteria in the stool during administration, which persisted after administration (Abdulkadir et al., 2016).

1.2.6.2 Antibiotics

Antibiotics have a key impact on the preterm gut microbial landscape, with changes in both relative abundance and richness of bacterial species, especially with broad spectrum antibiotics such as meropenem (Gibson et al., 2016). This effect is however short-lived and not all infants exhibit the same response in gut microbiota to antibiotics (Korpela et al., 2018).

1.2.6.3 Iron

Iron supplementation is common in preterm infants due to their increased needs and irons low concentration in MOM. Whilst no studies have addressed associated gut microbial changes in a preterm population, iron supplementation is associated with an increased abundance of *Enterococcus* in the first year of life in term infants (Jaeggi et al., 2015). This would be consistent with iron being preferentially utilised by certain bacteria.

1.2.6.4 Multivitamins

As mentioned in Section 1.1.6.2, MOM is low in certain micronutrients, this means that **multivitamins** are commonly used on the neonatal unit. Specific bacteria are known to synthesise vitamins (LeBlanc et al., 2013) which may in turn influence the composition of the microbiota. There are no studies assessing the influence of vitamins on the gut microbiota in preterm infants. However, Talsness et al. performed an observational study involving qPCR on stools of 616 term infants, hypothesising that vitamin D supplementation (maternal or infant) would result in changes in abundance of gut microbial OTUs. There was no difference comparing infants who received or did not receive supplementation with vitamins A-D

(Talsness et al., 2017). Similarly, Sordillo et al postulated that maternal vitamin D exposure would lead to changes in the gut microbiome in a group of 333 infants. They found that vitamin D levels in umbilical cord blood were negatively associated with *Lactococcus* and positively associated with *Lachnobacterium* (Sordillo et al., 2017).

1.2.7 *NEC, LOS and the gut microbiota*

As mentioned in Sections 1.1.2 and 1.1.3, the gut microbiota of infants developing NEC and LOS differs from healthy controls. However, a limitation of comparing gut microbial changes, from stool, is the fact that these two diseases often cause an ileus leading to infrequent stooling, especially in severe disease, meaning that data is limited at a time of interest (Roze et al., 2017). Whilst multiple bacteria or bacterial patterns have been associated with NEC, these have not been consistent between studies (Morrow et al., 2013, Warner et al., 2016, Sim et al., 2015) however infants who display more fluctuation in their microbiome composition appear at increased risk of NEC (Stewart et al., 2016). It has also been suggested that immunoglobulin A (IgA) from MOM plays an important role. A study involving humans and mice demonstrated that the IgA unbound bacteria may be important in the development of NEC (Gopalakrishna et al., 2019) . This highlights an interaction between the diet, microbiota and immune system that is described in more detail in Section 1.4.2.2.2. As described in section 1.1.3, an increased abundance of the bacteria causing LOS has been demonstrated in the gut microbiota prior to disease onset (Stewart et al., 2017b, Carl et al., 2014).

1.2.8 *Influence of the gut microbiota on the host*

There is limited knowledge of how the microbial changes associated with LOS and NEC may be having detrimental interactions within the host. In order to investigate possible mechanisms, studies in preterm infants have focused on the interaction between the microbiome and metabolites, volatile organic compounds and immune cells.

1.2.8.1Metabolites

Whilst the inferred metabolomic potential of bacteria can be derived from microbiome datasets, this type of analysis is limited if using 16S due to the inability to determine the species or strain of bacteria present. One study quantifying faecal metabolites in infants with LOS demonstrated that metabolites correlating with *Bifidobacterium* were increased in control samples. Furthermore, when samples were categorised into PGCT's based on the abundance of different bacterial genera, the PGCT characterised by *Bifidobacterium* was never present in

LOS infants (Stewart et al., 2017b). In a subsequent study, Stewart et al. demonstrated five metabolites which increased in abundance prior to the onset of NEC, and then decreased following disease (Stewart et al., 2016). Wandro et al. performed a similar study and did not find any gut metabolites associated with NEC, however they were limited by having only three NEC cases in a group of 32 infants (Wandro et al., 2018).

1.2.8.2Volatile organic compounds (VOC)

VOC are the gaseous carbon-based compounds produced following metabolic or physiological processes in the body. Microbial VOC often have distinctive odours and are measurable by flowing the ionised VOC through an electric field and separating them based on their ionic charge (Berkhout et al., 2019). Infants who subsequently develop NEC or LOS show an alteration in VOC profiles 2-3 days before clinical diagnosis compared to controls (Berkhout et al., 2019, de Meij et al., 2015). No difference was found when comparing the VOC profiles of preterm infants fed either CMF or MOM at specific time-points (El Manouni El Hassani et al., 2018).

Should clear perturbations in VOC be found in health or disease, VOC analysis offers the potential for a point of care diagnostic tool through the use of an electronic nose (Wilson, 2018). However, due to their volatile nature, VOC are difficult to measure. Furthermore, VOC analysis requires comparison between two groups to define normality or abnormality. The development of a normal comparison group would likely have to be NICU site-specific due to the variation in gut microbiota as described above.

1.2.8.3 Immune system

There are limited data exploring interactions between the microbiome and immune system in preterm infants. However, a study using mass cytometry and 16S on the blood and stool respectively of preterm infants found that those with an altered gut microbiome early in life had a perturbed immune system development at 3 months (Olin et al., 2018). Section 1.3 explores further what is known about the T lymphocytes in preterm infants, and what can be derived from laboratory studies regarding the interaction between the microbiome and T lymphocytes. The term T cells will refer to T lymphocytes in the remainder of this thesis.

1.2.9 *Areas of uncertainty regarding the preterm gut microbiome*

The preterm gut microbiome is mainly influenced by age, diet and antibiotic administration and differs systemically from that of a term infant, likely due to variations in their

environmental exposures (i.e. relatively sterile NICU vs home). An assumption of infant gut microbial analysis is that the pattern seen in MOM-fed infants is optimal, therefore the aim for dietary interventions is for resemblance to this microbial pattern. Whether the preterm infant's diet consists solely of MOM, DHM, or CMF produces differences in both composition and diversity of the gut microbiome. Feeding with a solely DHM diet leads to a gut microbiome more closely resembling a MOM diet than a CMF diet.

However, there are few mechanistic studies exploring current feeding strategies and hence able to support typical clinical decision-making with regards to feeding, namely:

1. What should be given should there be a shortfall of MOM; decision in first few days?
2. If a fortifier is deemed necessary for weight gain, should this be derived from human (HMF) or cow's milk (BMF); decision following establishment of MOM feeding?

As described in Section 1.1.5, the diet of preterm infants is predominantly MOM, with RCT's exploring diet commonly reporting more than 70% MOM diet (Corpeleijn et al., 2016, Sullivan et al., 2010). It is therefore important to explore this situation and establish if DHM can be shown to have an influence. Should changes in the gut microbiome be identified, correlation with changes in immune or metabolic function would suggest a possible mechanism and increase confidence in a real effect on the infant.

1.3 T Cells in preterm infants

The immune system of preterm infants differs from that of adults, being generally hyporesponsive. T cells are a population of lymphocytes that develop in the thymus, before further differentiation in the periphery. T cells are an important component of the adaptive immune response, although some T cells display innate properties. T cells are of particular interest when considering the newborn, who is exposed to many potential pathogens soon after birth. At birth, infants are thought to be antigen naïve leading to an increased reliance on the innate immune system, however some T cells in the fetal blood and intestine have been shown to have a memory phenotype raising the possibility of fetal antigen exposure (Schreurs et al., 2019, Li et al., 2019, Zhang et al., 2014).

A challenge of studying T cells in preterm infants is the difficulty of obtaining sufficient samples of a suitable tissue, commonly peripheral blood. Umbilical cord blood is often available in sufficient volumes and has been taken to represent the immune system of preterm infants, however a systems level analysis by Olin et al., revealed it to be a poor surrogate for peripheral blood. This included when umbilical cord blood was compared to blood taken on the day of birth suggesting that umbilical cord blood is more representative of the placental environment. The same analysis showed that the immunophenotype of leucocytes in preterm infants differs from term infants after birth with a convergence over the first months of life, highlighting the importance of early life exposures (Olin et al., 2018).

The next section is a description of how immune cells can be identified and categorised. I will then focus on fetal T cell development before describing what is known about the T cell compartment of preterm infants. Based upon laboratory and animal studies I will then describe what is known about the interaction between diet, gut microbiome and T cells.

1.3.1 *Identification of immune cells*

1.3.1.1 Categorisation

Immune cells are distinguished by a combination of appearance, function and expression of cell surface proteins. Over the last 40 years a large number of cell surface proteins have been identified and given cluster of differentiation (CD) numbers, allowing for the standardised description of immune cells based on their expression of combinations of these proteins. For example, T cells are CD45+CD3+, and further split into two major categories of CD4+ T cells and CD8+ T cells. Proteins expressed intra-cellularly are given names but not CD numbers, such

as the transcription factor, T-bet. The function of a cell can be marked by the expression of cytokines (e.g. Interleukin (IL) 4, interferon-γ (IFN-γ)). Cells typically require stimulation in the laboratory in order for cytokine expression to be detected.

Several techniques can be used to measure the protein expression of cells including Western Blots, ELISA, immunofluorescence and flow cytometry. These techniques all depend on the recognition of specific antigens and are limited by the number of different antibodies that can be detected within any individual assay. In the last 20 years technologies have been developed that allow for a greater number of gene products to be detected simultaneously. These include techniques that identify cells based on their genetic material (e.g. transcriptomics) or antibody labelling with rare heavy metals (e.g. mass cytometry). Flow cytometry and mass cytometry are two commonly used techniques for measuring protein expression on a single cell basis, however both have limitations.

1.3.1.2 Flow Cytometry

Flow cytometry describes a technique to analyse cells in suspension on a single cell level. A typical experiment will involve a sample of cells being stained with antibodies that have been tagged with a variety of known fluorochromes. A suspension of these cells is run through a flow cytometer. Cells can initially be identified on the basis of light scattering at two angles, with one perpendicular to the other, named forward scatter and side scatter. Forward scatter, when the light is behind the cell, gives an idea of cell size. Side scatter, which is perpendicular to this, gives an idea of cell granularity. In this way, lymphocytes, granulocytes and debris can be easily separated. Furthermore, lasers can be used to identify cells by the emission, or lack of emission, of fluorescent light. This technique can allow for the identification of over 20 antibodies in a cell suspension, however a typical experiment might only involve 5-10 antibodies. The two main limitations of this technique are overlap between different spectra of the light emitting fluorochromes therefore making analysis difficult, and as a result limiting the number of antibodies that can be analysed at any single time-point. These limitations have been overcome to a degree by the introduction of increasingly sophisticated fluorochromes and instrumentation.

1.3.1.3 Mass Cytometry

Mass cytometry combines a flow cytometer with a time-of-flight mass spectrometer. A sample is stained with antibodies tagged with rare heavy metals (of the lanthanide group). The sample is then introduced into the instrument and nebulised, using an inert gas (argon), into a single cell suspension. The cells are vaporised with an argon laser creating an ion cloud. Light

molecules (such as oxygen and nitrogen) are removed by a quadrupole leaving the heavy metal ions. Ions are projected and their molecular weight calculated based on their time-of-flight. A cell is identified based on the presence of their signature of heavy metal ions. This technique has the advantage compared to flow cytometry of the possibility to detect over 100 antibodies; but currently a panel of around 40 antibodies is achievable due to the difficult in obtaining lanthanide metals and maintaining purity. Mass cytometry provides less spillover between channels giving cleaner data than conventional flow cytometry. Importantly, mass cytometry does not support the measurement of side or forward scatter so cell populations are identified based on the presence of cell surface or intracellular proteins. The limitations include that the signal, in comparison to a fluorochrome, is much weaker and there are challenges with analysing the resultant large datasets. Furthermore, cells are destroyed as part of the detection process, and data acquisition is currently slow (Gadalla et al., 2019).

1.3.1.4 Analysis

Techniques such as mass cytometry produce huge datasets creating difficulties in analysis. Standard techniques for analysing flow cytometry data involve the identification of populations that are positive or negative for a particular antibody by successive gating (often presented in a dot plot). The main limitation in this approach is bias in the choice of which antibodies to gate on, which could potentially miss important information. The emergence of mass cytometry has led to the use of statistical techniques that allow for visualisation and comparison of a whole population of cells (Olsen et al., 2019, Bendall et al., 2011). The techniques used to define a population involve less bias, however, do require knowledge to gain meaningful conclusions from the data and are dependent on users having bioinformatic expertise.

1.3.2 *T-Cell development*

1.3.2.1 Thymic development

Lymphocytes (B, NK and T cells) have common progenitors in the bone marrow. T progenitor cells exit the bone marrow into the vasculature, and enter the thymus around 8 weeks gestation, where they are then termed thymocytes. This has been demonstrated by the presence of thymocytes in fetuses at 8 but not 7 weeks gestation (Haynes and Heinly, 1995). T cell maturation in the thymus is described related to CD4 and CD8 expression. Thymocytes are initially double negative (CD4-CD8-) before becoming double positive (CD4+CD8+), and finally single positive (CD4+CD8- or CD4-CD8+) T cells (Shah and Zuniga-Pflucker, 2014). They must go

through a process of positive and negative selection. Positive selection is to ensure that the T cell receptor (TCR) is able to bind to a major histocompatibility complex (MHC) molecule, and therefore support tonic TCR signalling. If this is not the case, the cell undergoes apoptosis. Negative selection ensures that the TCR does not bind too strongly to MHC molecules as this could lead to self-reactivity as seen in auto-immune conditions. T cells predominantly display a TCR with both α and β chains however approximately 5% of T cells display γ and δ chains of TCR and are known as γδ T cells. The double negative stage of thymic development is when the thymocytes undergo rearrangement of the TCR gene loci to either become αβ or γδ T cells (Krueger et al., 2017). Whilst early thymocytes do not display the full range of TCR, by 16 weeks gestation the full range is established (Michaelsson et al., 2006). T cells are present in the fetus as early as 8 weeks' GA and fetus-derived T cells have the capacity to proliferate and produce cytokines (Michaelsson et al., 2006). In order to proliferate, T cells need to bind their TCR, as well as receive secondary activation through binding other molecules. CD28 is an example of a secondary activation molecule of CD4+ T cells.

1.3.2.2 T cell migration

T cell migration to the peripheries has been confirmed with identification of T cells in the mesenteric lymph nodes as early as 12 weeks' GA (Michaelsson et al., 2006) and the spleen and intestinal mucosa as early as 14 weeks' GA (Darrasse-Jeze et al., 2005, Li et al., 2019, Schreurs et al., 2019).

The destination of the T cell is influenced by chemokines. Chemokines are soluble signaling proteins that guide immune cells to different locations in the body. Chemokine receptors (CCR) on T cells provide information to guide the tissue homing of that cell. Broadly, the intestine-homing chemokine receptor is CCR9, the skin homing chemokine receptor is CCR4 whilst the secondary lymphoid tissue homing receptor is CCR7. CCR4 is also associated with a T_h2 response whilst another chemokine receptor, CXCR3, is associated with a T_h1 response (see Section 1.3.4) (Bonnechi et al., 1998).

1.3.2.3 Maturation

T cells that egress from the thymus are naïve. Upon recognition of their cognate antigen presented by antigen presenting cells (APC's), T cells proliferate and may develop effector function. A population of T cells then go on to develop memory function in case of future antigen exposure. The maturity of T cells can be identified by the expression of cell surface antigens. CD45 has two major isoforms, CD45RA is expressed by naïve T cells whilst CD45RO is expressed by memory or effector T cells. Furthermore, CCR7 can be used to categorise CD45RA

negative cells as either effector memory T cells (CCR7-) or central memory T cells (CCR7+) (Romero et al., 2007)(Figure 4). Central memory T cells reside in the secondary lymphatic tissue, hence the expression of CCR7, whilst effector memory T cells circulate in the blood and enter the tissues. CD8 T cells can differentiate further into a population that is CD45RA+ and CCR7-, termed T_{emra} (effector memory RA+ T cells). These cells lack the capacity to persist long-term but have increased capacity to secrete perforin/granzyme B, cytolytic proteins associated with cytotoxic function (Sallusto et al., 1999).

Unsurprisingly, the majority of T cells derived from cord blood in preterm or term infants are naïve (Quinello et al., 2014). Using CD45RA as a marker of naivety, the proportion of naive T cells has been shown to remain high (median 85%) at 6-8 weeks of age, although significantly less than term infants at the same postnatal age (Berrington et al., 2005).

1.3.3 *Compartmentalisation of T cells*

A feature of immune cell populations is that their composition and function may vary depending on tissue location, including term cord blood. The expression of cell surface markers and cytokines is distinct for CD4+ and CD8+ T cells in each tissue (Wong et al., 2016, Thome et al., 2016).

This is highlighted by findings in a study on human fetal tissue. Fetal CD4 T cells in the intestinal lamina propria but not spleen or liver were shown to have a predominantly memory phenotype (Schreurs et al., 2019, Li et al., 2019). This suggests prior antigen exposure and these cells were shown to have the capacity to produce a broad cytokine repertoire upon stimulation (Li et al., 2019). Fetal mesenteric lymph node (MLN) derived T cells have been shown to proliferate and secrete IFN-γ. Both proliferation and cytokine production are increased with the removal of regulatory T cells (T_{reg}s) (Michaelsson et al., 2006).

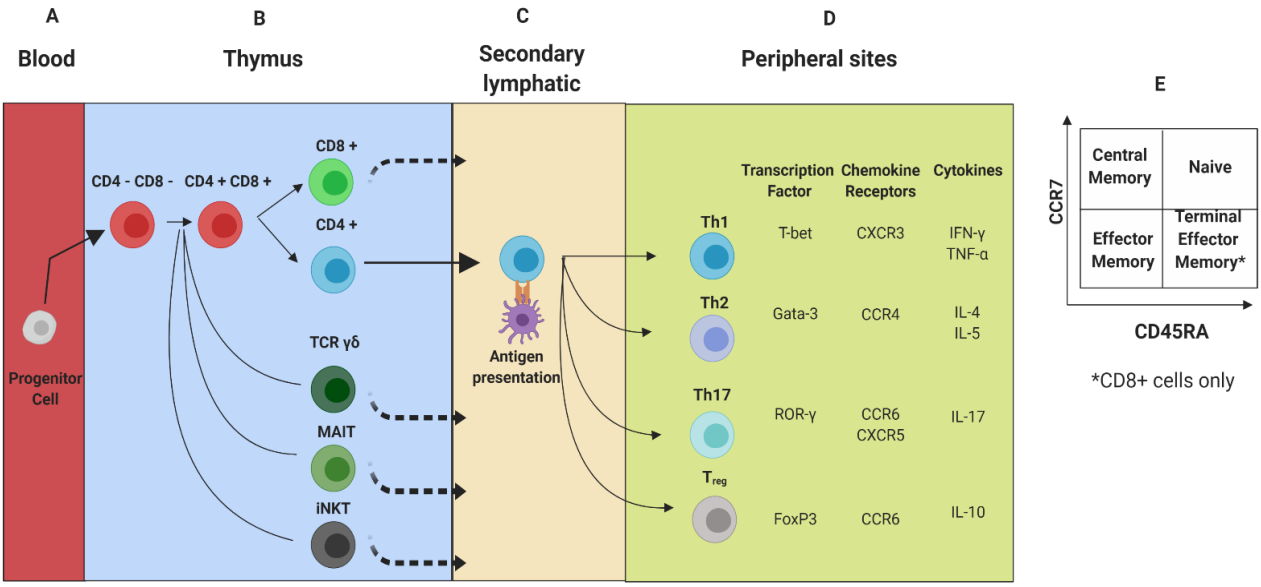


Figure 4 - Graphical representation of T cell development(A-C), differentiation of T_h cells following antigen presentation(D) and a gating strategy to define maturation of T cells (E)

1.3.4 *CD4+ T cells (T-helper (T_h) cells)*

Naïve T cells in the peripheral blood of infants are predominantly CD4+ and hence recognise peptide antigen in the context of class II MHC. CD4+ T (T_h) cells play a central role in the adaptive immune system by supporting the function of other immune cell populations through cytokine secretion, as well as regulating immune responses. Preterm infants have a decreased number and abundance (as percentage of lymphocytes) of T_h cells at 6 weeks compared to term infants (Berrington et al., 2005).

1.3.4.1 The T_h Response

Naïve T_h cells can differentiate based on the cytokines to which they are exposed. Interleukin (IL) -12 leads to the promotion of a T_h1 response. IL-4 exposure leads to a T_h2 response, Transforming Growth Factor – beta (TGF-β) induces T_{reg}s whilst TGF-β, IL-16 and IL-23 lead to a T_h17 response. Furthermore, a T_h19 response is elicited by IL-19 and a T_h22 response is elicited by IL-22 (Raphael et al., 2015). Lastly, T follicular helper cells (Tfh), which are important for the development of germinal centres in secondary lymph nodes, are promoted by exposure to IL-21 (Eyerich and Zielinski, 2014). Each T_h response results in different actions through cytokine secretion, for example a T_h1 response drives cell mediated immunity whilst a T_h2 response helps B cells produce antibodies. T_h subsets are ideally identified by the cytokines the cells

produce. However, T_h subsets also show distinctive patterns of expression of various transcription factors and chemokine receptors. Figure 4 lists the main ways of identifying a T_h1, T_h2, T_h17 or T_{reg} population based on protein expression (Eyerich and Zielinski, 2014, Raphael et al., 2015).

Whilst T_h1 and T_h2 responses are mutually exclusive there does appear to be overlap between T_h1 and T_h17 responses as well as the T_h2 response and T_{reg}s (Kunicki et al., 2018). Importantly, an excessive T_h response is associated with the development of disease. An excessive T_h1 response can lead to tissue damage whilst an excessive T_h2 response can result in atopy (Berger, 2000). In contrast, an inadequate T_h response can result in disease, an infant's predisposition to viral infections has been related to the inability of T cells to secrete sufficient IFN-γ (Melville and Moss, 2013).

1.3.4.2 T_h2 skewing in preterm infants

Preterm infants have been suggested to have a skewed T_h2 response based on increased production of IL-5 (T_h2 cytokine) and decreased production of IFN-γ (T_h1 cytokine) upon stimulation of peripheral blood compared to term cord blood (Dirix et al., 2013). The skewed T_h2 response appears to start in utero and may be driven by antigen-presenting cells (APCs) as T cells cultured with fetal APCs have been shown to produce significantly more IL-4 (T_h2 cytokine) in comparison to T cells cultured with adult APCs (McGovern et al., 2017). This skewed T_h2 response is possibly a protective mechanism to limit reaction to maternal antigens. Pregnant women have similarly been found to have a skewed T_h2 response, due to suppression of the T_h1 response which is felt to be important to prevent fetal rejection (Sykes et al., 2012).

Postnatally, the ability of preterm T_h cells to secrete IFN-γ remains low, however αβ T cells are able to secrete a significantly increased amount of IL-2 (T_h1 cytokine) at 1 month compared to birth (Gibbons et al., 2009). This is consistent with single cell RNA-seq analysis comparing peripheral blood mononuclear cells (PBMC's) of term and preterm infants at 12 weeks of age demonstrating upregulation of genes that suppress IFN-γ (Olin et al., 2018).

1.3.4.3 IL-8 Secretion

An effector function of T cells in newborns that is distinct from adults is a significantly increased ability to produce IL-8/CXCR8 upon stimulation (Gibbons et al., 2014). IL-8 is a cytokine associated with myeloid and endothelial cells and is involved in neutrophil migration towards a site of inflammation (Gibbons et al., 2014). Olin et al. found an upregulation of IL-8

transcripts in preterm compared to term infants blood at 12 weeks age, as well as an increased plasma concentration of IL-8 in early preterm peripheral and cord blood, compared to early term or later preterm blood (Olin et al., 2018).

Pekalski et al. demonstrated that recent thymic emigrants (RTE's) derived from term cord blood expressed more IL-8 than older infants (Pekalski et al., 2017). Scheible et al. explored RTE's further demonstrating that both the population of CD4+ RTE's and the production of IL-8 from T_h cells correlated positively with GA at birth. They demonstrated that IL-8 production was increased in RTE's compared to the rest of the T cells (Scheible et al., 2018).

Combined, these data suggest that newborn infants' T cells are able to secrete more IL-8 compared to adults and there is an increased concentration in the blood of preterm infants in the first months of life. These data suggest an important role of T cells in innate immunity in preterm infants.

1.3.4.4 Intestinal T_h population

A skewed T_h2 response in the peripheral blood contrasts with the environment in the fetal intestine. Schreurs et al. have shown that fetal gut T_h cells have an increased tendency to secrete tumour-necrosis factor α (TNF-α) and IL-2, cytokines associated with a T_h1 response, when compared to infant intestinal samples. TNF-α was shown, in a human fetal organoid model, to be important in intestinal epithelial growth via its effects on intestinal stem cells, however high levels of TNF-α suppressed epithelial growth. Furthermore, this study demonstrated an increased production of TNF-α from T_h cells in intestinal samples of infants with NEC compared to controls (Schreurs et al., 2019). McGovern et al. (2017) explored the interaction between APC's and T cells in fetal life. They demonstrated that TNF-α expression by T cells was reduced in fetal spleen compared to adult samples. They attributed this effect to the expression of arginase-2 by fetal T cells, independent of Toll-like receptor (TLR) stimulation. The exact mechanism by which arginase-2 expression increased was not found, however arginase is known to deplete the tissue of L-arginine which is essential for TNF-α production (Morris, 2010). The expression of TNF-α was increased with the introduction of arginase inhibitors to the experiment (McGovern et al., 2017).

Together, this work highlights the compartmentalisation of T cells and suggests that fetal intestinal T_h cells may influence intestinal mucosal integrity. Excessive TNF-α production by preterm T cells may be implicated in the pathogenesis of NEC. Should this be proven, modulation of L-arginine levels may be a potential therapeutic intervention, through its alteration of TNF-α signalling.

1.3.5 *Regulatory T cells (T_{reg}s)*

Regulatory T cells (T_{reg}s) are defined by their suppressive function towards effector T cell responses. They are characterised by expression of CD4, high levels of CD25, low levels of CD127 and the transcription factor FOXP3. T_{reg}s are known to develop in the thymus as well as being induced in the peripheral immune system, with different transcriptome signatures respective of location (Miragaia et al., 2019). T_{reg}s comprise approximately 8% of the T_h cells in peripheral blood of preterm infants, this is significantly more that term infants and adults and is inversely correlated with GA (Pagel et al., 2016, Zahran et al., 2019). T_{reg}s help regulate the immune response through suppression, which could be particularly important in early life when the newborn is exposed for the first time to a large number of bacterial antigens. These exposures could otherwise lead to an excessive immune response, which in turn would cause tissue inflammation and potentially gut failure.

1.3.5.1 Fetal development of T_{reg}s

In humans, T_{reg}s have been identified in the thymus as early as 13 weeks gestation and in the periphery (spleen) from 14 weeks gestation (Darrasse-Jeze et al., 2005) The proportion of T_{reg}s in the fetal MLN's has been shown to be significantly increased compared to adults (Schreurs et al., 2019, Michaelsson et al., 2006).

1.3.5.2 Suppressor function of T_{reg}s

Whilst the population of circulating T_{reg}s is increased in preterm infants compared to term infants, they are predominantly naive (Dirix et al., 2013, Zahran et al., 2019). The PBMC's of preterm infants show transcriptional up-regulation of genes associated with IL-10 secretion compared to term infants (Olin et al., 2018). IL-10 is a known mediator of the suppressive function of T_{reg}s as well as inducing their expansion (Raphael et al., 2015). Michaelsson et al. demonstrated this with fetal T_{reg}s: in culture, stimulated T cells were able to proliferate and increase cytokine production after removal of T_{reg}s, confirming a tonic suppressor effect (Michaelsson et al., 2006)

In contrast to the blood and MLN's, the abundance of T_{reg}s in the fetal lamina propria is reduced compared to infants (median age 4 months) but it is not known if this in true in preterm infants. Interestingly the ability of fetal lamina propria T_{reg}s to produce IL-10 is greatly reduced compared to infants, and is absent in the presence of NEC (Schreurs et al., 2019).

Together, this suggests that T_{reg}s play an important role in the preterm immune system including through their secretion of IL-10, however whilst the population is expanded in preterm infants, they may have impaired effector function.

1.3.6 *CD8+ T cells (Cytotoxic T(T_c) cells)*

T_c cells represent approximately 30% of the peripheral T cell population in the preterm infant (Berrington et al., 2005). They display mainly a naive phenotype in the first 12 weeks of life (Walker et al., 2010). The absolute T_c cell population is significantly expanded compared to adults but similar to term infants (Olin et al., 2018, Scheible et al., 2015).

1.3.6.1 Fetal development and effector function of T_c cells

T_c cells begin to exit the thymus between 7 and 16 weeks GA and have been demonstrated in the fetal intestine as early as 16 weeks GA (Zhang et al., 2014, Michaelsson et al., 2006).

There is little known about the effector function of T_c cells in peripheral blood of preterm infants. However, using cord blood-derived T_c cells from varying gestations (23-41 weeks GA), Scheible et al. demonstrated there is an increased ability to secrete IFN-γ, TNF-α and IL-2 at earlier gestations in response to a cognate antigen. However, they found that the significance of IFN-γ was lost when correcting for prolonged rupture of membranes in the infants' mothers. The authors suggested the finding of enhanced effector response was due to a pro-inflammatory state at earlier gestations (Scheible et al., 2015).

1.3.6.2 Newborn T_c cells have a distinctly different phenotype to adults

Using a transcriptome analysis of the cord blood from term infants, Galindo-Albarran et al. were able to demonstrate that T_c cells from cord blood had a completely different phenotype to adults. The cord blood had upregulation of genes associated with innate immunity, viral infection and cell cycle processes, whilst adult blood upregulated genes were associated with cytotoxicity and effector function (Galindo-Albarrán et al., 2016). Functionally, however, fetal (term and preterm cord blood) and adults have been shown to develop a similar memory T_c cell population that is able to produce perforin and cytokines in response to Cytomegalovirus (CMV) infection (Marchant et al., 2003).

1.3.7 *Gamma-delta (γδ) T cells*

γδ T cells comprise 4-10% of T cells in the peripheral blood of adults compared to 1-3% and <1% of T cells in the cord blood of term and preterm infants respectively (Li et al., 2013). They

are distinct from αβ T cells in that they show reduced clonal diversity of an alternative antigen receptor that is not restricted by classical MHC molecules. γδ T cells seem particularly important early in life. They have an innate-like ability to become activated and display effector function without prior clonal expansion (Vantourout and Hayday, 2013).

1.3.7.1 Maturity and effector function of γδ T cells

Whilst γδ T cells are predominantly naïve in the cord and peripheral blood of preterm infants, they develop a memory (CD45RO+) phenotype in the first month of life (Gibbons et al., 2009). Unlike αβ T cells, a large number reside in non-lymphoid tissues including the gut, spleen and lungs.

The ability of peripheral blood γδ T cells from preterm infants to secrete IFN-γ and IL-10 on stimulation with PMA and ionomycin is increased compared to both γδ T cells in term infants and αβ T cells (Gibbons et al., 2009). However, upon influenza virus stimulation, cord blood γδ T cells were found to have a decreased ability to proliferate and produce IFN-γ compared to term infants and adults (Li et al., 2013). There is known to be interplay between T_{reg}s and γδ T cells through mouse models of intestinal inflammation: mice deficient in T_{reg}s have intestinal inflammation mediated by γδ T cells which is suppressed following the introduction of T_{reg}s (Yurchenko et al., 2011). Furthermore, γδ T cells have been shown to have important roles in responding to chronic bacterial as well as viral infections (Zhao et al., 2018)

Li et al. used cord blood for their analysis whilst Gibbons et al. used peripheral blood following birth from preterm infants, so it may be that IFN-γ production improves with age. This suggests that preterm γδ T cells may have an important role in T cell response to viruses and immune regulation, due to changes in their ability to secrete IFN-γ and IL-10 respectively.

1.3.8 *Invariant natural killer T (iNKT) cells*

Natural killer T (NKT) cells are a population of T cells that bear TCR's restricted by the non-classical MHC molecule CD1d, and express cell surface antigens associated with NK cells. There are two main populations of NKT cells, type 1 NKT (or iNKT) cells express the invariant TCR (Vα24-Jα18) whilst type 2 NKT cells express a broader range of TCR's. Type 1/invariant NKT cells have been of particular interest, partly as it is difficult to conclusively identify type 2 NKT cells (Balato et al., 2009).

1.3.8.1 Abundance of iNKT cells

iNKT cells are rare and vary hugely in abundance depending on tissue. iNKT cells comprise just 0.01-0.1% of lymphocytes in peripheral blood but 10% of lymphocytes in the omentum (Berzins et al., 2011). There is an increased abundance of iNKT cells in preterm cord blood compared to term cord blood (Ladd et al., 2010).

1.3.8.2 Effector function and association with disease

iNKT cells are able to secrete cytokines that mediate T_h1, T_h2 and T_h17 responses, as well as recognise specific glycolipid antigens, which are presented by a MHC-1 like molecule, CD1d (Moreira-Teixeira et al., 2011). In mice, the alternative cytokine secretion patterns are reflected in the nomenclature for iNKT subsets iNKT1, iNKT2 and iNKT17 cells by analogy with T_h1, T_h2 or T_h17 responses(Krovi and Gapin, 2018, Crosby and Kronenberg, 2018). The ability of preterm iNKT cells to produce these cytokines is not known, however term infant cord blood iNKT's are able to secrete IL-17, IL-4, IL-10, but not IFN- γ. Their secretion of IL-10 is particularly increased compared to adult PBMC's suggesting a regulatory role (Moreira-Teixeira et al., 2011).

iNKT cells have been of particular interest due to their association with autoimmunity, colitis and cancers (Berzins et al., 2011). In a mouse model, mice that were bred to be deficient in iNKT cells have been shown to have a predisposition to cancer (Swann et al., 2009). Whilst deficiency of iNKT cells may lead to disease, an experimental colitis has been shown to be driven by iNKT cells producing IL-13 (Heller et al., 2002). Subsequently, it has been demonstrated that germ-free mice have an increased proportion of iNKT cells in their colon associated with increased susceptibility to colitis, however colonisation of the intestine with bacteria early in life leads to a decreased iNKT population and provides protection from colitis (Olszak et al., 2012). Together, these studies suggest a regulatory role of iNKT cells which is influenced by environmental factors.

1.3.9 *Mucosa-associated invariant T (MAIT) cells*

MAIT cells are a population of T cells predominantly found in the lung and intestinal mucosa. They develop in the thymus before post-thymic maturation in the mucosal tissues and liver (Leeansyah et al., 2014). Similar to iNKT cells, they recognise non-peptide antigens presented in the context of a non-classical major histocompatibility complex (MHC) 1 molecule, MR1. MAIT cells represent less than 1% of peripheral blood T cells in infants (term or preterm) compared to up to 10% in adults (Ben Youssef et al., 2018).

MAIT cells are of particular interest as MR-1 presents microbial derived metabolites of riboflavin (Le Bourhis et al., 2011), which is an essential vitamin found in many foods. Riboflavin is present in human milk but the concentration varies depending on maternal intake (Hampel et al., 2017). Furthermore, multi-vitamin supplements commonly given to preterm infants contain riboflavin. Riboflavin is synthesised by over 80% of gut commensal bacteria, preferentially the bacterial phyla Proteobacteria and Bacteroidetes (Tastan et al., 2018). The preterm gut microbiome has a very low abundance of Bacteroidetes making it likely that Proteobacteria species are the predominant bacteria involved in riboflavin synthesis in preterm infants.

MAIT cells are CD4- T cells identified by the expression of CD161 as well as the semi-invariant TCR (Vα 7.2 – Jα3.3/20/12 (Vα 7.2)) which binds the MR-1 tetramer. The expression of the transcription factor PLZF, CD45RO and the β chain of the CD8 TCR are associated with a mature MAIT cell state and effector function (Ben Youssef et al., 2018). MAIT cells are described as innate-like as they do not require repeated antigen exposure to display effector functions (Ben Youssef et al., 2018).

1.3.9.1 Fetal development of MAIT cells

MAIT cells, as defined as CD45+CD3+CD161+Vα 7.2+, were demonstrated by Leeansyah et al. to be present in the fetal thymus between 18 to 23 weeks gestation. In the thymus, MAIT cells do not express the activation marker, PLZF. PLZF induces proliferation of NKT cells and is expressed in MAIT cells derived from the fetal intestine, MLN and spleen suggesting activation in the periphery. Supporting this, fetal intestinal, liver and lung but not thymic or spleen MAIT cells were shown to produce IFN-γ following stimulation (Leeansyah et al., 2014).

1.3.9.2 MAIT cell maturation over the first months of life

The dramatic difference between the abundance of MAIT cells in infants compared to adults suggests a postnatal expansion which happens over years. Whilst the expansion does not occur quickly, MAIT cells develop a mature phenotype, defined by the expression of CD45RO and the β-chain of the CD8 TCR in the first 2 months of life (Ben Youssef et al., 2018, Walker et al., 2014).

1.3.9.3 Population expansion and effector function

Upon recognition of a cognate ligand MAIT cells are able to produce inflammatory cytokines (IFN-γ, TNF-α and IL-17) as well as cytotoxic molecules such as granzyme B and perforin in

adults (Ghazarian et al., 2017). Functionally, production of granzyme and perforin enables MAIT cells to exert cytotoxic function via degranulation (Kurioka et al., 2015).

The role of MAIT cells in disease is not clear, although the number and capacity of MAIT cells has been shown to vary in a number of autoimmune conditions. They have been suggested to have immune-suppressive effects in multiple sclerosis and to migrate to inflamed tissue in inflammatory bowel disease (Chiba et al., 2018). The latter effect was however not evident when Ben Youssef et al. (2018) examined the intestinal mucosa samples of three infants with necrotizing enterocolitis. Mucosal MAIT cells were however found to have a memory phenotype, contrary to the naïve MAIT cells found in the thymus or peripheral blood (Ben Youssef et al., 2018). What role, if any, these cells have in the pathogenesis of disease remains to be elucidated.

Table 5 - Summary of differences between peripheral blood T cell populations in preterm infants compared to term infants, *indicates cord blood

	Relative abundance **(Preterm compared to term infants, population as % of T cells)**	**Relative Function** **(Preterm compared to term infants)**
T_h cells **(Gibbons et al., 2014)**	↑↑	↑ IL-5 ↓ IFN-γ
T_{reg}s **(Dirix et al., 2013)**	↑↑	↑ IL-10
T_c cells **(Berrington et al., 2005, Scheible et al., 2015)**	↔	Not known ↑IFN-γ* ↑TNF-α* ↓IL-2*
γδ T cells **(Gibbons et al., 2009, Li et al., 2013)**	↓↓	↑ IFN-γ ↑ IL-10
iNKT cells **(Ladd et al., 2010)**	↑*	Not known
MAIT cells **(Ben Youssef et al., 2018, Walker et al., 2014)**	↓	Not known

1.3.10 *Summary of T cells in preterm infants*

The T cell population in preterm infants differs in both composition and function from both adults and term infants.**Error! Reference source not found.** highlights the key differences b etween preterm infants and term infants.

Preterm infant T cells are characterised by an increased T_{reg} proportion, an increased ability to produce T_h2 cytokines and a decreased ability to produce T_h1 cytokines in peripheral blood.

However, other important differences compared to older humans have been highlighted such as an increased ability of γδ T cells to produce IFN-γ and IL-10, as well as T_h cells to produce IL-8.

T cells have the ability to influence both the innate and humoral immune systems. What role innate like cells such as iNKT and MAIT cells play in this has yet to be unravelled, however studies on iNKT cells in animals are compelling and imply these cells may have an important role preventing intestinal inflammation in infants.

The number of studies performed on the peripheral blood or tissues of preterm infants is small, and findings from fetal or umbilical cord samples may not be representative of preterm infants.

1.4 Interaction between the diet, gut microbiome and T cells

1.4.1 *Interaction between the diet and the immune system*

1.4.1.1 Colostrum

As described in Section 1.1.6, colostrum has a markedly higher concentration of bioactive components than other breastmilk. Ren et al. postulated that in preterm pigs at increased risk of NEC, following intrauterine intraamniotic LPS injection, a bovine colostrum diet would protect against NEC and systemic inflammation. Pigs were euthanised on day 5 of life. Pigs that had been fed colostrum did not have a significant reduction in NEC however there was a trend towards decreased NEC severity. Pigs fed colostrum did have decreased levels of distal ileal IL-8 and IL-1β, as well as increased distal ileal intestinal fatty acid binding protein and goblet cell density. This was combined with increased abundance of blood T_h cells (as % of lymphocytes). The authors have concluded that a colostrum diet led to maturation of the intestinal mucosa as well as the systemic immune system (Ren et al., 2019).

There has been some work exploring bovine colostrum in preterm infants albeit without exploration of either gut microbial or immune cell changes. Juhl et al. (2018) conducted a small RCT (n=40) aiming to identify if it is safe and feasible to supplement MOM with bovine colostrum during the first weeks of a preterm infant's life. They concluded that bovine colostrum appeared safe, although there was a rise in serum tyrosine in infants given bovine colostrum, this was felt to reflect a higher content of this amino acid in bovine colostrum and thought unlikely to be of clinical relevance (Juhl et al., 2018). This RCT has led to a further larger RCT which is due to complete recruitment in late 2020, which aims to identify if supplementation with bovine colostrum improves feed tolerance (Clinical Trials.gov identifier NCT03085277).

1.4.1.2 Human Milk Oligosaccharides (HMO'S)

As described in Section 1.1.6.3.1, HMO's are an important component of MOM. HMO's have been shown to promote the expansion of *Bifidobacterium* species in vivo whilst suppressing potentially pathogenic organisms including Group B streptococcus and Candida albicans in vitro (Gonia et al., 2015, Bode, 2012).

Maternal HMO secretion is dependent on the maternal Lewis Blood Group (Bode, 2015, Azad et al., 2018). HMO's have been suggested to modulate the immune system. This was demonstrated in a study that isolated HMO's from both colostrum and mature milk. Specific

HMO's were present in high concentrations only in colostrum. The HMO's were added to in-vitro organ cultures of fetal intestines (GA 7-22 weeks) before measuring gene expression of chemokines and chemokine receptors in the intestine. There was modulation of gene expression by colostrum HMOs, but not mature milk HMO's, affecting transcripts associated with signaling pathways relevant to T_h cell differentiation. Specifically, the observed transcriptional modulation following exposure to colostrum HMO's was deduced to promote a T_h1 response whilst suppressing the T_h17 response, T_h2 response and IL-8 expression. In order to investigate the influence of colostrum HMO's in inflamed tissue, intestinal epithelial cells were co-cultured with colostrum HMO's before the addition of pattern associated molecular protein (PAMP) stimulation. An upregulation in transcripts associated with recruitment of immune cells for bacterial clearance (neutrophils and macrophages) as well as tissue repair was found when colostrum HMO's were present, this was associated with a reduction in the concentration of IL-8, which was used as a marker of inflammation (He et al., 2014). Whilst this suggests that HMO's have local immunomodulatory effects especially in early life, they have also been detected in the plasma of term infants suggesting the potential for systemic effects (Ruhaak et al., 2014).

Together this suggests that HMO's could impact the T cell population in the intestine, including at the preterm stage.

1.4.1.3 Immune cells

As described in Section 1.1.6.3.4, human milk is known to contain both stem cells and leucocytes (Foteini et al., 2012, Trend et al., 2015), the numbers of which increase in response to not only maternal (mastitis) but also infant illness (viral infection) (Hassiotou et al., 2013). A number of studies have reported that immune cells can translocate from the intestine into the circulation or distal organs of animals (Ma et al., 2008, Hassiotou et al., 2015, Aydin et al., 2018). Stem cells derived from milk have been shown to translocate to the brain of mice, and once there they have been shown to be widespread and differentiate into glial and neuronal cells (Aydin et al., 2018).

In mouse models, maternal T cells are the predominant cell type that transfers across the intestinal epithelium despite their relative paucity in milk (Cabinian et al., 2016, Ma et al., 2008). T_c cells translocate into the intestinal Peyer's patches, reflected in their expression of a gut homing receptor (CCR9). The T_c cells present in mouse milk show an increased ability to produce cytokines upon stimulation compared to peripheral blood T_c cells (but similar to mature mouse T_c cells) (Cabinian et al., 2016). It is unclear whether milk-derived T cells migrate beyond the murine gut: Ma et al. found maternal derived T cells in the spleen and thymus of

pups, whilst Cabinian et al. only found maternal derived T cells localised to the intestinal mucosa, Peyer's patches and mesenteric lymph nodes (Cabinian et al., 2016, Ma et al., 2008).

1.4.1.4 Lactoferrin

Lactoferrin is the major whey protein of human milk. It is a glycoprotein that has been the subject of research for over 50 years due to its ability to inhibit bacteria through the sequestration of iron (Oram and Reiter, 1968). Lactoferrin affects bacteria by other mechanisms including cell membrane disruption, biofilm disruption and inhibiting bacterial adhesion to cells (Ward and Conneely, 2004). Lactoferrin has also been shown to bind to immune cells and influence their function (Legrand, 2016).

1.4.1.4.1 Interaction with bacteria and T cells

In vitro, lactoferrin inhibits the growth of bacteria that rely heavily on iron for growth such as *E. Coli* (Ward and Conneely, 2004). This is combined with an ability to increase the growth of *Bifidobacterium*. Interestingly bovine and human lactoferrin influence growth of different Bifidobacteria (Petschow et al., 1999).

Most lymphocytes can express a lactoferrin receptor including αβ and γδ T cells, however only when stimulated (Mincheva-Nilsson et al., 1997). Furthermore, T cells in the lamina propria of pigs have been shown to bind lactoferrin (Nielsen et al., 2010). In a mouse colon cancer model, T cells in the lamina propria were shown to proliferate in mice with and without cancer when administered enteral bovine lactoferrin. These T cells secreted IFN-γ and IL-18 suggesting that they were immunologically active (Wang et al., 2000).

1.4.1.4.2 Clinical trials

As a result of laboratory work demonstrating effective actions in inhibiting bacteria that are associated with LOS, it was hypothesised that administration of bovine lactoferrin to preterm infants would reduce LOS rates (Group, 2013). Promisingly a study in Italy did show a reduction in LOS (n= 472, RR 0.34 (CI 0.17-0.7)). The trial had 2 groups exposed to lactoferrin, with and without probiotics, and the reduction in LOS occurred in both intervention groups compared to the control group (Manzoni et al., 2009). The validity of these results was explored in a UK based RCT, Enteral Lactoferrin in Neonates (ELFIN) noting that a reduction in the rate of fungal sepsis seen in the Manzoni et al. trial explained some of the effect. Fungal sepsis is rarely seen in many UK NICUs since the introduction of antifungal prophylaxis (Cleminson et al., 2015).

The ELFIN study was a double blinded, multi-centre, randomised controlled trial that consisted of 2203 infants born less than 32 weeks GA. There was 90% power to detect a relative risk

reduction of 24-28% in the rate of LOS. The study found no difference in the rate of LOS between study groups (RR 1.05 (CI 0.87-1.26)) (Griffiths et al., 2019).

Both trials mentioned above used bovine lactoferrin, which though very similar to human lactoferrin has potentially important differences. Therefore, there is a question of whether supplementation with lactoferrin derived from humans could affect disease.

There has been one clinical trial exploring the use of recombinant human lactoferrin in low-birth weight (750-1500 grams) infants (n = 120). The trial was not powered to detect a reduction in infection but concluded that human lactoferrin was safe and there was a trend towards less infectious morbidity. The trial was limited by an unclear definition of hospital acquired infection (Sherman et al., 2016).

1.4.2 *Interaction between the microbiome and the immune system*

Much of our knowledge of the interaction between the microbiome and the immune system derives from the use of animal models. Germ-free animals can be bred offering a model that can then be manipulated to investigate this interaction. These mice can be inoculated with single bacterial strains or combinations (Hooper et al., 2012). The studies below refer to mouse models unless otherwise stated.

1.4.2.1 Increased innate inflammatory response in preterm infants

There are important deficiencies in the preterm intestine highlighted in Section 1.2.5. This includes an increased innate inflammatory response. Intestinal epithelial cells have TLR's on their cell surface and NOD-like receptors (NLR's) intracellularly. These receptors sample the intestinal contents for possible pathogens and recognise microbe associated molecular patterns (MAMP's) such as lipopolysaccharide (LPS).

As part of normal development, preterm infants are known to express high levels of TLR-4 in the intestinal mucosa compared to term infants(Hackam and Sodhi, 2018). TLR-4 is the receptor for LPS, which is the outer component of gram-negative bacteria. Increased expression of TLR-4 has been associated with the development of NEC (Afrazi et al., 2011, Hackam and Sodhi, 2018). Egan et al. showed that preterm infants and mice with NEC have an increased expression of the chemoattractant receptor CCL25 and IL-17RA, as a result of increased TLR-4 expression. These molecules lead to the induction of T_h17 cells which weaken the IEC tight junctions and infiltrate the intestinal mucosa, predisposing mice to develop NEC. This study demonstrated the plasticity of T_h cells to be induced to either display a T_h17 or T_{reg}

response. They also demonstrated that a dietary intervention with all-trans retinoic acid to suppress the T_h17 response could decrease the risk of disease in mice (Egan et al., 2016).

Pharmacological intervention with a TLR-4 inhibitor is possible and is being explored (Neal et al., 2013). However, human milk is already known to contain a number of molecules, including lactoferrin and HMO's, which have been shown to inhibit TLR-4 expression (He et al., 2016).

1.4.2.2 Immunoglobulin

The ability of the newborn intestine to secrete immunoglobulins develops over the first few weeks of life, being low at birth. Rognum et al. analysed duodenal samples from infants that had died postnatally and stained with antibodies to immunoglobulin and Human Leukocyte Antigens (HLA). This demonstrated low levels or absence of immunoglobulin (Ig) (IgA, IgG and IgM) and HLA at birth with a gradual increase in specimens from the first weeks of life. All the preterm infants in the study died in the first 3 days of life so it is unclear whether they had the same postnatal maturation as term infants (Rognum et al., 1992).

1.4.2.2.1 Specific Immunoglobulin production

MacPherson et al. demonstrated that commensal gut bacteria influence specific IgA production in mice. Germ-free mice were administered *Enterobacter cloacae* enterally, before mesenteric adenectomy. The presence of viable *Enterobacter cloacae*, using culture and 16S, was demonstrated in MLN's and dendritic cells. The number of bacteria correlated with the quantity of bacteria administered to the mice, however, when comparing intravenous injection of bacteria with gastric lavage, bacteria were only present in the mesenteric lymph nodes following gastric lavage. Furthermore, translocation was demonstrated to be through the Peyer's patches and not the lamina propria by identification of APC combined with bacteria in the Peyer's patches. Immunoglobulin A was found to be significantly increased in the intestinal mucosa and serum of mice inoculated with live but not dead bacteria. They concluded that dendritic cells originating in the Peyer's patches engulf bacteria following penetration of the intestinal barrier before transporting them to the mesenteric lymph nodes, where specific immunoglobulin is produced (Macpherson et al., 2000, Macpherson and Uhr, 2004). IgA is known to be important in the intestinal lumen by blocking bacterial access to epithelial receptors, keeping the bacteria within the mucus layer to facilitate removal from the body (Mantis et al., 2011).

Antigen specific IgA production has been shown to be heavily reliant on MyD88 signaling in T cells (Kubinak et al., 2015). However, in the mesenteric lymph nodes, IgA production is not entirely T cell dependent (Macpherson et al., 2000). B cells are not thought to populate the gut

mucosa until a few weeks of age, meaning host IgA production would be dependent on other mechanisms before then (Macpherson et al., 2000, Macpherson and Uhr, 2004). After this time, B cells are thought to secrete IgA into the intestinal lumen, as demonstrated by a linear relationship between age and fecal IgA in CMF fed infants (Gopalakrishna et al., 2019).

1.4.2.2.2 IgA impact on the microbiome and disease

The timing of secretion of IgA by host B cells is compelling as a study by Gopalkrishna et al., involving preterm infants and a mouse model, demonstrated that the IgA bound component of bacteria was protective of NEC. In particular they found an increased abundance of Enterobacteria in the IgA unbound bacteria. IgA was demonstrated to offer mice protection from NEC (Gopalakrishna et al., 2019).

Combining this data suggests that IgA is important for preventing the action of specific bacteria in the gut. In the first weeks of life, breast milk derived IgA may be particularly important before the intestinal mucosa has the capacity to produce IgA.

1.4.2.2.3 Should there be further clinical trials involving IgA?

Unfortunately, trials examining the use of oral immunoglobulin to prevent NEC in the late 1980's and early 1990's found no evidence of protection (Foster et al., 2016). With one exception, these trials used IgG, with the trial of IgA having a high risk of bias. A further study comparing oral IgA administration with oral gentamicin administration found a rate of NEC of 13% vs 1% respectively (n=200) (Fast and Rosegger, 1994).

Following on from the study by Gopalakrishna et al., the question of IgA supplementation remains, possibly with the use of specific IgA, or IgA originating from human milk rather than plasma.

1.4.2.3 Induction of Regulatory T cells by commensal bacteria

1.4.2.3.1 Clostridia species

Similar to humans, mice harbour high concentrations of T_{reg}s in the intestine compared with other organs. In mice, the introduction of gut bacteria induces an increase in the population of colonic but not small intestinal T_{reg}s which appears driven by a *Clostridium* species. The T_{reg}s are thought to have been peripherally induced (rather than thymus-derived) on the basis that they lack expression of the transcription factor Helios (K. et al., 2011).

1.4.2.3.2 Bacteroides fragilis

Induction of T_{reg}s does not seem to be isolated to *Clostridium* species. *Bacteroides fragilis* has also been shown to induce T_{reg}s (Round et al., 2011, Ramakrishna et al., 2019). Production of

polysaccharide A (PSA) by *Bacteroides fragilis* in the presence of TLR-2, has been shown to drive this change. Induction of T_{reg}s with treatment by PSA from *Bacteroides fragilis* has been implicated in the suppression of not only intestinal diseases (colorectal cancer, colitis) but encephalitis suggesting a systemic effect of this interaction (Sittipo et al., 2018, Ramakrishna et al., 2019, Round et al., 2011). These induced T_{reg}s have been shown to produce IL-10. Interestingly, a mutation causing loss of function in the IL-10 receptor has been found in human infants with early-onset enterocolitis (Glocker et al., 2009), highlighting the importance of IL-10 in intestinal health in early life.

1.4.2.4 Induction of iNKT cells by gut bacteria

Shen et al. demonstrated that mice deficient in iNKT cells had a different gut microbial composition compared with wild-type mice. Furthermore intestinal leucocyte infiltration, especially neutrophils, was increased in iNKT deficient mice, not dependent on the microbial composition (Shen et al., 2018).

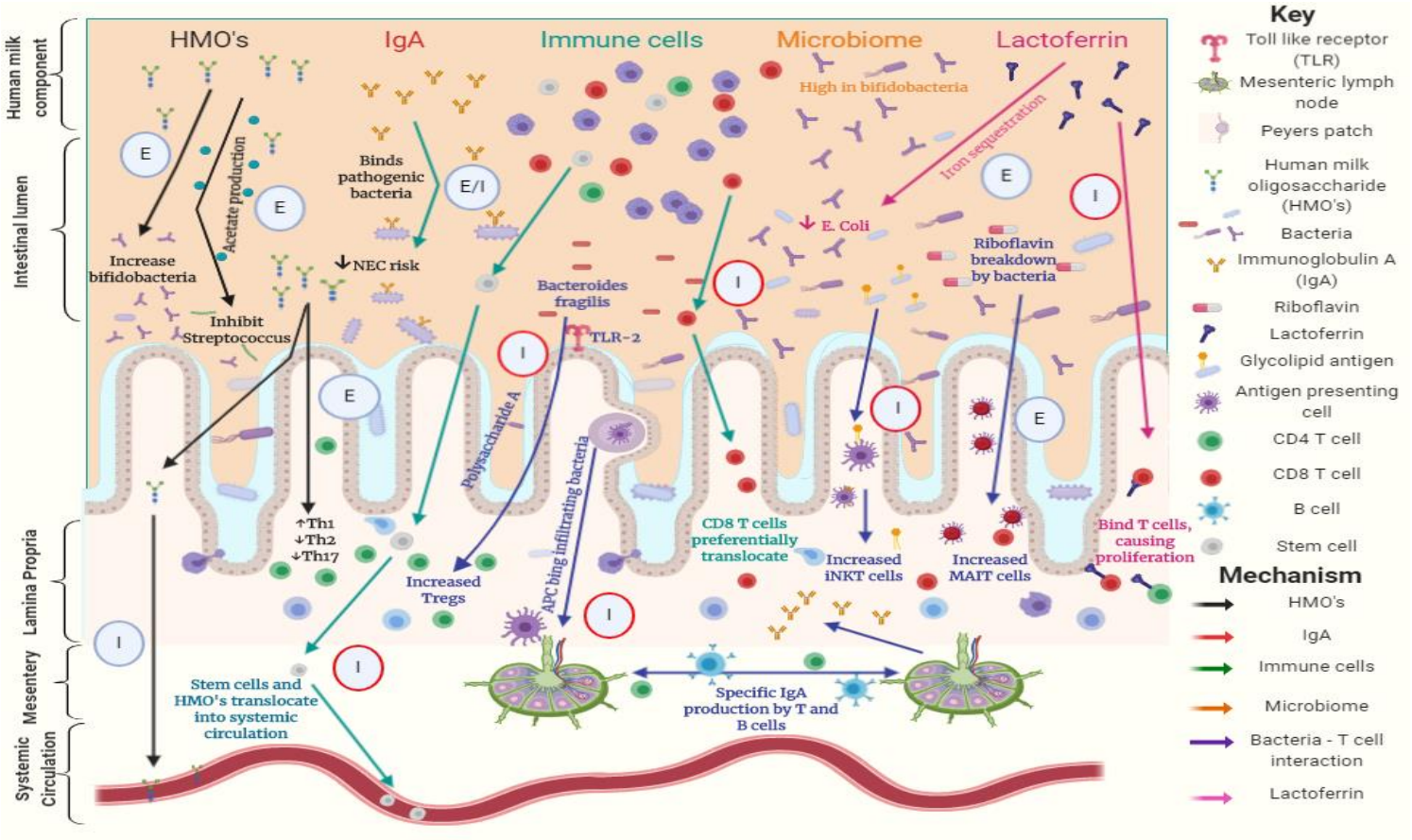


Figure 5 – Graphical summary of possible interactions between components of human milk, the gut microbiome and T cell populations in the preterm intestine. This includes work using human, animal and laboratory-based experiments. E indicates ex-vivo experiment, I indicates in-vivo experiment. Experiments with animal tissue are indicated by red circle whilst all grey circles indicate experiments using human tissue.

References

HMO's (Gonia et al., 2015, Bode, 2012, He et al., 2014, Ruhaak et al., 2014)

IgA (Gopalakrishna et al., 2019)

Immune cells (Aydin et al., 2018, Hassiotou et al., 2015, Ma et al., 2008, Cabinian et al., 2016)

Microbiome (Biagi et al., 2017)

Bacteria – T cell interaction (Round et al., 2011, Ramakrishna et al., 2019, Macpherson and Uhr, 2004, Olszak et al., 2012, Leeansyah et al., 2014)

Lactoferrin (Wang et al., 2000, Ward and Conneely, 2004, Nielsen et al., 2010)

1.4.3 *Summary of introduction*

The optimal diet of preterm infants is MOM due to a wide range of benefits including the reduced risk of serious disease, such as NEC and LOS. There are a number of mechanisms by which the diet, gut microbiome and the immune system have been demonstrated to interact in human, mouse and laboratory experiments, some of which are summarised in Figure 5. Currently, there is uncertainty as to the relative importance of different constituents of MOM in mediating the net protective effect.

When there is a shortfall of MOM, the impact of DHM on NEC and LOS compared to CMF appears inconsistent from the published literature to date, exemplified by the studies described in Table 1, and hence further study is merited. DHM differs from MOM in part due to the effects of pasteurisation and storage, especially on cellular constituents. If DHM were to have a beneficial effect over CMF it may be due to HMO's or lactoferrin, as the concentration of these molecules is relatively stable upon processing of DHM. There are likely to be other mechanisms, and it is also possible that there are detrimental effects of CMF e.g. foreign proteins inducing an allergic type response. However, it is plausible that DHM may be inferior to CMF in certain situations possibly due to variation in the processing of DHM as well as the natural variation in bioactive components.

However, given the interactions that have been described, it is reasonable to hypothesise that any difference in disease risk consequent upon infant dietary manipulations would be accompanied by alterations in the gut microbiome and/or T cells.

Chapter 2. Study Design

This thesis concerns infants recruited in Newcastle as part of the Interactions between the diet, gut microbes and body composition (INDIGO) study.

This thesis describes a pilot study aiming to identify interactions between the diet, microbiome and T cell populations in preterm infants. Figure 6 offers a graphical representation of the study hypotheses.

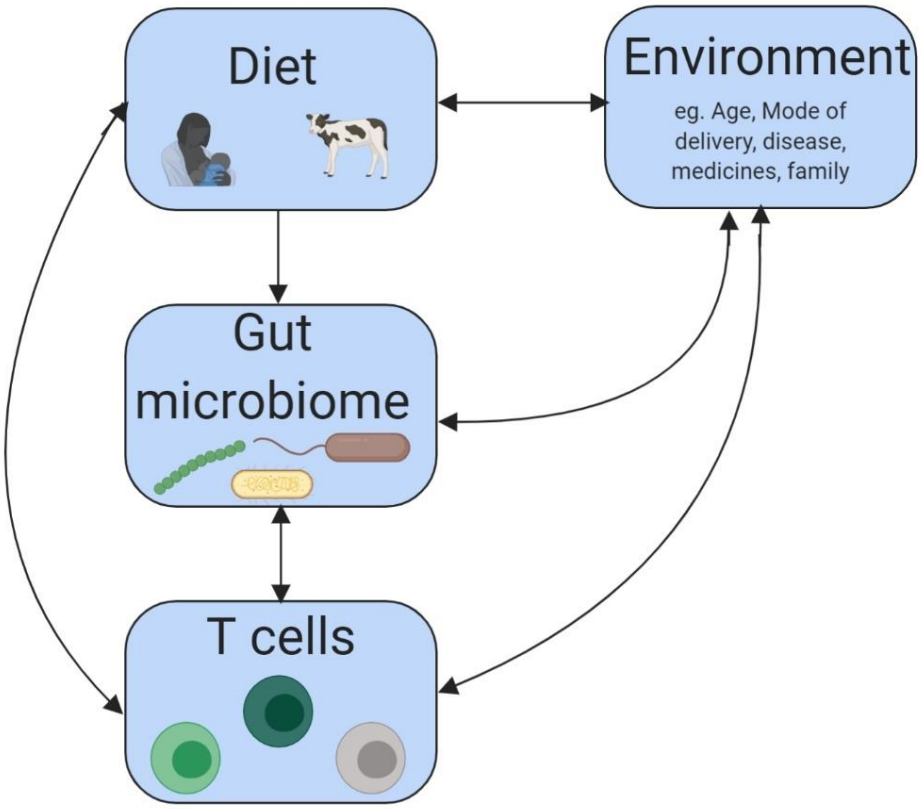


Figure 6 - Graphical representation of the hypotheses of the study

2.1 Aims

The aim of the INDIGO study was to test the impact of an exclusively human milk diet compared to a diet containing bovine products on gut microbiota, growth and metabolic outcomes.

There were two main aims of the pilot study described in this thesis:

- To identify alterations in the gut microbiome and T cell population, associated with the exposure of preterm infants to different diets.
- To describe the peripheral blood T cell immunophenotype in the preterm infant

2.2 Study Hypothesis

The hypotheses in the study are derived from evidence in the literature. They are presented as null hypotheses.

1. There is no association between the dietary intervention and the gut microbiome, either in terms of alpha diversity, beta diversity or composition at genus level, at 34 weeks corrected GA.
2. There is no association between the dietary intervention and peripheral blood T cell subsets, specifically the abundance of T_h1, T_h2 or T_h17 cells or abundance of iNKT cells, MAIT cells or T_{reg}s
3. There is no association between gut microbiome composition and peripheral blood T cell subsets, specifically the abundance of T_h1, T_h2 or T_h17 cells or abundance of iNKT cells, MAIT cells or T_{reg}s
4. There are no significant differences in iNKT, MAIT or T_{reg} cell abundance in preterm infants compared to adults

Chapter 3. Methods

3.1 Graphical Methods

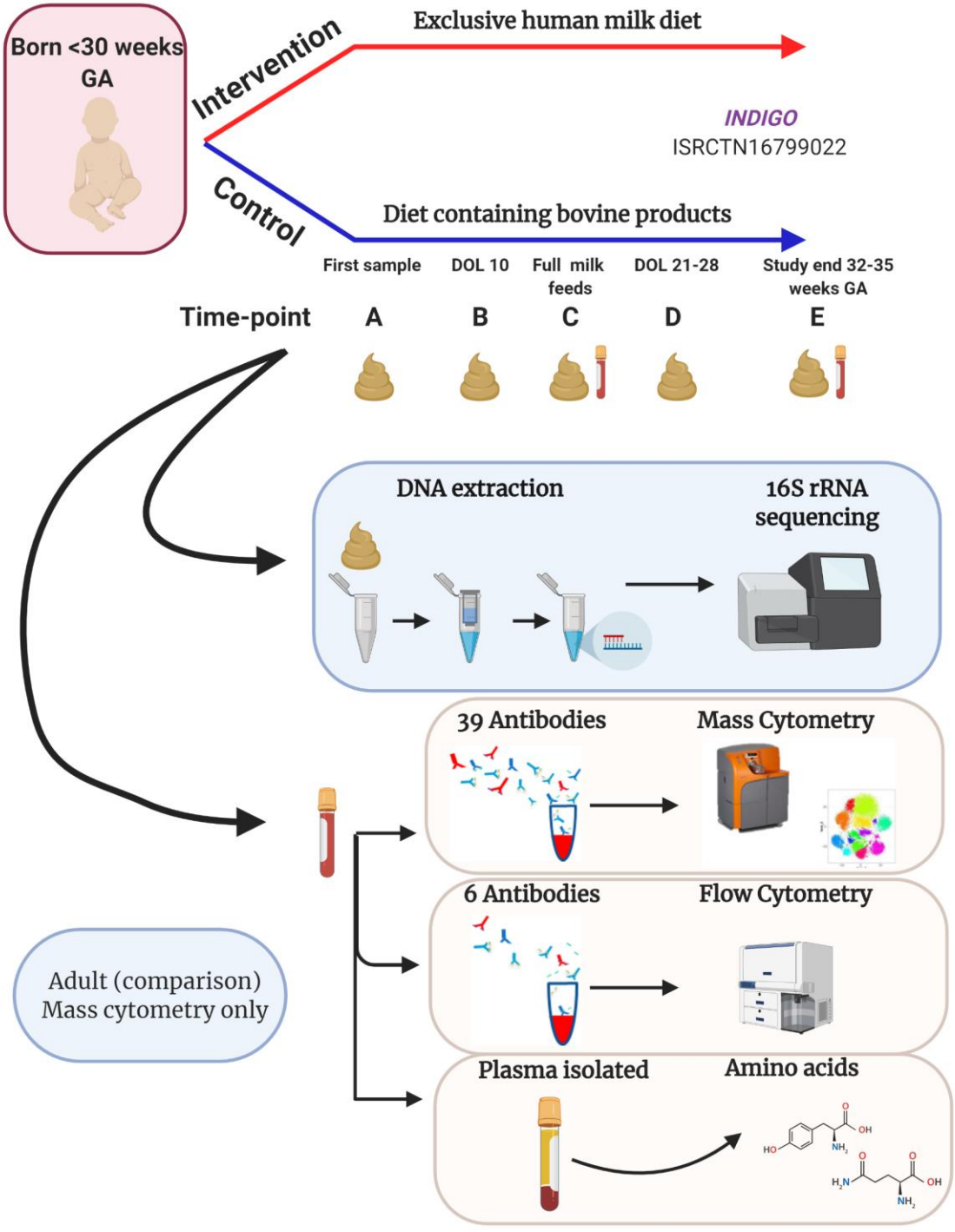


Figure 7

3.2 Resource Table

Table 6 – Resource Table *denotes conjugation using antibody labelling kit with metal listed in Section 3.5.6

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies **(Antibody (Clone) – Isotope of metal (unless conjugated on site) or fluorochrome)**		
CD45 (HI30) – 89Y	Fluidigm®	Catalog # 3089003B
CD19 (HIB19)*	Biolegend®	Catalog # 302247
CD196/CCR6 (GO34E3) – 141Pr	Fluidigm®	Catalog # 3141003A
CD152/CTLA-4 (14D3) *	Thermofisher	Catalog # 16-1529-82
TCR y/d (B1) *	Biolegend®	Catalog # 331202
CD3 (UCHT1) *	Biolegend®	Catalog # 300443
CD4 (RPA-T4) -145Nd	Fluidigm®	Catalog # 3145001B
CD8a (RPA-T8) – 146Nd	Fluidigm®	Catalog #3146001B
CD161 (HP-3G10) *	Biolegend®	Catalog # 339919
CD86 (IT2.2) *	Biolegend®	Catalog # 305402
CD25 (2A3) – 149Sm	Fluidigm®	Catalog # 3149010B
CD199/CCR9 (L053E8) *	Biolegend®	Catalog # 358902
CXCR3 (G025H7) *	Biolegend®	Catalog # 353702
CD1d (51.1) *	Biolegend®	Catalog # 350302
CD56 (REA196) *	Miltenyi Biotech®	Catalog # 130-108-016
CD49b (P1E6-C5) *	Biolegend®	Catalog # 359301
CD45RA (HI100) -155Gd	Fluidigm®	Catalog # 3155011B
CD335 (9E2) *	Biolegend®	Catalog # 331902
Anti-FITC (FIT-22) *	Biolegend®	Catalog # 408305
Va24-FITC (6B11)	Biolegend®	Catalog # 342906
CD279(PD-1) (EH12.2H7) *	Biolegend®	Catalog # 329941
CD197/CCR7 (G043H7) – 159Tb	Fluidigm®	Catalog # 3159003A
Vα7.2-APC	Biolegend®	Catalog # 351708
Anti-APC (APC003) *	Biolegend®	Catalog #408005
Anti-Human/Mouse Tbet (4B10) – 161Dy	Fluidigm®	Catalog # 3161014B
Anti-Human FoxP3 (259D/C7) – 162Dy	Fluidigm®	Catalog # 3162024A
Anti-Human CD294/CRTH2 (BM16) – 163Dy	Fluidigm®	Catalog # 3163003B
CD69 (FN50) *	Biolegend®	Catalog # 310939

REAGENT or RESOURCE	SOURCE	IDENTIFIER
TIGIT (A151536) *	Biolegend®	Catalog # 372702
Granzyme B/Perforin (REA226) *	Miltenyi Biotech®	Catalog # 130-108-055
Anti-Human/Mouse Gata3 (TWAJ) – 167Er	Fluidigm®	Catalog # 3167007A
ROR-y (600214) *	R&D Systems®	Catalog # MAB6109
CXCR5(RF8B2) *	BD Biosciences®	Catalog # 552032
5-OP-RU (tetramer) – PE	NIH tetramer core facility	Not applicable
Anti-PE (PE001) *	Biolegend®	Catalog # 408105
CD28 (CD28.2) *	Biolegend®	Catalog # 302937
CD14 (M5E2) *	Biolegend®	Catalog # 301802
CD117/c-kit (104D2) *	Biolegend®	Catalog # 313223
HLA-DR (L243) -174Yb	Fluidigm®	Catalog # 3174001B
CCR4 (L291H4) – 174Lu	Fluidigm®	Catalog # 3175035A
CD127 (A019D5) – 175Yb	Fluidigm®	Catalog # 3176004B
CD16 (3G8) – 209Bi	Fluidigm®	Catalog # 3209002B
CD66b (G10F5) – FITC	Biolegend®	Catalog # 305104
Metals		
Maxpar® X8 antibody Labelling Kits (multiple – see Section 3.5.8 for details)	Fluidigm®	Catalog # (multiple) 201141A - 201176B
Biological Samples		
Adult peripheral blood samples	Newcastle University	Permissions: AWERB Project ID: 633
Preterm peripheral blood samples	NICU, RVI	Permissions: ISRCTN16799022
Preterm stool samples	NICU, RVI	Permissions: ISRCTN16799022
Reagents		
Red Cell lysis Buffer	Biolegend®	Catalog #420302
Lymphoprep™	Stemcell technologies™	Catalog # 07851
Maxpar® water	Fluidigm®	Catalog #201069
Maxpar® cell staining buffer	Fluidigm®	Catalog # 201068
Maxpar® Fix and Perm buffer	Fluidigm®	Catalog # 201067

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Cell-ID™ intercalator 125uM	Fluidigm®	Catalog # 201192A
Triton X-100	Sigma-Aldrich®	Product no. X100-100ml
Formaldehyde 16%	Thermofisher®	Catalog # 28906
Hanks Balanced Salt Solution (HBSS)	Sigma-Aldrich®	Catalog # H9394-500ml
Dimethyl sulfoxide (DMSO)	Sigma Aldrich®	Catalog # D9170-5VL
Dulbecco's Phosphate Buffer Solution (PBS)	Sigma-Aldrich®	Product no. D8537
Fetal bovine serum (FBS) Gibco™	Thermofisher®	Catalog # 10270-106
Cell-ID™ Cisplatin – 194Pt	Fluidigm®	Catalog # 201194
Laboratory Equipment		
Mr Frosty™ Freezing Container	ThermoFisher®	Catalog # 5100-0001
NanoDrop™ One	Thermofisher®	Catalog # ND-ONE-W
Helios™, mass cytometer	Fluidigm®	
FACSCanto™, flow cytometer	BD Biosciences®	
MiSeq™ Personal Sequencer	Illumina®	
Amino Acid Analyser	Biochrom®	
Critical Commercial Assays		
Powerlyzer® powersoil® DNA extraction kit	Qiagen®	Catalog # 12855-100
MAXPAR® Antibody Labelling Kit	Fluidigm®	Catalog # 201142
Multitest™ 6-colour TBNK	BD®	Catalog # 644611
Key Documents		
INDIGO study protocol version 1.3	Available on request	Details at www.isrctn.com ISRCTN: 16799022
INDIGO study patient Information Leaflet (PIL)	Appendix A	
INDIGO study consent form	Appendix B	
INDIGO study data collection forms (Study enrolment and daily collection)	Appendix C	
Software and Algorithms		
CyTOF software v6.7	Fluidigm®	https://www.fluidigm.com/software
FCS Express 6	DeNovo software®	https://www.denovosoftware.com

REAGENT or RESOURCE	SOURCE	IDENTIFIER
R 3.5.1	R Core Team 2018	https://www.r-project.org
GraphPad Prism 8©	GraphPad Software	https://www.graphpad.com/
Sealed Envelope	Sealed Envelope Limited	https://www.sealedenvelope.com/
REDCap	REDCap Consortium	https://www.project-redcap.org/
Measurement tool		
Leicester incubator measure	Harlow Healthcare®	Catalog Category: Height/Length Measuring Equipment

3.3 Patient Cohort – INDIGO study

3.3.1 *Study Design*

The Interactions between the diet, gut microbes, metabolism and body composition (INDIGO) study was a multicentre, stratified (using gestational age <25 weeks or 25-30 weeks, hospital site, and multiple pregnancy) with balanced randomisation (1:1), non-blinded, parallel group study conducted in the United Kingdom (initially 2 sites). The INDIGO study will now be referred to as "the study". The study was designed to identify the impact of human milk-based products compared to cow's milk-based products in the feeding regime of extremely preterm infants, when there is a shortfall of MOM.

3.3.2 *Ethics Approval*

The study was approved by the National Health Service (NHS) Health Research Authority (HRA) North East – Tyne and Wear (T&W) South Research Ethics Committee (REC) on the 29th June 2017. The Integrated Research Applications System (IRAS) Identification for the study is 215037 and the International Standard Randomised Controlled Trial Number (ISRCTN) is 16799022.

3.3.3 *Changes to study design*

Whilst there were no major changes to protocol design, challenges in recruitment were experienced in the Chelsea and Westminster (C&W) Hospitals study site in early 2018 possibly due to the requirement of infants needing an MRI scan after 34 weeks GA. In addition, there

were a number of infants transferred to other hospitals before study completion at the Royal Victoria Infirmary (RVI) study site. As a result, three amendments to the eligibility criteria were submitted to the trial steering committee. Approval for these amendments (described below) was sought and obtained from the T&W South REC. This approval was obtained on the 13th June 2018.

The amendments were as follows:

1. Maximum gestational age of recruitment increased from 29 to 30 weeks
2. The time to obtain consent and perform randomisation increased from 48 to 72 hours
3. Permission to access ongoing care information using existing database systems if the infant was transferred to a different hospital so that outcome information of relevance to the study could be obtained

Subsequently, due to ongoing recruitment issues within the funded time available in C&W Hospitals, permission was sought to open two further recruitment sites (James Cook University Hospital, Teesside and William Harvey Hospital, East Kent). This was agreed by the trial steering committee and T&W South REC approval for a non-substantial amendment was obtained on the 3rd October 2018.

3.3.4 *Participants*

Inclusion criteria for the study were any infants born before 29 weeks (September 2017 – June 2018) or 30 weeks (June 2018- March 2019) gestational age, who did not have any severe immediate life-threatening condition or congenital abnormalities. Infants were excluded if they had received any other milk apart from MOM or their mother had no intention to provide any MOM. To be eligible, infants had to be randomised before 72 hours of age (48 hours of age (September 2017 – June 2018)).

Only those infants recruited in the Neonatal Intensive Care Unit (NICU) in the Royal Victoria Infirmary (RVI), Newcastle Upon Tyne Hospitals are described in this work. The NICU at the Royal Victoria Infirmary provides local neonatal care to a maternity unit with 6000-8000 deliveries per year, as well as providing neonatal surgical care for infants from Cumbria, Northumberland, County Durham and Darlington and Teesside. The RVI NICU cared for 153 infants born less than 32 weeks gestational age in 2017 (Fenton, 2017). The standard care in the RVI NICU at the time of the study was for cow's milk formula(CMF) should there be a shortfall of MOM, and cow's milk-based fortification(BMF) if fortification was clinically indicated.

3.3.5 *Consent process*

The consent process followed those outlined in the Good Clinical Practice (Authority, 2017). A study investigator (NDE, JEB, JG, AK, TDRS) would approach the parents of an infant typically within the first 24 hours of life. The study would be verbally explained in language that was easy to understand and a Patient information leaflet (PIL) given (Appendix A). Parents were made aware that the routine care of their infant would not be affected regardless of their participation and that they could withdraw from the study at any time. Parents were typically given 12-24 hours to consider the study before being approached by the study investigator for their decision regarding participation. Following verbal consent, participation was confirmed upon completion of a consent form which re-iterated the key aspects of the study (Appendix B).

3.3.5.1 Language Difficulties

The study protocol does not specify action with regards to parents who do not speak English. An infant was therefore deemed eligible if their parent could read English and gave informed consent following discussion with the aid of an interpreter. This ensured that all parents had both verbal and written information regarding the study before participation.

3.3.6 *Randomisation*

Infants were randomised to either an exclusive human milk diet (intervention) or standard care (control). Both diets are detailed in Section 3.3.7. Randomisation was performed using an online software package with minimisation designed for this purpose (www.sealedenvelope.com). Randomisation was performed on a 1:1 basis, stratified by hospital site, gestational age (<25 weeks, or 25-29+6 weeks) and multiple pregnancy. Multiple deliveries were randomised individually. Once randomisation was performed, the parents, bedside nurse and attending clinician were informed of the randomisation allocation (i.e. no blinding).

3.3.7 *Milk intervention*

Following informed consent and randomisation, the attending medical team were advised to commence enteral feeds before 72 hours unless there was a medical contraindication. Donated human milk was supplied for this study by Prolacta Biosciences© (California, USA) and consisted of unfortified human milk (HM), fortified human milk (RTF 26) and human milk-based human milk fortifier (P+6). These products are produced exclusively from human milk and were couriered frozen from California to Newcastle in batches, to ensure there was no

shortfall in supply. The standard practice in the RVI NICU is for SMA Gold prem 1 (GP1) (Nestle©, UK) to supplement MOM should there be a shortfall, and for Cow and Gate© nutriprem breast milk fortifier (Nutricia©, UK) to be used to fortify MOM once preterm infants are fully enterally fed (defined at 150ml/kg/day). This standard practice was the diet used in the control arm of the study.

The dietary intervention in the study had two components. The two components relate to the point at which an infant achieved full enteral feeds (defined as the day when 150ml/kg milk was tolerated).

1. At any time in the study period (until 34 weeks CGA), infants were supplemented with RTF 26 (Intervention arm) or GP1 (Control) should there be a shortfall of MOM.
2. From the day of achieving full enteral feeds (defined as 150ml/kg/day), fortification of MOM was commenced with either Prolacta P+6 (Intervention arm) or Cow and Gate nutriprem breast milk fortifier (Control).

Table 7 and Table 8 compare the macronutrient content of the formulas and fortifiers used in the Control and Intervention arms of the study.

Table 7 - Comparison of macronutrient concentrations of formulas used in the study

	Control **GP1/100ml**	**Intervention** **RTF 26/100ml**
Energy (Kcal)	80	90
Protein (g)	2.9	2.6
Carbohydrates (g)	8.1	8.1
Fat (g)	4	5.2

Table 8 - Comparison of macronutrient and folate concentrations of fortifiers used in the study. Concentrations of macronutrients were derived from (Stoltz Sjostrom et al., 2014) per 100 ml MOM as follows; 70kcal energy, 1.8g protein, 6.8g carbohydrates, 4g fat. *not including MOM

	Control **Cow and Gate Nutriprem Breast Milk Fortifier (100ml MOM and fortifier)**	**Intervention** **Prolacta P + 6** **(70ml MOM and 30ml P+6)**
Energy (kcal)	85	88
Protein (g)	2.9	2.8
Carbohydrates (g)	9.5	7.6
Fat (g)	4	5.2
Folate/Folic Acid * (micrograms)	30*	11.7*

A discrepancy in the concentration of folate between the two fortifiers was noticed following the commencement of the study. Measurement of serum folate to ensure there was no deficiency was therefore arranged, combining with other blood tests to ensure no extra venepuncture for the study infants.

3.3.8 *Data Collection*

Information regarding infants was collected using data collection forms (Appendix C) as outlined in the study protocol. Admission details were collected at the time of recruitment. Anthropometry was collected weekly. Nutritional information was collected daily.

3.3.8.1 Admission details

Admission details included birth mode, exposure to antenatal factors (prolonged rupture of membranes, reduced end-diastolic flow), condition at birth (heart rate at 5 minutes of age) and maternal details (NHS Ethnicity and antenatal steroids).

3.3.8.2 Anthropometry

Anthropometry included measurements of head circumference, length and weight. Weight measurements were made as part of routine infant care and the data recorded from the bedside nursing chart. Head circumference was measured using a tape measure with the occipito-frontal circumference (OFC) recorded. Length measurements were made using the Leicester Incubator measure by a researcher and an assistant, typically the infant's bedside nurse. This enabled measurements of infants with minimal disturbance as has been reported (King and Dogra, 2011). Length was measured without the use of the length board in order to further minimise disruption of the infant, although this may impact the accuracy of results (Wood et al., 2013). For both OFC and length measurements, the largest of 3 recordings was recorded provided there was no more than 0.2cm between recordings, if there was greater than 0.2cm between measurements the recordings were repeated.

3.3.8.3 Nutritional Data

Nutritional data included any exposure that may impact the gut microbiome such as probiotics, antibiotics, vitamins and iron on a per sample (blood or stool) basis. Exposure at the time of sampling (Yes or No) and number of days the patient was exposed before sampling were recorded. With regards to milk, the following nutritional information was recorded:

- Type (MOM, Prolacta Ready to Feed 26 (RTF), CMF, TPN or other) and volume (nearest milliliter)
- Fortifier exposure and type (Human or bovine)
- If the patient had been nil by mouth for greater than 4 hours on that day

Following identification of the stool and blood samples for analysis, samples were categorised based on the percentage of enteral nutrition that was MOM in the 72 hours before a sample (including day of sample) using the equation:

Total volume of MOM (ml) / Total volume of enteral nutrition (ml) x 100.

The categories used were:

- 70-100% MOM

- 30-69% MOM
- 0-29% MOM

An infant receiving fortified MOM in the intervention arm could receive a maximum of 70% MOM (as the fortifier comprises 30ml for every 70ml MOM) therefore using these cut-offs offers the opportunity of a relevant comparison between the two study arms.

3.3.8.4 Clinical Outcomes

Usual practice at the RVI NICU is for preterm infants to be transferred to the paediatric service after they are post-term. Following transfer to paediatric services the availability of detailed information decreases. Therefore, details regarding clinical outcomes were collected at hospital discharge or 4 weeks post-term (whichever occurred first). Clinical outcomes included information about morbidity, mortality and hospital stay details.

Morbidities were defined as:

- **LOS**: microbiologically confirmed infection of blood or cerebro-spinal fluid with any organism, including coagulase-negative staphylococcus (CONS) only if the intention was to treat as LOS. LOS was defined after 72 hours of life and before 44 weeks postmenstrual age or discharge, whichever occurred first. This definition was used in a recent large multi-centre neonatal RCT (Costeloe et al., 2016)
- **NEC**: at least one clinical and one radiological sign, as well as intention to treat for greater than or equal to 5 days with metronidazole. Clinical signs; abdominal distension, blood in stool, billous gastric aspirate or vomiting. Radiological (AXR) signs; pneumatosis, hepato-biliary gas, pneumoperitoneum. Alternatively, a diagnosis could be made histologically; either at post-mortem or following surgery. This is based on the definition used by the NNAP (RCPCH, 2017). Each case of NEC was independently reviewed by an additional consultant neonatologist to ensure robust agreement.
- **Intra-ventricular haemorrhage (IVH**): Any stage recorded by radiologist on cranial ultrasound, stage defined as per ICD-10 (Organisation, 2004).
- **Retinopathy of Prematurity (ROP):** Any stage recorded by consultant ophthalmologist. Screening undertaken as per local NICU guidance (any infant less than 32 weeks gestational age or <1500g at birth)
- **Chronic Lung Disease**: Defined as mild (oxygen or respiratory support at 28 days), moderate (oxygen requirement at 36 weeks corrected gestational age) or severe (respiratory support at 36 weeks corrected gestational age) (Jobe and Bancalari, 2001)

Should a patient die before discharge or 4 weeks post-term, the primary causes of death as determined by multidisciplinary case review with clinicians independent of the research team including information from post-mortem (if performed), were recorded. Furthermore, factors that may have contributed to death, namely infection, NEC, brain injury, decision to withdraw intensive care, congenital anomaly and respiratory failure were recorded in a binary fashion.

3.3.8.5 Patient Transfer

Should a patient be transferred to another hospital before completion of the clinical study (defined as 34 weeks CGA), study outcome data was collected from the BadgerNet neonatal electronic patient record as per the study amendment detailed in Section 3.3.3. Additionally, should the transferred patient be in the study control group, a request was made to the local neonatal department to obtain detailed feeding and anthropometry information. Ethical approval and clinical governance issues did not permit the use of the intervention products (HMF) outside the recruiting NICUs, so growth and nutritional intakes from those infants were considered off-protocol.

3.3.8.6 Data storage

Data storage was conducted as outlined in Good Clinical Practice. The study site file was stored in a secure room. Original copies of consent forms were kept in the site file, any complete data collection forms were securely stored. A secure web-based application named research electronic data capture (REDCap®) was used to collate and store all clinical data in a central location. REDCap® is designed for data management and storage and allows for the creation of a bespoke data collection form. REDCap® was hosted at the RVI, Newcastle (Harris et al., 2019).

3.3.9 *Study withdrawal*

Should an infant withdraw from the study, data was collected up until the point of withdrawal from the study. If any stool or blood samples were collected before withdrawal from the study these were analysed. If any samples were collected following study withdrawal they were discarded.

3.3.10 *Study completion*

An infant was deemed to have completed the study at 34 weeks CGA. However, as stool samples depend on stooling habits and collection by nursing staff, there would not necessarily be a sample available after 33 weeks CGA. Preferentially, a sample obtained between 33-34

weeks CGA was used for the analysis, however infants who completed the intervention for at least 4 weeks and had stool samples available between 32-35 weeks corrected gestational age were deemed to have complete the study providing they had not deviated from the study intervention before 34 weeks CGA.

3.3.11 *Sample Size*

The study was powered on detecting a difference in the gut microbiome between the study groups, specifically a 0.5 standard deviation (SD) difference in alpha-diversity between study groups. 40 infants in each study group would give a 95% power at a 5% significant level to detect this.

For this pilot study, 16 infants in each study group would give an 80% power at a 5% significance level to detect a 0.5 SD difference in alpha-diversity between the study groups.

3.4 Samples

3.4.1 *Stool collection and storage*

Stool samples were collected from participants during routine care times by either their parent or the bedside nurse. A maximum of one sample was collected per day. Stool samples were collected in sterile glass containers before being transferred with patient details to a -20 Celsius(C) refrigerator on the NICU. The samples were then anonymised by giving them a unique identification number. Samples were transferred weekly to a -80 C freezer for storage until DNA extraction.

3.4.1.1 Selection of stool samples for analysis

In order to understand about the effect of the diet on the gut microbiome, stools were identified from five time-points.

The five time-points selected for analysis were:

A. Earliest sample (before DOL 10)
B. DOL 10 (+/- 2 days)
C. Two days following full enteral feeding (150ml/kg/day)
D. DOL 21-28
E. Last sample before study end (range 32-35 weeks CGA)

Suitable stool samples were identified using their unique identification number in an online database. If a sample was identified that overlapped with a later time-point they were defined by the latter time-point.

3.4.2 *Blood collection*

3.4.2.1 Infants

Blood samples were collected from study infants as per the study protocol and timed to coincide with clinical blood tests.

Blood samples were collected at time-points C and E to coincide with stool sampling.

Clinical Research guidelines recommend a maximum draw of 1% of blood volume at any single time-point or 3% within a 4-week period ((CHMP) and (PDCO), 2008). Preterm infants are estimated to have a blood volume of 80-90ml/kg. When the infant weighed less than 1100 grams and therefore would have a circulating blood volume of less than 100ml, the clinical full blood count was salvaged following completion of the clinical test. Analysis of salvaged blood samples was permitted as per the study protocol. Blood samples were collected as per Figure 8 based on the infants' weight.

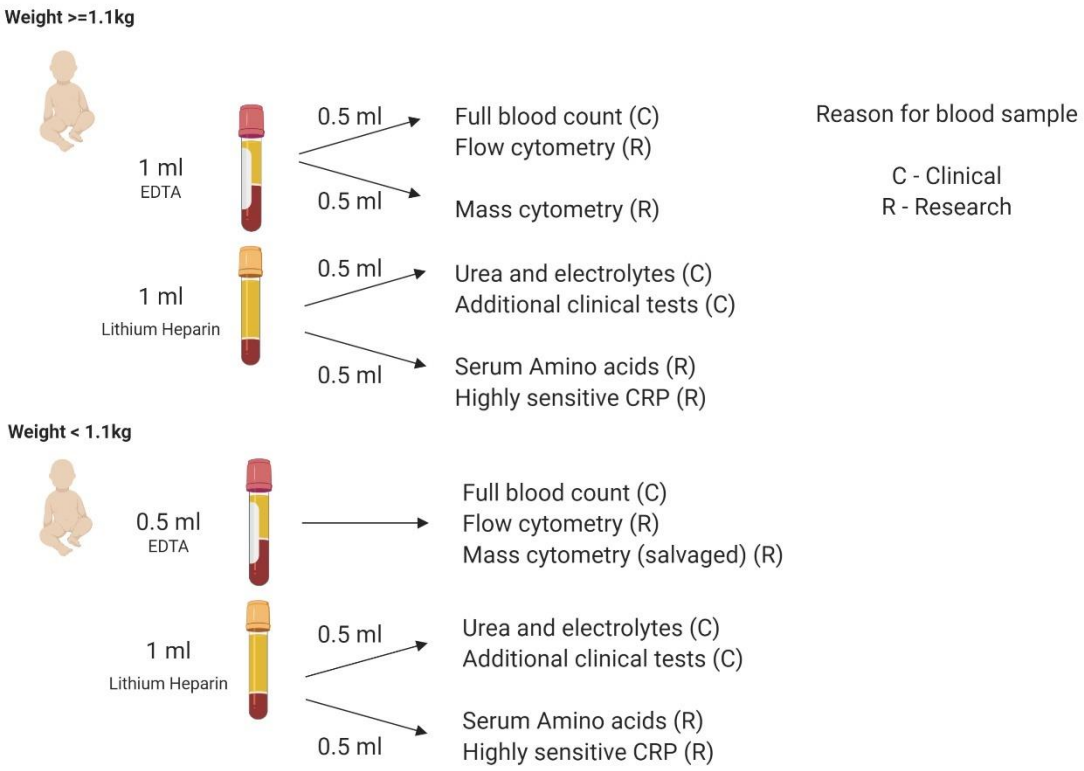


Figure 8 - Blood sampling during the INDIGO study based on infants' weight

Blood samples for amino acids analysis were centrifuged immediately, with the plasma isolated and then frozen at -20 C prior to analysis. Amino acid analysis occurred within 1 month of sample collection.

Blood samples to be analysed using mass cytometry were anonymised with the study identification number only. They were stored at room temperature prior to either red blood cell lysis or density gradient centrifugation, followed by staining with cell surface antibodies and freezing at -80 C within 24 hours of collection.

Due to the small blood volume of samples, red blood cell (RBC) lysis was initially undertaken. However, this resulted in a large number of granulocytes being analysed using mass cytometry. Following optimization, it was found possible to perform a density gradient centrifugation on small volume blood samples. This was the preferable method for cell preparation. There is therefore a combination of samples where cell reduction was performed using RBC lysis and density gradient centrifugation.

3.4.2.2 Adults

Blood samples from adults were to establish staining protocols and as technical controls for infant samples. These were collected as part of the Understanding Mechanisms of Immune

Mediated Diseases study (Animal Welfare Ethical Review Body Project ID: 0633) in Newcastle University. Written consent was obtained from participants and a trained phlebotomist collected 9ml of blood into an Ethylene-diamine-tetraacetic acid (EDTA) container using a butterfly needle. All samples were stored anonymously at room temperature prior to staining. Cell reduction methods used for adult samples mimicked that used for the infant samples that were stained simultaneously.

3.4.3 *Sample metadata*

Data were prospectively collected. After selection of the samples for analysis, this was collated against each sample time-point.

This information covered four areas:

1. **Demographics** Age, Sex, study group, admission details (see Section 3.3.8.1)
2. **Medications** before (number of days medication given before day of sample) and during (medication given on day of sample) sample (Antibiotics, probiotics, iron, vitamins)
3. **Nutrition** MOM category (see Section 3.3.8.3), fortifier exposure
4. **Disease** NEC, LOS (temporal relationship to sample i.e. days before/after sample)

3.5 Laboratory Techniques

3.5.1 *Bacterial DNA extraction from stool*

Bacterial DNA was extracted from 200ug of stool. A MoBio™ Powerlyzer® Powersoil® DNA extraction kit was used as per the manufacturer's instructions with the following 2 alterations. Bead beating combined with chemical lysis was undertaken for 20 rather than 10 minutes. Incubation with solutions C2 and C3, used to improve DNA quality by precipitating other organic material, was for 10 minutes rather than 5 minutes. A final volume of 100µl DNA was produced as per the manufacturer's instructions. This was stored at -80 C until 16S rRNA gene sequencing.

Bacterial DNA was extracted from 23 infant samples in each batch. A blank control sample was processed with each batch in order to detect background contamination.

3.5.2 *Sequencing*

Bacterial profiling was performed by the NI-OMICS DNA sequencing research facility at Northumbria University using the Illumina® MiSeq™ Personal Sequencer. This utilised the 16S rRNA gene targeting variable region 4 based on the Schloss wet-lab MiSeq standard operating

protocol (Kozich and Schloss, 2014). Fastq data files were created using the MiSeq™ platform (Illumina®). The primers used for amplification were GGACTACHVGGGTWTCTAAT and GTGCCAGCMGCCGCGGTAA3. Read pairs were demultiplexed based on the unique molecular barcodes, and reads were merged using USEARCH v7.0.10015, allowing zero mismatches and a minimum overlap of 50 bases. Merged reads were trimmed at first base with Q5. In addition, a quality filter was applied to the resulting merged reads, and reads containing above 0.05 expected errors were discarded.

Subsequently the fastq files were processed by the Alkek Center for Metagenomics and Microbiome Research (Baylor University, USA) to produce a biom and tre file containing OTU's and taxonomic information respectively for each sample as follows.

16S rRNA gene sequences were assigned to OTUs or phylotypes at a similarity cut-off value of 97% using the UPARSE algorithm. OTUs were then mapped to an optimised version of the SILVA Database6,7 containing only the 16S v4 region to determine taxonomies. Abundances were recovered by mapping the demultiplexed reads to the UPARSE OTUs. A custom script was constructed containing an OTU table from the output files (tre and biom files) generated in the previous two steps, which is then used to calculate alpha-diversity, beta-diversity, and provide taxonomic summaries that are leveraged for all subsequent analyses.

3.5.3 *Serum Folate analysis*

Serum folate was measured by the RVI Biochemistry laboratory. The test was performed on blood that had been taken for clinical or research purposes. Folate was analysed using the Elecsys Folate III kit as per the manufacturer's instructions and measured using a Cobas® e 411 analyser (Roche®)

3.5.4 *Plasma Amino acid analysis*

Plasma amino acids were analysed by the RVI Biochemistry department as follows.

Blood samples in tubes containing lithium-heparin were centrifuged within one hour of being received in the laboratory. The plasma was separated and then frozen at -20 C until analysis.

Following de-frosting, the plasma was de-proteinised with 10% sulphosalicylic acid containing 500 umol/L amino-L-cysteine as the internal standard. The sample was then centrifuged before the supernatant was injected onto an amino acids analyser (Biochrom™, UK).

The amino acids were separated by ion-exchange chromatography with lithium buffers of increasing pH and ionic strength. Post-column derivatisation with ninhydrin at 135°C formed

amino acid-ninhydrin conjugates with A_{max} 570 nm (440 nm for proline and hydroxyproline), which are detected on the photometric detector.

Data were recorded using a chromatogram with each amino acid separated and the peak area proportional to the amount of amino acid-ninhydrin conjugate formed. The area of each peak was compared with the internal standard. This was further compared with results obtained from a calibration standard containing 500 µmol/L of each amino acid to calculate the concentration in the sample.

3.5.5 *Flow Cytometry*

Blood sample analysis using conventional flow cytometry was performed in the Flow Cytometry Blood Sciences laboratory, RVI, Newcastle upon Tyne Hospitals.

Blood samples were collected in EDTA tubes. The blood sample was gently mixed before 50µl of blood was added to a Trucount™ tube (BD®, USA) containing 10µl of BD Multitest™ 6-colour TBNK (BD®, USA) and 25µl of HLA-DR. Samples were incubated for 15 minutes. 450µl of FACSlyse (BD®, USA) (10:1 dilution with distilled water) was added before further incubation for 15 minutes.

Samples were then analysed using a FACSCanto (BD®, USA) flow cytometer, acquiring 10,000 events using a lymphocyte stopping gate based on forward and side scatter. The gating strategy used to identify T, B and NK cells is defined by the manufacturer.

3.5.6 *Red Blood Cell (RBC) lysis*

For blood samples that underwent red blood cell lysis the protocol was as follows:

1. RBC Lysis Buffer (Biolegend®) (1 x diluted with MaxPar water (Fluidigm®)) was added to blood at a ratio of 1:10 (i.e. 1ml of blood to 10ml of RBC Lysis Buffer)
1. The solution was incubated for 5-10 minutes at room temperature inverting the tube every few minutes. RBC lysis was deemed complete when the solution was opaque
2. Samples were centrifuged at 400*g* for 5 minutes to pellet white blood cells (WBC)
3. WBC were washed with 5ml of PBS (Sigma-Aldrich®), then centrifuged at 400*g* for 5 minutes, carefully pouring off the supernatant
4. Step 4 was repeated
5. WBC were counted using a counting chamber and a 10x microscope
6. 3-6 x 10^6 WBC per well were placed into a 96 well plate. If there was more than 6 x 10^6 WBC then the sample was split into 2 aliquots. One aliquot was re-suspended in freezing medium (10% DMSO (Sigma-Aldrich®) and 90% FBS (Thermofisher®)) and

placed in either a polystyrene box or Mr. Frosty™ (Thermofisher®) freezing container before entering a -80 C freezer.

3.5.7 *Density Gradient Centrifugation*

For the isolation of peripheral blood mononuclear cells (PBMC's) from whole blood the following protocol was used with a typical sample volume of 350-500ul (infant) or 2-3ml (adult)

1. Blood samples were re-suspended in HBSS (Sigma-Aldrich®) to final volume of 1.2ml (infant) or 4-6ml (adult)
2. In a universal container, the blood suspension was gently pipetted on top of, 5ml (infant) or 20ml (adult), lymphoprep (Stemcell technologies©)
3. The sample was centrifuged at 1200 *g* for 20 minutes with an acceleration of 6 and deceleration of 1.
4. The mononuclear layer, containing PBMC's, was identified and aspirated with a 2ml Pasteur© pipette.
5. PBMC's were washed twice with PBS
6. PBMC's were counted using a counting chamber and a 10x microscope
7. 3-6 x 10^6 PBMC's were isolated for staining. Should a sample contain more than 6 x 10^6 PBMC's, the sample was split and frozen as described in the RBC lysis section.

3.5.8 *Mass cytometry*

3.5.8.1 Antibody Panel design

A mass cytometry panel was designed to identify populations of T cells. The primary focus was the T_h response using chemokine receptors and transcription factors. The panel was designed to simultaneously explore T_{reg}, MAIT and iNKT cell populations. This was similar to a study exploring an adult population identifying T_h cells based on chemokine receptors and transcription factors (Kunicki et al., 2018). Dr Rebecca Payne designed the initial panel, using the MaxPar® Panel designer (Fluidigm®) to identify possible channel spillover issues, as well as published literature regarding T cell markers that tolerate freezing (Sumatoh et al., 2017). Table 9 details the initial antibody panel (Panel A) and the subsequent panels that were optimised over time. Panels C and D were used for the analysis of samples in this study.

Table 9 - Antibody panels used for mass cytometry. Changes made between panels are highlighted in red. *metal isotopes that were conjugated on site to a pure antibody

Metal Isotope	Panel A	Panel B	Panel C	Panel D
89Y	CD45	CD45	CD45	CD45
CD113*	CD19	CD19	CD19	CD19
141Pr	CD196 (CCR6)	CD196 (CCR6)	CD196 (CCR6)	CD196 (CCR6)
142Nd*	CTLA4	CTLA4	CTLA4	CTLA4
143Nd*	TCRgd	TCRgd	TCRgd	TCRgd
144Nd	CD3	CD3	CD3	CD3
145Nd	CD4	CD4	CD4	CD4
146Nd	CD8a	CD8a	CD8a	CD8a
147Sm*	CD161	CD161	CD161	CD161
148Nd*	CD86	CD86	CD86	CD86
149Sm	CD25 (IL-2R)	CD25 (IL-2R)	CD25 (IL-2R)	CD25 (IL-2R)
150Nd*	CD199 (CCR9)	CD199 (CCR9)	CD199 (CCR9)	CD199 (CCR9)
151Eu*	CXCR3	CXCR3	CXCR3	CXCR3
152Sm*			CD1d	CD1d
153Eu*	CD56	CD56	CD56	CD56
154Sm*	CD49b	CD49b	CD49b	CD49b
155Gd	CD45RA	CD45RA	CD45RA	CD45RA
156Gd*	CD335 (NKp46)	CD335 (NKp46)	CD335 (NKp46)	CD335 (NKp46)
		CD66b	CD66b	va24-FITC (iNKT)
157Gd*		ant-FITC	anti-FITC	anti-FITC
158Gd*	CD279 (PD-1)	CD279 (PD-1)	CD279 (PD-1)	CD279 (PD-1)

Metal Isotope	Panel A	Panel B	Panel C	Panel D
159Tb	CD197 (CCR7)	CD197 (CCR7)	CD197 (CCR7)	CD197 (CCR7)
	Valpha7.2-APC (MAIT)	Valpha7.2-APC (MAIT)	Valpha7.2-APC (MAIT)	Valpha7.2-APC (MAIT)
160Gd*	anti-APC	anti-APC	anti-APC	anti-APC
161Dy	Tbet	Tbet	Tbet	Tbet
162Dy	Foxp3	Foxp3	Foxp3	Foxp3
163Dy	CD294 (CRTH2)	CD294 (CRTH2)	CD294 (CRTH2)	CD294 (CRTH2)
164Dy*	CD69	CD69	CD69	CD69
165Ho*	TIGIT	TIGIT	TIGIT	TIGIT
166Er*	Granzyme B/perforin	Granzyme B/perforin	Granzyme B/perforin	Granzyme B/perforin
167Er	Gata3	Gata3	Gata3	Gata3
168Er*	ROR gamma (RORC, NR1F3)	ROR gamma (RORC, NR1F3)	ROR gamma (RORC, NR1F3)	ROR gamma (RORC, NR1F3)
169Tm*	CXCR5	CXCR5	CXCR5	CXCR5
	va24-PE (iNKT cells)	va24-PE (iNKT cells)	va24-PE (iNKT cells)	5-OP-RU(MAIT)-PE
170Er*	anti-PE PE001	anti-PE PE001	anti-PE PE001	anti-PE PE001
171Yb*	LAG3		CD28	CD28
172Yb*	CD14	CD14	CD14	CD14
173Yb*	CD117 (ckit)	CD117 (ckit)	CD117 (ckit)	CD117 (ckit)
174Yb	HLADR	HLADR	HLADR	HLADR
175Lu	CCR4	CCR4	CCR4	CCR4
176Yb	CD127 (IL-7Ra)	CD127 (IL-7Ra)	CD127 (IL-7Ra)	CD127 (IL-7Ra)

Metal Isotope	Panel A	Panel B	Panel C	Panel D
209Bi	CD16	CD16	CD16	CD16

3.5.8.2 Conjugation of monoclonal antibodies

As the number of conjugated antibodies commercially available for mass cytometry is currently limited, 23 antibody conjugations were performed on site. Purified antibodies were purchased as detailed in the resource table (Table 6) and conjugated using a Maxpar® X8 antibody Labelling Kit (Fluidigm) as per the manufacturer's instructions. A Nanodrop One™ (Thermofisher Scientific) was used before and after conjugation to measure the protein concentration. A titration experiment was then performed to identify optimal antibody concentration.

3.5.8.3 Staining of PBMC and WBC for mass cytometry analysis

Following isolation of PBMC or WBC, samples were placed in a 96 well plate. The protocol had the following steps:

1. Live dead stain using Cell-ID™ Cisplatin (Fluidigm®). Cells were required to be in concentration of $10x10^6$/ml. Cell-ID™ Cisplatin was added to give a final concentration of 5µM before incubation for 5 minutes
2. Primary antibody stain with antibodies bound to fluorochromes for 30 minutes.
3. Secondary antibody stain with remaining cell surface antibodies for 30 minutes.
4. Cell fixation; samples were suspended in 100µl PBS and 100µl 3.2% formaldehyde (16% formaldehyde diluted with PBS) added, before incubation for 1 hour.
5. 200µl of freezing media (10% DMSO in FBS) was added to a sample in 20-50µl PBS following centrifugation
6. Samples were immediately wrapped in parafilm and placed in either a polystyrene box or Mr. Frosty™, in a -80˚C freezer
7. One day prior to mass cytometry analysis, samples were defrosted at room temperature
8. Cells were permeabilised with 200µl Triton Perm Buffer (PBS + 2% FCS + 0.1% Triton X-100)
9. Cells were suspended in 50µl heparin (100 IU/ml) for 10 minutes
10. Intracellular antibodies were added and samples were incubated overnight

11. Cell-ID™ Intercalator (Fluidigm®) diluted to a concentration 1:1000 with Maxpar Fix/perm buffer. Cells were suspended in 100µl of solution for 1 hour.
12. Cells were washed twice with Maxpar® or MilliQ® water (Sigma Aldrich®) and filtered through a 40-micron nylon mesh into a Falcon™ tube (BD®)

Each staining step was performed in 50µl Maxpar® cell staining buffer with 2% FBS. Cells were washed with either PBS or MaxPar® cell staining buffer between each step.

3.5.8.4 Optimisation of mass cytometry panel

Following panel design, before the use of patient samples, optimisation was required to ensure effective antibody staining and to detect any significant spillover between channels.

3.5.8.4.1 Difficulty to detect antibody expression

As with any cytometry panel, optimisation to ensure optimal staining by each marker in the panel was required. Following conjugation of each antibody, titrations were performed to ensure specific staining of the cells of interest and to ensure there was not significant spillover into adjacent channels. The expression of LAG3 was not detected, likely either due to poor antibody binding to cells or metal isotope. As this marker was not crucial it was removed from the panel.

3.5.8.4.2 Signal spillover

Signal spillover can commonly occur in two ways in mass cytometry:

1. The signal in a channel for mass cytometry can spill over into the directly adjacent channels (+1 or –1).
2. Following ionisation by the argon plasma in the mass cytometry instrument, ions can form oxides, causing signal spillover into the +16 channel. Oxidation is more likely to happen with the heavy metals; lanthanum (La), cerium (Ce), praseodymium (Pr), and neodymium (Nd) (Fluidigm, 2016).

Unfortunately, there was spillover demonstrated in our panel from the 146Nd channel to 162Dy channel. This resulted in CD8+ cells in the 146Nd channel appearing to express FoxP3 in the 162Dy channel as demonstrated in Figure 9. As a result, FoxP3 was not used for clustering of cells as described in Section 3.6.5.

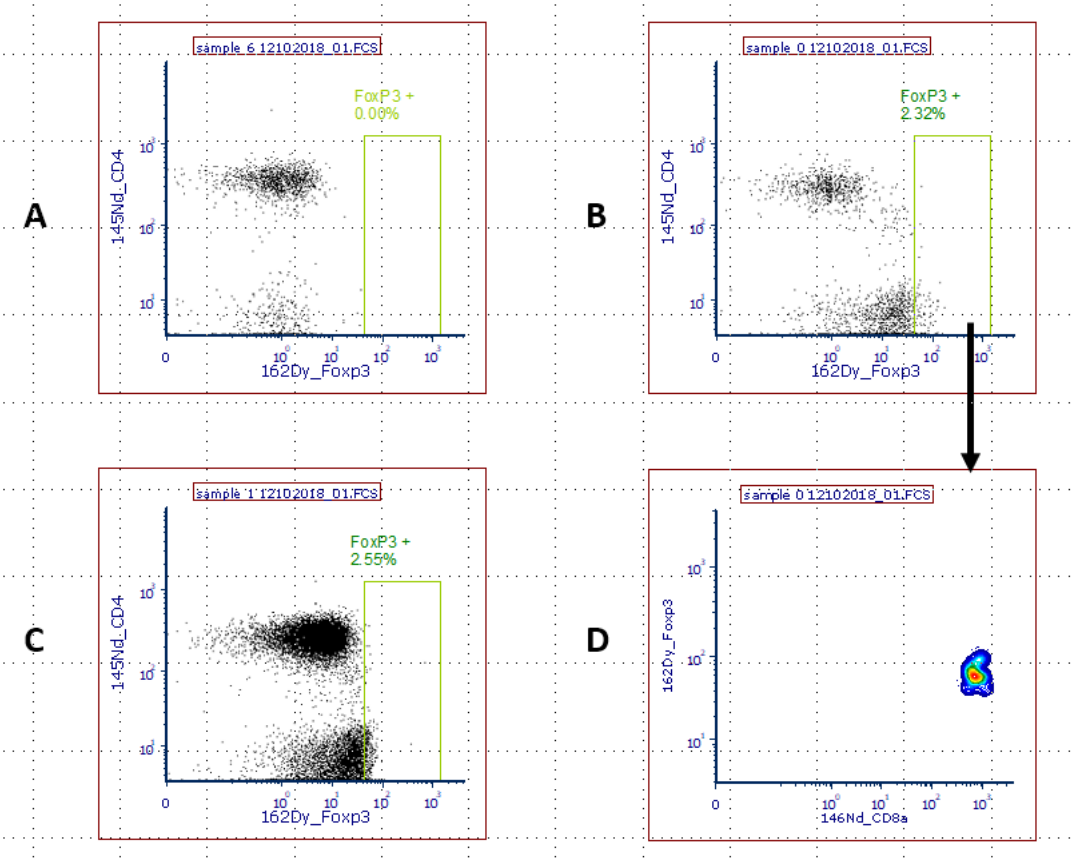


Figure 9 – Dot plots displaying T cells (CD45+CD3+), demonstrating spill over of signal from 146Nd to 162Dy channel. Axes have been scaled using Arcsinh to display all T cells. (A) Staining without 162Dy-FoxP3 or 146Nd-CD8α antibody. (B) Staining with 146Nd-CD8α antibody. (C) Staining with both 162Dy-FoxP3 and 146Nd-CD8α antibody. (D) Density plot demonstrating that FoxP3 + cells in (B) express CD8α

3.5.8.4.3 Heparin to reduce non-specific eosinophil staining

It has been shown that heparin reduced the non-specific binding of antibodies to eosinophils after cells have been fixed and permeabilised (Rahman et al., 2016). This was therefore incorporated into the staining with intracellular markers in our protocol as described in Section 3.5.8.3.

Following analysis of all samples, the expression of the transcription factors T-bet, Gata-3 and ROR-γ was found to be below that expected biologically (Kunicki et al., 2018, Pandya et al., 2016). A comparison was therefore made between samples stained using heparin and without heparin, which as demonstrated in Figure 10 showed that heparin reduces the binding of antibodies to intracellular T cell antigens. Figure 10 included only PBMC's, as such there should be no eosinophils present.

This result suggests the transcription factor markers used in the analysis cannot be relied upon to define biological populations, however the expression of transcription factors could be compared between different groups of samples stained using the same technique.

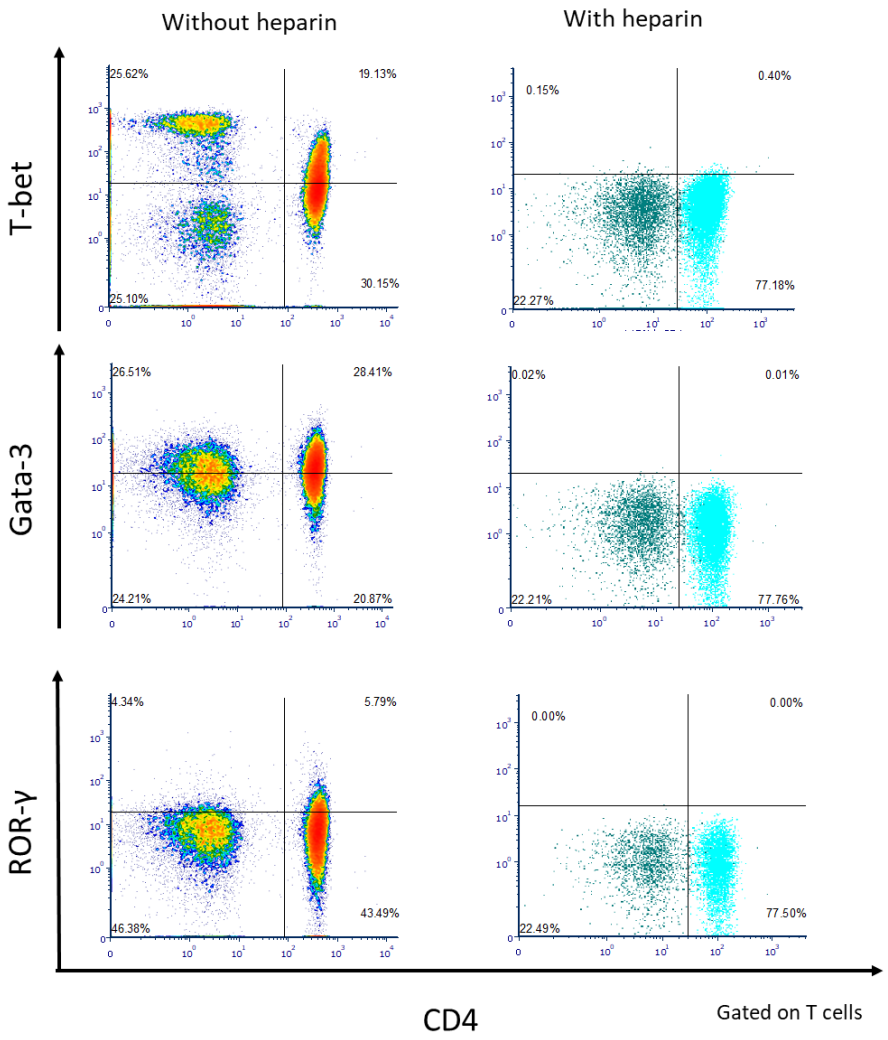


Figure 10 - Dot plots comparing intracellular staining of T cells with and without heparin. Axes have been scaled using Arcsinh to display all cells. Gates used to offer comparison between staining not to define cell populations. The plots on the left display density dot plots whilst the plots of the right are colour dot plots.

3.5.8.5 Control samples

3.5.8.5.1 Experimental controls

One adult control sample was stained every time an infant sample was stained. If multiple infant samples were stained then only one adult sample was simultaneously stained as a technical control for the staining procedure. This provided an experimental control to ensure consistent staining over time, whilst the batch controls (described below) controlled for

consistent analysis on the mass cytometry instrument over time. To assess whether there was consistent staining between samples, control samples were compared against each other using dot plots gating for each antigen. This demonstrated the staining of each marker and minimal change between controls.

3.5.8.5.2 Batch controls

Batch controls were used to ensure there was minimal variation between runs on the mass cytometer. Blood obtained from a single adult donor underwent cell surface staining and freezing in multiple aliquots. One batch control sample was defrosted and analysed with each mass cytometry run. The variation in batch controls was assessed by using a multi-dimensional scaling (MDS) plot as described in Section 3.6.5.1 which demonstrated that batch controls clustered closely together (Figure 11).

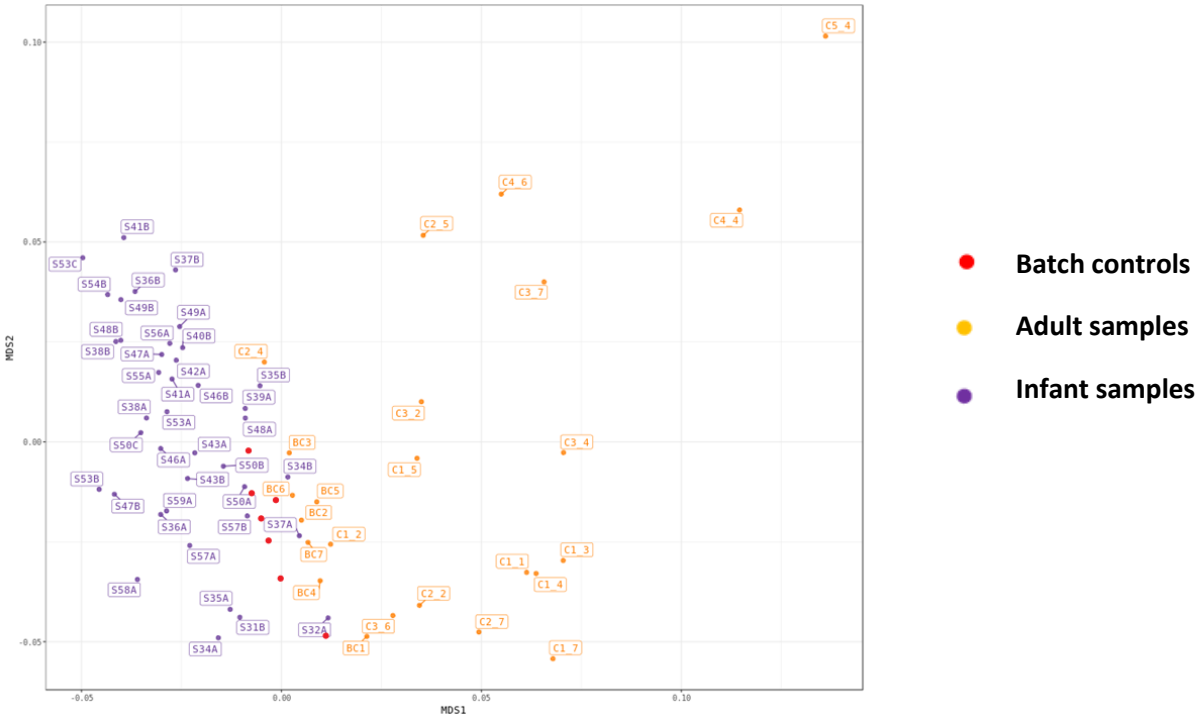


Figure 11 – Multi-dimensional scaling plot displaying infant and adult samples based on the median marker expression of 35 antibodies. Batch controls, adult and infant samples are represented in different colours as per the key to the right of the plot

3.5.8.6 Data acquisition

Prior to data acquisition, cells were counted using a BD Accuri™ C6 flow cytometer. Samples were then diluted in milli Q water containing 10% EQ™ Four Element Calibration Beads (Fluidigm®) to a final concentration of 5 x 10^5 cells/ml. Samples were acquired on a Helios mass cytometer (Fluidigm®). Sample acquisition occurred at a flow rate of 30 µL/min. The acquisition of samples was performed by a member of the Newcastle University Flow

Cytometry Core Facility. Each sample was run for 30 minutes aiming for an acquisition of 5 x 10^5 events per sample.

3.6 Statistical Analysis

When reporting the study, the Consolidated Standards of Reporting Trials (CONSORT) guidelines were used. Data from REDCap® were exported as a Microsoft Excel file. Statistical analyses of 16S and mass cytometry data were performed in R. All other analyses were performed in GraphPad Prism 8©.

When analysing clinical, biometric and nutritional data, the Kolmogorov-Smirnov, Shapiro-Wilk and Anderson-Darling tests were used to assess normality. Normality was assumed if 2/3 of these tests assessed the data as normally distributed. For normally distributed data, an unpaired t-test was applied. In the case of multiple comparisons, the two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, was used with a Q value of 5% to minimise the false discoveries. For non-normally distributed data a Mann-Whitney U test was performed.

3.6.1 *Growth*

Weight gain per day was calculated by dividing weight gain between first and last measurements in study period, by the difference in CGA between the two measurements. Change in length (cm/week) and change in OFC (cm/week) were calculated in a similar fashion. Rate of weight gain (g/kg/day) was calculated by dividing the change in weight by the average weight over the time period and then by the number of days between measurements. Rate of weight gain (g/kg/day) was calculated for the whole study, and from full enteral feeds to the study end. The calculation is again described in Section 4.1.6.

Z-scores derived from an international database of growth references were used (Fenton and Kim, 2013). A limitation of this growth reference is that it is based on birth-weights, however another publicly available growth standard based on postnatal growth only provides data from 27 weeks GA (Villar et al., 2015). Use of a combination of these reference data sets was considered. However, one database describes a North Atlantic population whilst the other describes infants from all over the world (predominantly Asia, Africa and South America), therefore they may not be comparable in terms of growth.

3.6.2 *Amino Acids*

Values for the 22 amino acids analysed were given as µmol/L. The samples were grouped by study group (Control or Intervention) and time-point (C or E). Samples were compared between study groups and time-points.

3.6.3 *Gut microbiome*

16S data was converted into tre and biom files as described above.

Together with a metadata file, the tre and biom files were uploaded onto a graphical user interface called ATIMA (Agile Toolkit for Incisive Microbial Analyses) developed by the Alkek Center for Metagenomics and Microbial Research (Baylor University, USA), which is run through R.

For identification of the impact of diet, samples were stratified by time-point (A-E) and then by study group. Furthermore, the use of MOM categories as described in Section 3.3.8.3 allowed the stratification of samples based on the amount of MOM received.

Mass cytometry data from blood samples (taken within 3 days of stool samples) were integrated into the sample metadata so that correlations with T cells populations could be explored. Linear regression analysis was used for these correlations.

Analysis of taxa (OTU) abundance focused on the genus and phylum taxonomic levels.

3.6.4 *Flow Cytometry analysis*

The Flow Cytometry Blood Sciences laboratory, RVI, Newcastle upon Tyne Hospitals provided both absolute cell counts (cell/microliter) and percentage of lymphocytes for T, B, NK, CD4 T, CD8 T and HLA-DR positive T cells. This was using a gating strategy as per the manufacturer's instructions.

In order to align flow and mass cytometry data, an unconventional gating method was used. Lymphocytes are identified using CD45 in mass cytometry whereas ordinarily forward and side scatter would be used in flow cytometry to identify the lymphocyte population. To offer comparison, lymphocytes were identified using CD45 in flow cytometry as shown in Figure 12.

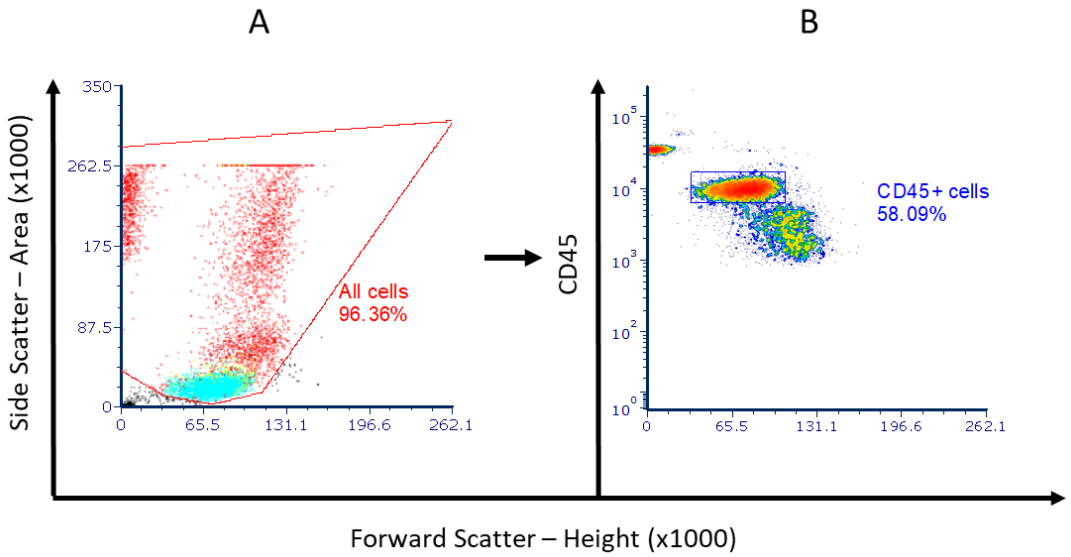


Figure 12 – Flow cytometry gating used to enable comparison with mass cytometry. A - All Cells were gated using forward and side scatter, lymphocytes highlighted in light blue are usually identified this way. B - CD45+ cells are identified for mass cytometry comparison, debris or calibration beads noted to be CD45+ on the y axis

3.6.5 *Mass Cytometry*

Following acquisition, data were exported in a flow cytometry standard (FCS) file format file. The FCS file was processed using CyTOF software (Fluidigm®). CyTOF software performed normalisation using the EQ™ Calibration beads followed by Gaussian normalisation of the data.

The resultant FCS files were then imported into FCS Express (De Novo Software) for the next part of the analysis. The gating strategy implemented in FCS Express, began with the exclusion of dead cells and doublets followed by positive gating of CD45+CD3+ T cells (Figure 13). They were then ensured to be CD19, CD66b, and CD14 negative. To explain further, Equilibrium beads were removed from the analysis in the first plot (Figure 13A). Live cells are then identified as cisplatin negative and iridium (DNA intercalator) positive (Figure 13B). Cytometry analysis can result in doublets, when two cells are processed at the same time, resulting in more DNA intercalator than single cells. Figure 13C displays the gating strategy used to identify single cells. The T cell gate was exported as a separate FCS file for analysis in R using the analysis described in Section 3.6.5.1.

The T cell population was further split into iNKT, MAIT and T_{reg} cells. iNKT cells were defined as T cells (CD45+CD3+) displaying the invariant TCR Vα24-Jα18. MAIT cells were defined as CD8+ T cells expressing CD161 and the invariant TCR Vα7.2. During the study, a tetramer for MR-1 was

supplied by the NIH Tetramer Core Facility (Atlanta, USA). This allowed for the definitive identification of MAIT cells. T_{reg}s were identified as CD4+T cells that were $CD25^{high}$ CD127- and FoxP3+. All populations were assessed to ensure there was no spillover from adjacent channels. FCS files containing these cells were then exported and uploaded into R for analysis.

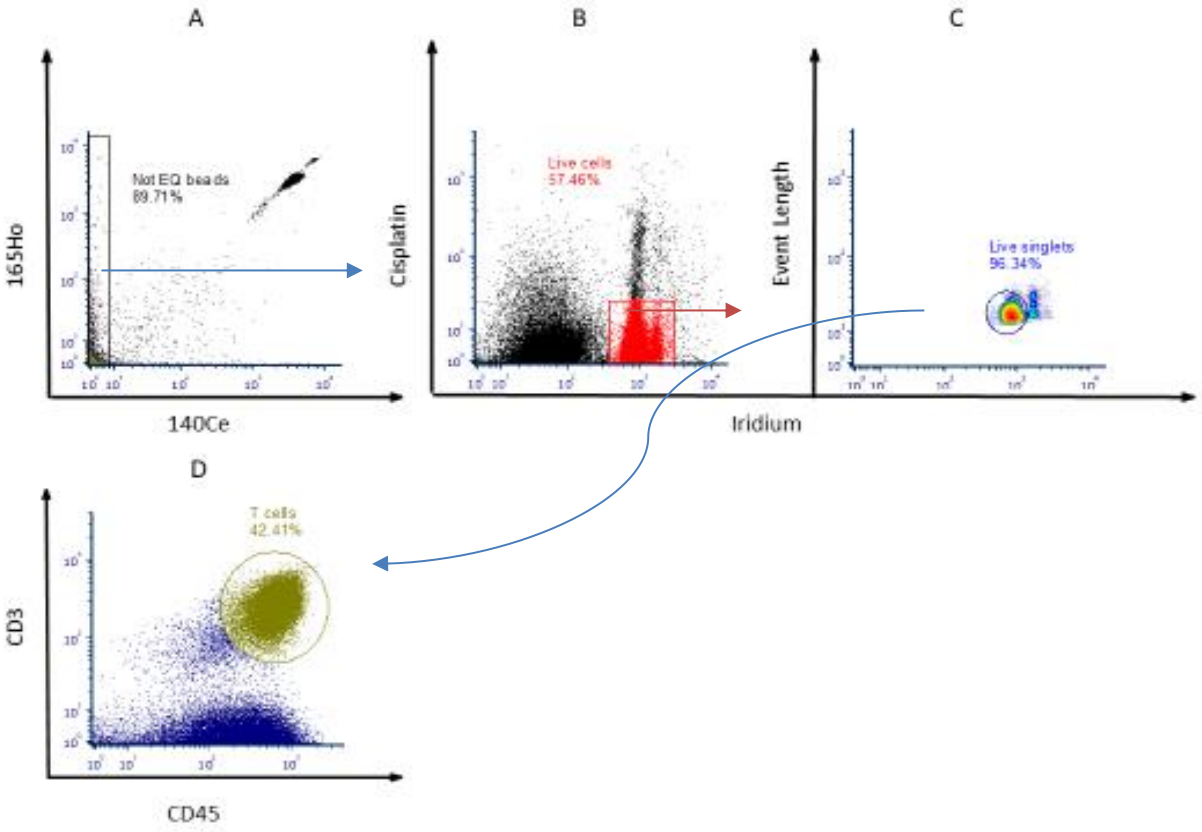


Figure 13 - Gating strategy used for identifying T cells using mass cytometry. A - Gate to remove Equilibrium beads (EQ) (on channels 140Ce and 165Ho). B - Identification of live cells. Live cells are iridium positive and cisplatin negative. C – identification of single live cells. Cells are iridium single positive, note the population to the right of the gated poplulation representing doublets. D – Identification of T cells as CD45+ CD3+.

3.6.5.1 Nowicka Analysis

FCS files containing only T cells were analysed using a workflow in R devised by Nowicka et al. (Nowicka et al., 2017). Briefly, an arcsinh transformation of the data to allow symmetrical measurement is performed. An initial description of the data using a histogram of marker expression and a bar graph of cell numbers is produced. A multi-dimensionality scaling (MDS)

plot is then generated offering a point for each sample based on the expression of each marker, as shown in Figure 11.

The MDS plot offers a broad comparison of the data stratified by variables. A non-redundancy score (NRS) is calculated from the MDS plot which indicates which markers have the most influence of the position of any sample in the MDS plot. The NRS was used to identify the markers to define the clusters in the t-distributed stochastic neighbour embedding (t-SNE) plots later in the workflow, as these markers were likely to contribute most to the differences between clusters. The NRS score for each marker in the mass cytometry panel is displayed in Figure 14. Using the NRS, and knowledge of important markers, 14 markers were chosen for the t-SNE plots in the further analysis. The cells in each sample were then down-sampled to 2000 cells, using known cell numbers, in order to limit the computational memory required and so that each sample could be compared. Dimensionality reduction was then performed using t-SNE. Clusters of cells were defined using Flow self-organising maps (FlowSOM). A heatmap is generated with the t-SNE plots, which allows for the identification of cell populations using a cluster dendrogram and their corresponding expression of multiple markers. A generalised linear mixed model is then used to compare the normalised expression of other markers (not used in the initial clustering) between cell clusters and defined variables (Nowicka et al., 2017).

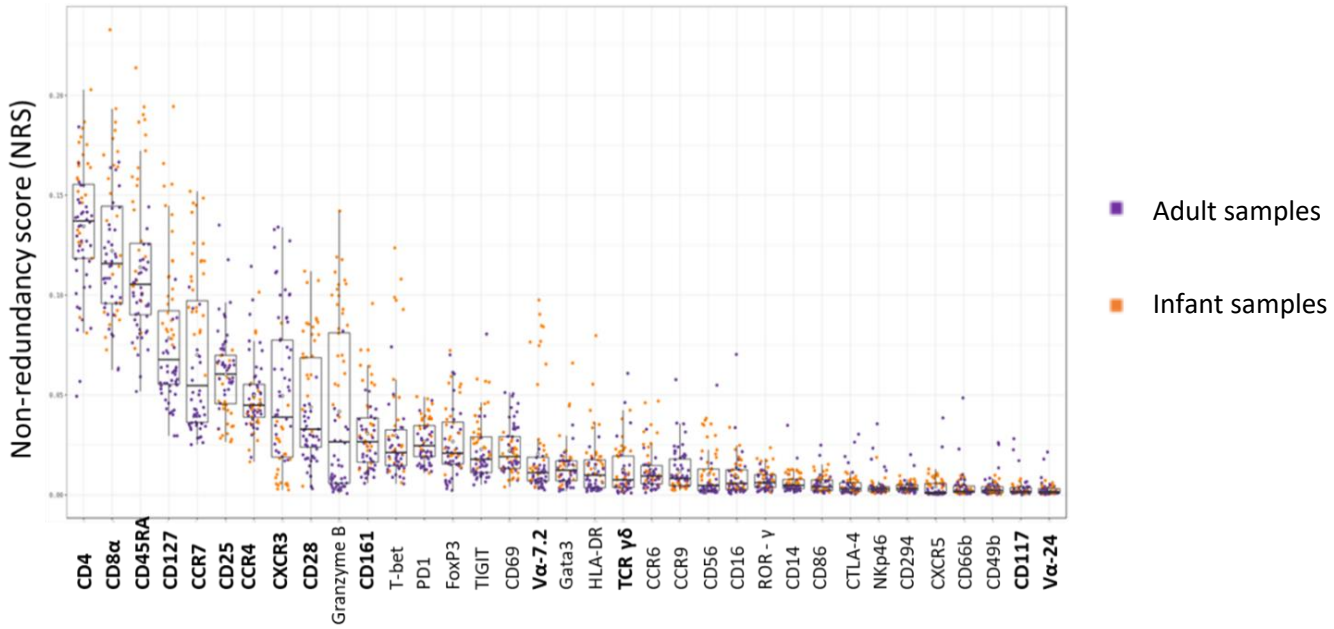


Figure 14 – Box plot displaying the NRS for each marker. 14 markers used to define clusters in the latter t-SNE plots are highlighted in bold.

3.6.5.2 Cytofkit analysis

Analysis of T cell populations (iNKT and MAIT cells) was performed using a bespoke package in R, named cytofkit (Chen et al., 2016). Cytofkit offers a number of ways to visualise (e.g. t-SNE) and interpret (eg. heatmaps) mass cytometry data. Unlike other mass cytometry packages there is an integrated user interface which makes analysis possible without bioinformatic expertise.

Populations of iNKT and MAIT cells were made into individual FCS files. These files were uploaded into cytofkit. All cells in each file were analysed. The FCS files underwent an arcsinh transformation before clustering using Rphenograph and visualisation using t-SNE plots, based on the expression of 34 markers. Clusters were labelled and merged using a heatmap of the marker expression. Cytofkit was used to assess expression of individual markers in expression plots. Differences in cluster abundance was compared between sample variables. Cluster abundance was derived from the cytofkit analysis however the statistical analysis was performed in Graphpad Prism 8©.

Chapter 4. Results

4.1 Clinical Data

4.1.1 *Participant Flow*

The study was open to recruitment at the NICU, Royal Victoria Infirmary on the 29th September 2017. Figure 15 describes the study recruitment and outcomes for infants between the 29th September 2017 and the 8th March 2019.

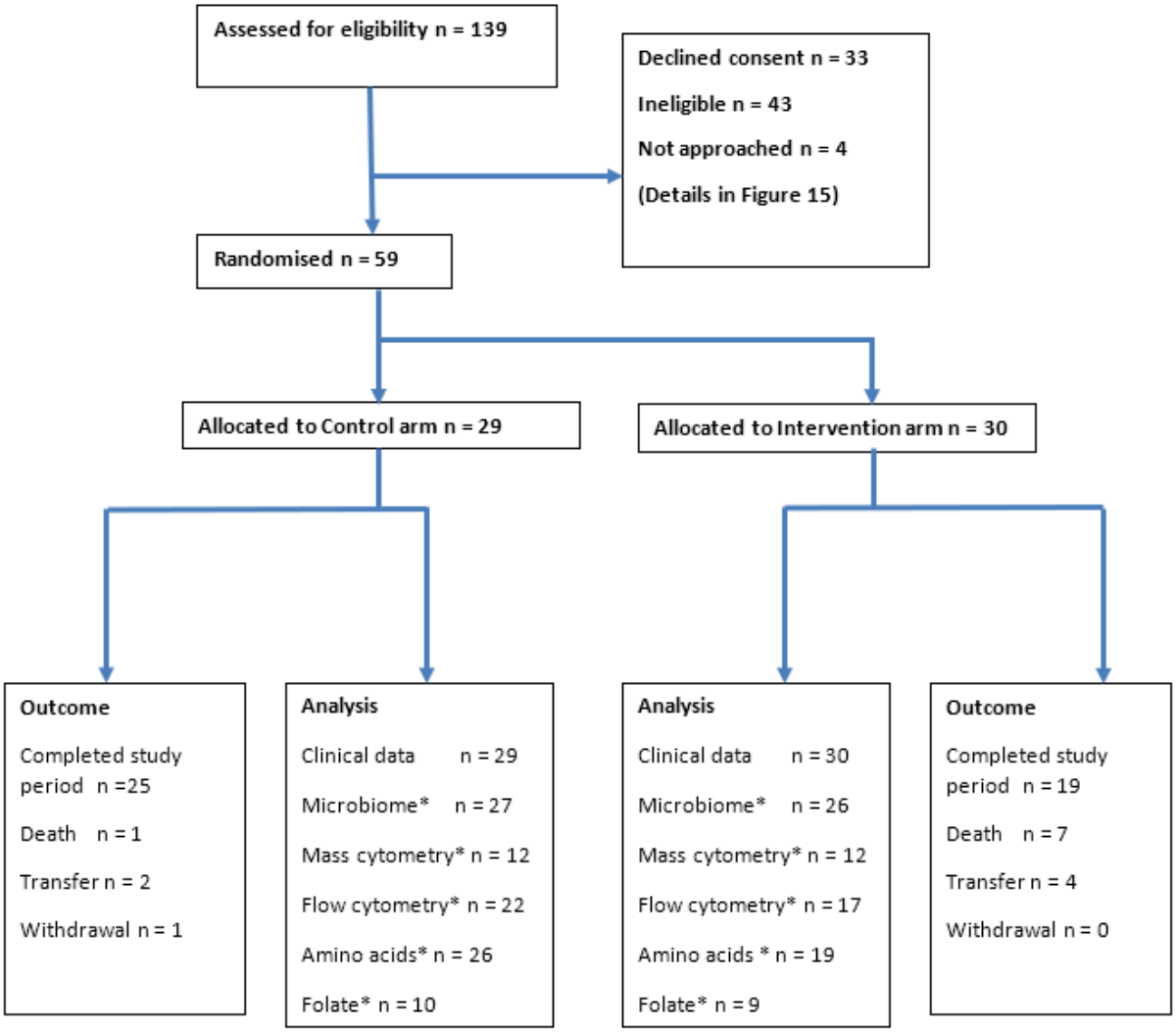


Figure 15 - Recruitment Flow for the INDIGO study *indicates at least one sample[+] provided clinical data following transfer as per study amendments

Declined consent n = 33

- No reason given	n = 17
Reasons given:	
- Not willing to give DHM	n = 7
- Twin randomisation	n = 2
- Baby unwell	n = 3
- Not willing to remain at study site	n =1
- Sibling unwell	n =1
- Religious reasons	n =1
- Struggling with preterm birth	n =1

Eligible, but not approached n = 4

Reason for not approaching:	
- Lack of parental capacity	n = 1
- Sibling demise	n = 1
- Study investigator unavailable	n = 1
- Unknown gestation	n = 1

Ineligible n = 43

Reason for ineligibility:	
- Transfer out before 72 hours	n = 11
- Transfer in after 72 hours	n = 25
- Congenital abnormality	n = 1
- Unlikely to survive	n = 2
- Likely to be transferred before study end	n = 3
- Unable to read PILS	n =1

Figure 16 - Reasons for non-participation of infants in the INDIGO study

All NICU admissions born less than 30 weeks GA were screened for eligibility. As detailed, 59 infants out of 92 eligible mother-infant pairs gave consent to the study giving a recruitment rate of 64%. Considering the infants not recruited (Figure 16), 4 potentially eligible families not approached for consent, one was due to lack of availability of a study investigator during a holiday period. Other families were all approached unless ineligible or deemed not appropriate to approach by the study investigators. However, it should be noted that 14 families were not approached either because they were transferred to another hospital within 72 hours of birth or were likely to be transferred before the end of the study period. The reason for this is that the RVI NICU is a surgical referral centre in the North East of England and so has a high demand on the cots for infants, as well as a large number of infants transferred into the NICU more than 72 hours after birth. The most common reason given for not participating in the study was that the parents were not willing for their child to have DHM.

4.1.2 *Demographics of study infants*

Table 10 - Admission demographics of study infants separated by study groups. P-values were calculated using Student t-test or Mann Whitney U test. Kolmogrov-Smirnov test was used to compare cumulative distributions.

	Control (n= 29)	Intervention (n =30)	p-value
White British (%)	24 (83)	29 (97)	0.93
Birthweight (g), mean (SD)	879 (261)	974 (295)	0.25
Gestation at birth (weeks), median (IQR)	27 (26-28)	27 (26-28)	0.87
Male (%)	16 (55)	19 (63)	1
C-section (%)	18 (62)	15 (50)	0.98
Multiple pregnancy (%)	5 (17)	6 (20)	1
Antenatal steroids, 2 doses (%)	18 (62)	18 (60)	1
Prolonged rupture of membranes (%)	6 (21)	7 (23)	1
Reversed placental end diastolic flow (%)	9 (31)	4 (13)	0.74
Heart rate >100 at 5 mins age (%)	28 (97)	26 (87)	1

The majority of the population in the North East of England is white British reflected in the study population. The control group had a lower mean birth weight, corresponding with a higher rate of reversed end diastolic flow(rEDF) in the umbilical artery. The groups were otherwise well matched for demographic information (Table 10). The rate of caesarean section of 50-62% is reflective of the preterm population, which is likely increased in the control group due to the increased rate of umbilical artery rEDF leading to caesarean section for fetal reasons.

4.1.3 *Dietary intervention*

During the study period, the 59 infants received a total of 428.3 litres of nutrition (parenteral or enteral), 354.2 litres of this nutrition was given enterally, and 215.7 litres was MOM. MOM therefore comprised 60.9% of all enteral nutrition by volume (total MOM volume / total enteral volume) for all infants in this pilot study.

Four infants in the intervention group did not complete the protocol in terms of the dietary intervention. Two infants received CMF 1 day before study completion at 33^{+6} weeks GA, 1 infant incorrectly received bovine fortifier at 33^{+0} weeks GA and 1 infant received a specialised feed (monogen) from 1 week of age due to a persistent chylous pleural effusion. All infants were included as part of an intention to treat analysis but only those samples (blood and stool) taken whilst adhering to the dietary intervention were used for the microbiota analysis.

As this study was not blinded, it was important to identify if there were variation in practice or exposures between study groups unrelated to study protocol which may have been related to bias in clinicians. Table 11 displays a comparison of dietary interventions and outcomes during the study period.

Table 11 - Dietary outcomes during study period for infants that completed the study. * Normally distributed data given as mean, SD; # Non-normally distributed data given as median, IQR; + Not all Infants received fortifier therefore number of infants contributing to analysis in brackets. Days of feeding exposures summed prior to 34 weeks CGA. p-values were calculated using an unpaired sample Student t test or Mann Whitney U test.

	Control (n =25)	**Intervention (n = 19)**	**p-value**
First feed, day of life	2, 2-3	2, 2-3	0.567
Day of life fully enterally fed #	13, 11-18	12, 12-15	0.814
Number of days enteral feeding withheld more than 4 hours #	0, 0-3	0, 0-5	0.975

	Control (n =25)	Intervention (n = 19)	p-value
Total days of parenteral nutrition #	12, 9-19	11, 9-19	0.777
Days fully enterally fed prior to 34 weeks GA*	32, 9	31, 9	0.61
Days on which any MOM received*	34, 16	32, 16	0.75
Days fully MOM fed #	23, 8-33	19, 0-28	0.290
Day of life fortifier started (n) #+	17, 11-21 (22)	13, 11-21 (15)	0.544
Receiving any MOM at study end, n (%)	16 (59)	10 (53)	

There was no significant difference in any of the dietary outcomes measured in Table 11. Only one infant did not tolerate full feeds by 34 weeks CGA, this infant was in the control group and had a high stoma output following necrotising enterocolitis therefore was still receiving some parenteral nutrition. At the end of the study period, 59% and 53% of infants in the control and intervention group respectively were receiving some MOM as part of their diet, suggesting the intervention did not influence mothers' decision to continue expressing milk.

The donated human milk fortifier used in the study was a liquid of 30ml volume whereas the cow's milk fortifier was powder. This meant that every 100ml of "fortified MOM" contained 100mls MOM in the control group, but 30ml human milk fortifier(P+6) and 70ml MOM in the intervention group. We therefore measured the actual volume of MOM received in each study group.

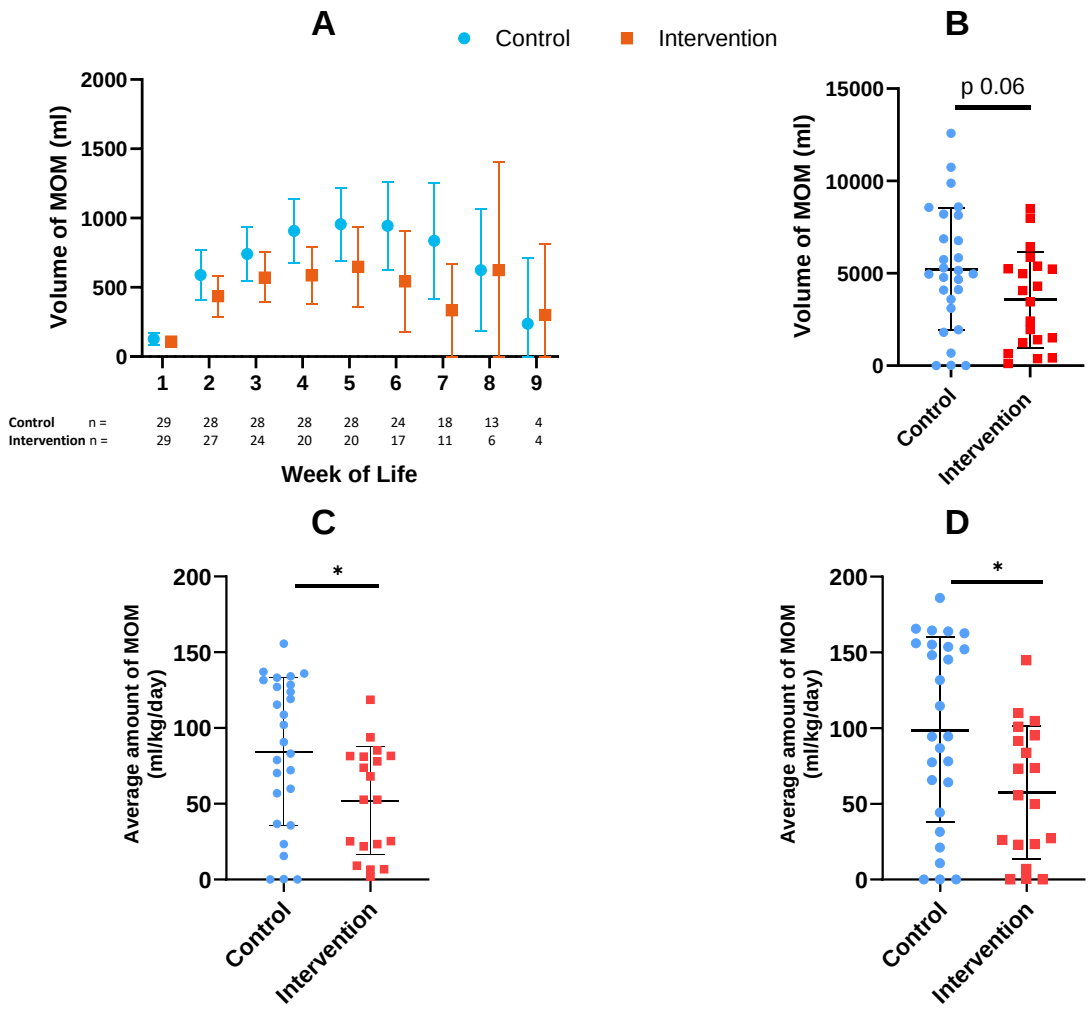


Figure 17 - Comparison of volume of MOM received between study groups. (A) Volume of MOM per study group over weeks of the study. Box represents mean and whiskers represent 95% confidence interval (CI) of the mean. (B) Scatter plot comparing total volume of MOM received in study period. (C) Scatter plot comparing total volume of MOM as ml/kg/day over whole study period. (D) Scatter plot comparing total volume of MOM as ml/kg/day from DOL 14 to study end. For scatter plots, line represents mean, whiskers represent SD. Values were compared using unpaired Student t test with a Welch correction, p values are displayed above the plots (p <0.05, *; <0.01, **; <0.001, *)**

A comparison was made between the weekly intake of MOM between study groups (Figure 17A). No statistical difference was found when comparing each week between study groups, although there was a clear trend towards higher MOM intake among infants in the control group.

Equally the total volume of MOM received by infants was not significantly different but there is a trend towards infants receiving more MOM in the control group. Figure 17B represents the participants for whom there was complete nutritional data at 34 weeks GA (control n =27, intervention n = 19). The characteristics of the infants that completed the study varied both in terms of weight, with a mean birth weight (BW) of 1054 (SD 356) vs 883 (SD 302) in the intervention and control groups respectively and in terms of duration, with a median number of days in the study of 43 (IQR 42-54) vs 45 (IQR 39-56) in the control and intervention group respectively. In order to correct for this discrepancy, I compared the volume each infant received as ml/kg/day. The infants in the intervention group received significantly less MOM than the infants in the control group after adjusting for weight and time in the study as shown in Figure 17C. Furthermore, I aimed to ensure that this difference was not due to difficulties in the mothers establishing good volumes of expressed MOM in the intervention group. I therefore compared the amount of MOM that each infant received between DOL 14 and the study end as most infants would be fully enterally fed around DOL 14 (Table 15). The significant difference in volume of MOM received between the two groups persisted after this comparison.

The most obvious reason for the difference in MOM received, shown in Figure 17, is that the fortifier used in the intervention group had a 30ml volume, whilst the control group used a powder fortifier which would not alter the volume of MOM. Therefore, if both groups had the same availability of MOM, the intervention group would receive 70% of the MOM that the control group received, when infants were fully enterally fed. However, when comparing the mean volume (ml/kg/day) of MOM received between the two groups from DOL 14 to the study end, the control group received a mean volume of 98.8 ml/kg/day whilst the intervention group received a mean volume of 57.4 ml/kg/day, meaning the intervention group mean was 59.2% (intervention mean volume / control mean volume) of the control mean. This is despite there being no significant differences in the variance of the values using an F test (p 0.157). The difference therefore, does not appear to be purely down to the fortifier volume but rather to limiting availability of MOM in the intervention arm of the study. This raises the possibility that this difference may reflect maternal attitudes towards expressing when infants are receiving DHM. There is a possibility that some mothers in the intervention group may have viewed DHM as ‘equivalent’ to MOM and therefore did not continue to express MOM as readily as mothers in the control group. It is also possible that the attitudes of staff towards the mothers differed, for example in continuing to support the mothers as actively. Alternatively, this finding may be due to chance as the study only compares small numbers of infants.

4.1.4 *Morbidity and mortality*

Tables 12-15 display the mortality and morbidity between study groups and relevant exposures. Nine infants died before hospital discharge with 8 infants dying during the study period. Mortality in the intervention group was higher than the control group. Deaths occurred at a median age of DOL 13 (Interquartile range (IQR) 9-22). The primary cause of death is given in Table 13. All causes of death are given in Table 14, categorised into 5 reasons as described in Section 3.3.8.4. All deaths were reviewed by the study investigators and no deaths were felt to be directly related to the study interventions.

Table 12 - Comparison of mortality between study groups and exposure to dietary interventions (during the study period)

	Control	**Intervention**
Death (number, %)	2 (7)	7 (23)
Day of life (values or median (+/- IQR))	34, 133	13 (8-17)
DHM or CMF as % of all nutrition (incl. TPN), values or median, range	0, 19	6 (0-31)
DHM or CMF as % of enteral nutrition, values or median, range	0, 100	55 (0-99)
Total volume of enteral nutrition (ml) (values or mean (+/- SD))	422, 742	572 (1046)
Total volume (ml) of CMF or DHM, (mean, SD)	0, 742	84 (132)
Exposed to fortifier, n (%)	0 (0%)	1(14%)

Table 13 – Details of deaths before hospital discharge, with regards to dietary exposures. NB: Only the primary cause of death is listed *Congenital anomaly was not recognised in this infant until after study recruitment

Study number	Study group	GA (weeks), birthweight (grams)	Age at death (DOL)	Prior to death:			Primary cause of death
				Days of MOM	Days exposed to RTF or CMF	Days of fortifier	
4	Intervention	27, 540	12	8	8	0	Respiratory failure
13	Intervention	24, 640	3	0	0	0	Intraventricular haemorrhage
16	Intervention	24, 470	19	1	15	0	Respiratory failure
19	Intervention	25, 770	6	3	2	0	LOS
24	Intervention	25, 895	10	4	7	0	Congenital anomaly*
25	Control	28, 500	133	0	117	0	Chronic Lung Disease
27	Intervention	26, 860	15	9	0	0	Intraventricular haemorrhage
33	Control	23, 545	34	23	0	0	LOS
52	Intervention	24, 550	32	29	0	0	Congenital infection

Table 14 – Comparison of cause of death between study groups categorised into common causes. *Congenital anomaly was not recognised until after recruitment (NB. a death may have been attributable to more than one cause)

	Control (n=2)	Intervention (n=7)
Respiratory Failure	2	5
Brain injury	1	2
Infection	1	2
Necrotising Enterocolitis	1	1
Congenital Anomaly	0	1*

Table 15 – Common morbidities of preterm infants per study group at hospital discharge or time of death. *serious morbidity was defined as LOS, NEC, IVH Grade 3 or 4, ROP requiring surgery, abdominal surgery(non-NEC) or moderate CLD (with oxygen requirement at 36 weeks CGA)

	Control **n = 29**	**Intervention** **n =30**
LOS, n (%)	5 (17%)	6 (20%)
NEC n (%) (medical, surgical)	4 (14%) (2,2)	1 (3%) (0,1)
NEC cases exposed to formula (RTF 26 or CMF), and/or fortifier, n	2	0
Retinopathy of prematurity, any, n (%)	10 (34%)	7 (24%)
Retinopathy of prematurity, treated (laser or injection), n (%)	3 (10%)	4 (13%)
Intraventricular haemorrhage (IVH), any, n (%)	8 (28%)	2 (7%)
IVH severity, grade 1-2, grade 3, grade 4, n	5 ,0, 3	0, 0, 2
Chronic Lung Disease, n (%) (mild, moderate, severe(n))	24 (83%) (1, 15, 7)	20 (67%) (6, 12, 3)
Spontaneous Intestinal Perforation (non-NEC), n (%)	1 (3%)	0 (0%)
Discharge without serious morbidity*, n (%)	5 (17%)	6 (20%)
Discharge without NEC or LOS, n (%)	21 (72%)	19 (63%)

Table 15 details the common morbidities found in this population. LOS was due to Coagulase negative Staphylococcus (n=7), *Staphylococcus aureus* (n=2) and *Enterococcus faecalis* (n=2) (Table 16). One infant, who sadly died, had *Enterococcus faecalis* isolated from blood cultures as well as *Serratia marcescens* isolated from throat swab, liver, spleen and lung tissue at post-mortem. A further case of *Enterococcus faecalis* LOS was preceded 20 days earlier by an episode of necrotising enterocolitis. Abdominal surgery (non-NEC) was performed in one case of spontaneous intestinal perforation.

Caution should be taken in interpreting the numbers in Table 15 as this study was not powered for clinical outcomes, there was disparity in the mortality rate between study groups and there was variable exposure to dietary interventions. An example of caution is in terms of NEC, as only 2/4 infants who developed NEC in the control group were exposed to bovine products, 1 to CMF and 1 to both CMF and fortifier. Survival to discharge without any serious morbidities was similar in the two groups. Serious morbidity was defined as NEC, LOS, Abdominal surgery (non-NEC), IVH (Grade 3 or 4), ROP receiving treatment or CLD (moderate or severe).

Table 16 - Comparison of organisms causing late onset sepsis between study groups. Numbers represent number of cases of LOS for each organism.

Organism	Control	Intervention
Coagulase negative Staphylococcus (n)	3	4
Staphylococcus aureus (n)	1	1
Enterococcus faecalis (n)	1	1

4.1.5 *Growth*

A concern of using unfortified DHM in preterm infants has been poor growth. This study used fortified DHM(RTF 26) and a HMF (P+6) which had a nutritional content that aims to match international recommendations and is similar to CMF and BMF as described in Section 3.3.7. We therefore did not expect any significant difference in the growth of infants depending on their study group.

Table 17 describes infants that had at least one measurement between 32-35 weeks corrected gestational age and had completed at least 4 weeks in the study.

In order to correct for the length of time each infant was in the study and starting weight I compared grams/kilogram/day change between study groups. The following calculation was used:

(Weight at end of study – birthweight(grams)) / (weight at study end – birthweight (kg))/2 / (number of days between measurements)

Using the whole study period includes comparison when infants were parenterally fed. In order to compare growth when infants were enterally fed I compared growth between DOL 14 and the end of the study period. Infants were fully enterally fed on a median of DOL 13 and DOL 12 in the intervention and control groups respectively (Table 11), hence the use of DOL 14. The following calculation was used to calculate growth (g/kg/day) in this period (DOL 14 to study end):

(Weight at end of study – weight at or after DOL 14 (grams)) / ((weight at study end – weight at or after DOL 14 (kg))/2) / (number of days between weight measurements)

Table 17 - Comparison of growth between study groups during study period. Z-scores were calculated using Fenton et al. database (Fenton and Kim, 2013). p-values were calculated using an unpaired sample t-test. [+] Mann-Whitney U test as not normally distributed, median values are given for weight z-scores

	Control (n =27)	**Intervention (n =19)**	**Standard error of mean (95% CI), p-value**
Rate of Growth from birth to study end (mean, SD)			
Weight (g/day)	17.8 (4.5)	17.5 (3.9)	-0.25 (-2.84 – 2.337), 0.85
Weight (g/kg/day)	10.4 (1.5)	9.4 (1.3)	**-1 (-1.9 to -0.2), 0.027**
Length (cm/week)	1.02 (0.25)	0.95 (0.23)	-0.06 (-0.22 to 0.09), 0.405
OFC (cm/week)	0.85 (0.21)	0.80 (0.19)	-0.05 (-0.17 to 0.07), 0.399
Rate of Growth from DOL 14 to study end (mean, SD)			
Weight (g/kg/day)	17.2 (4)	14.8 (3.3)	**-2.4 (-4.7 to -0.17), 0.035**
Mean change in Z-scores, SD (z-score at study end – z-score from first measurement)			
Weight[+]	-0.8 (-1.2 to -0.15) [+]	-1.1 (-1.4 to -0.9) [+]	-0.3, 0.0745[+]
Length	-0.78 (0.70)	-0.99 (0.64)	-0.22 (-0.65 to 0.21), 0.312
OFC	-0.23 (1.13)	-0.26 (1.08)	-0.04 (-0.72 to 0.65), 0.918

There was no significant difference comparing gross growth (g/day) between study groups, however when correcting for weight and time in study as described above there was a significant difference in rate of growth (g/kg/day). This persisted even when comparing rate of growth from full enteral feeds (DOL 14) to the study end suggesting it was related to the

enteral diet rather than growth whilst parenterally fed (Table 17 and Figure 18). The mean change in weight z-score was not significantly lower in the intervention group (p 0.0745). Fenton growth charts were used to calculate z-scores; therefore, the z-scores are compared against intra-uterine, not post-natal growth.

These data suggest that rate of weight gain is lower in infants in the intervention arm compared to the control arm. This result was not due to difference in energy intake as the calorific intake was matched between study groups as discussed in Section 3.3.7.

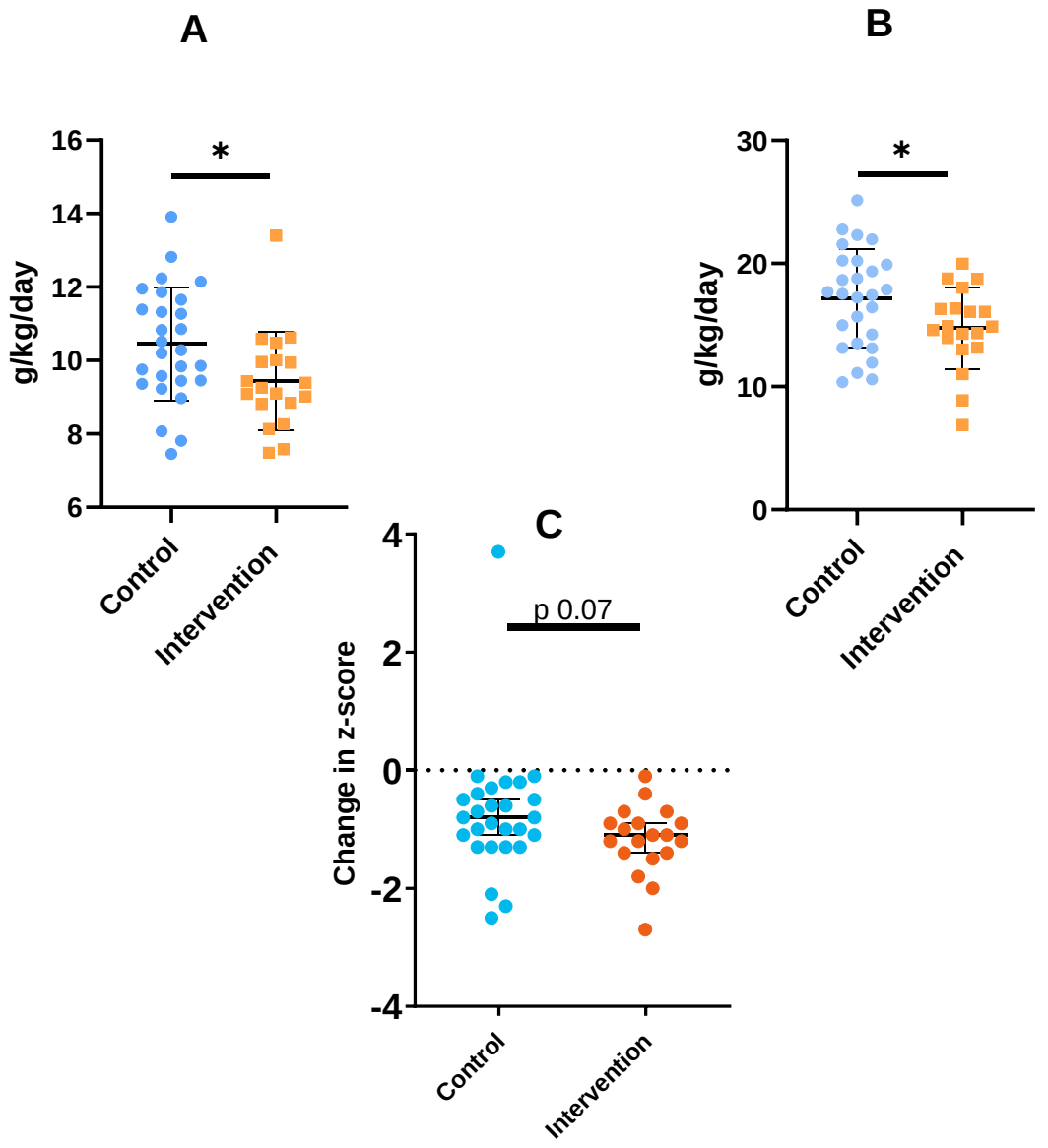


Figure 18 - Comparison of change in weight between study groups. A – Rate of growth during whole study period (g/kg/day). B – Rate of growth from DOL 14 to study end (g/kg/day) C - Change in weight z-score during whole study period (birth to end of study period). The bold line represents the mean with the error bars representing standard deviation. Comparison between groups was made using unpaired sample t tests.

4.1.6 *Other interventions likely to impact on gut microbiome*

As described in Chapter 1.2.6, a number of other factors especially those related to feeding can impact on the gut microbiome. A comparison between the factors that may impact on the gut microbiome is made between the study groups in Table 18.

Table 18 – Comparison of environmental exposures between study groups during the study period. Numbers represent days of exposure per infant between birth and 34 weeks corrected gestational age. Comparison made using Mann Whitney U test

	Control n = 27	Intervention n = 19	p-value
Antibiotics (median, IQR)	7 (4-16)	7 (3-11)	0.568
Probiotics (median, IQR)	36 (26-41)	34 (27-41)	0.654
Multivitamins (median, IQR)	31 (27- 36)	33 (24-39)	0.719
Iron (median, IQR)	0 (0-7)	0 (0-5)	0.916

There was little difference between the study groups. All infants received antibiotics, probiotics (Labinic drops™, Biofloratech Ltd.) and multivitamins. Only 14 infants received any enteral iron supplementation during the study period, 9/27 (33%) and 5/19 (26%) infants in the control and intervention groups respectively.

4.1.7 *Amino acid profiles*

A total of 70 samples from 47 infants were analysed for 22 different amino acids. Table 19 provides comparison between the study groups with regards DOL and CGA of sampling. I hypothesised that both dietary regimes would result in normal AA profiles and that there would be no significant differences between the groups.

Table 19 - Comparison of samples used for amino acid analysis between study groups.

Sampling time-point (see Graphical Methods Section 3.1)	**Time-point (TP) C**	**Time-point (TP) E**
n = number of samples (control, intervention)	n = 40 (23, 17)	n = 30 (18,12)
Control (n = 41) **DOL median (IQR)** **CGA (IQR)**	13 (11-15) 29^{+4} (28^{+2} – 29^{+6})	45 (39-48) 33^{+3}(33^{+0}- 33^{+4})
Intervention (n = 29) **DOL median (IQR)** **CGA (IQR)**	14 (13-17) 29^{+6}(28^{+4} – 30^{1})	38 (34-43) 33^{+2}(32^{+6}- 33^{+4})

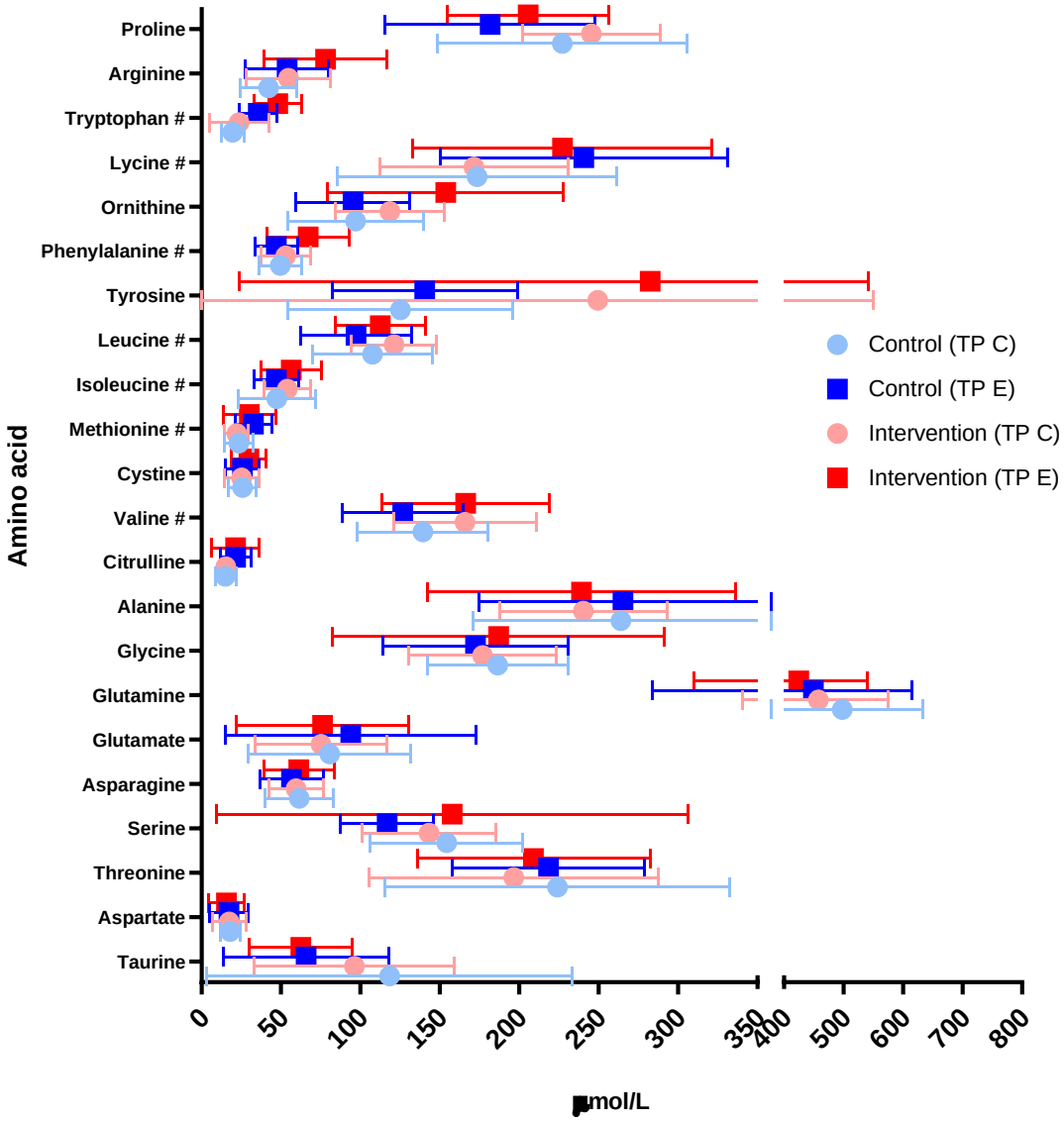


Figure 19 – Box and whisker plot comparing concentrations of 22 amino acids (9 essential) between study groups and time-points. Mean is represented by either circle or square, whiskers show standard deviation. TP C refers to time-point C, TP E refers to time-point E. A gap has been inserted into the x axis with a change in the right x axis tick interval. # indicates an essential amino acid

No statistical difference was found overall between the two dietary groups at either the early or later time point. However, when comparing samples in each study group separately, there was a significant increase in tryptophan in both the control infants (FDR p = 0.000003), and intervention group (FDR p = 0.019) between the two time-points. Whilst the majority of amino acid levels were within the normal range as defined by the clinical laboratory, four infants were found to have significantly increased tyrosine levels (3-5 times greater than upper reference range), and a fifth infant had borderline high tyrosine (1.5 times upper limit), all of which had resolved on repeat testing by 6 months of age. All 5 infants were in the intervention

group. The significance of this is unclear but transient tyrosinaemia has previously been reported in the preterm population (Juhl et al., 2018, Ventura and Brooke, 1987). Tyrosine can either be ingested in the diet or derived from the breakdown of another amino acid, phenylalanine in the liver, and is therefore not truly essential. The amount of tyrosine varies in human milk as the milk matures with higher concentrations in colostrum, with phenylalanine following a similar trajectory (Zhang et al., 2013). Transient tyrosinaemia, without an underlying error of metabolism, is thought to be benign with no impairment in neurological outcomes in preterm infants found, however there are a lack of recent studies in this area (Menkes et al., 1966). It has been suggested that term infants may have impaired neurological outcomes in later life following transient tyrosinaemia (Rice et al., 1989).

In order to ensure that these outliers were not having an undue effect on the analysis, a method that combines non-linear regression and outlier removal was applied to the dataset. The Q value was set to 0.1%, aiming for no more than 0.1% of the identified outliers to be false (Motulsky and Brown, 2006). This method removed 23 outliers from the analysis. A comparison was again repeated between study groups at the two time-points (Figure 20). Following a correction for false discovery using a Benjamini, Krieger and Yekutieli test, no mean concentrations of amino acids were found to be different at either time point between the two study groups (Benjamini et al., 2006).

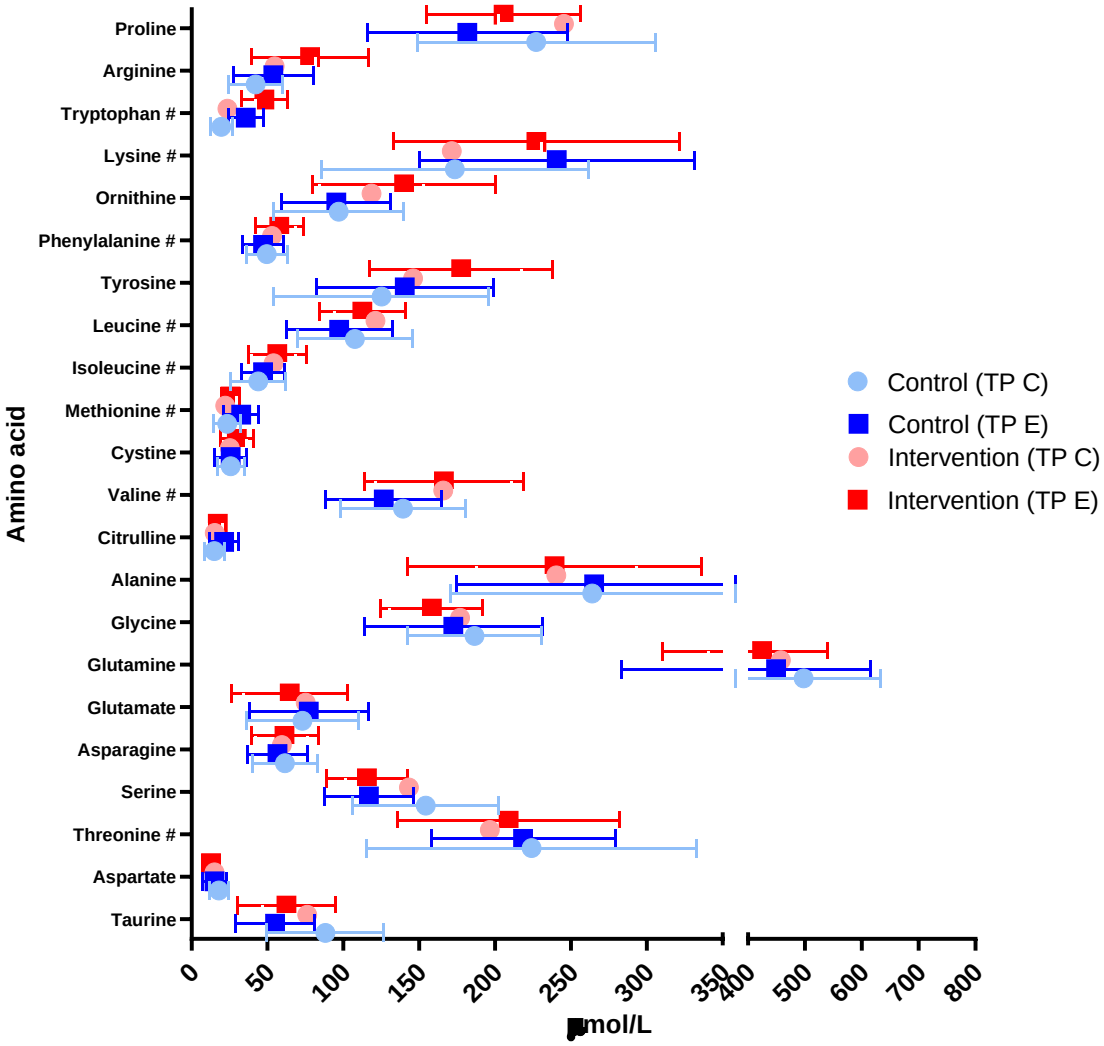


Figure 20 - Box and whisker plot comparing concentrations of 22 amino acids (9 essential) between study groups and time-points. Mean is represented by either circle or square, whiskers show standard deviation. TP C refers to time-point C, TP E refers to time-point E. A gap has been inserted into the x axis with a change in the right x axis tick interval. # indicates an essential amino acid. Groups were compared using student t test.

4.1.8 *Serum folate concentration*

There was a difference in the amount of folate provided by the two diets used in the study (Section 3.3.7) so we aimed to determine serum levels in a subset of patients.

Serum folate was measured in 19 patients (9 in the Intervention group, 10 in the Control group). Any results above 20µg/l were given as greater than 20 µg/l only. Serum folate was

measured at a median of 19 and 24 days following the start of fortifier i.e. when stable on full enteral feeds, in the intervention and control groups respectively.

The serum folate was greater than 20 µg/l in 9 of the 10 samples in the control group whilst the median value in the intervention group was significantly lower at 8.2 µg/l (range 3.5 – 12.6 µg/l). All but one value (3.5 µg/l) in the intervention group were within the normal range for serum folate used by the Newcastle Upon Tyne Hospitals laboratories, which is 3.9-12.6 µg/l. This was an infant born at 29 weeks GA who had folate measured at 33 weeks CGA. The infant was not symptomatic and had a normal haemoglobin and mean corpuscular volume in their full blood count.

Whilst the reported increase (control group) and reduction (intervention group) in serum folate did not appear clinically significant, these results emphasise that preterm infants are susceptible to biochemical disturbances, and clinicians should be vigilant as to nutrient intakes and the likelihood of them meeting recommended levels.

4.2 Gut microbiome

DNA was extracted from 218 stool samples from 53 study infants, and the 16S rRNA amplified using PCR. Rarefaction was performed at 3000 reads and 5/218 samples were excluded from analysis due to a low number of reads. Four of these samples were in the first 5 days of life with the other sample taken on day of life 11 and are likely to represent low levels of colonisation in an immature gut. Table 20 offers a comparison between stool samples after rarefaction in each study group at the 5 time-points (A-E). Samples in time-point C were taken at a median of 5 days earlier in the intervention group due to availability of samples. There was minimal overlap between time-points.

Table 20 - Comparison of the timing of stool samples analysed between study groups. Values represent day of life stool samples were taken with median (IQR) displayed

Sampling time-point	A	B	C	D	E
n = control, intervention	n = 23, 23	n = 21, 21	n = 20, 22	n = 22, 20	n = 23, 18
Control (median DOL (IQR)) n = 109	6 (4-8)	11 (10-12)	20 (15-23)	27 (25-28)	44 (39-50)
Intervention (median DOL (IQR)) n = 104	6 (4-7)	10 (9-11)	15 (14-21)	27 (24-28)	39 (34-46)

We sought to identify any differential change in α or β diversity between study groups at different time-points, and whether there was any difference in abundance of key taxa (OTU). Relative abundance was defined as the percentage of annotated reads at a particular taxonomic level. The null hypothesis was that there would be no association between the dietary intervention and the gut microbiome, either in terms of alpha-diversity, beta-diversity or composition at genus level, at 34 weeks CGA.

4.2.1 *No change in alpha diversity between study groups*

Figure 21-24 display the difference in alpha diversity. Figures 21 and 23 compare the Shannon diversity and observed OTU's (richness) between the study groups respectively, stratified by

time-point (A-E). Figures 22 and 24 demonstrate the change in Shannon diversity and observed OTU's (richness) over time respectively.

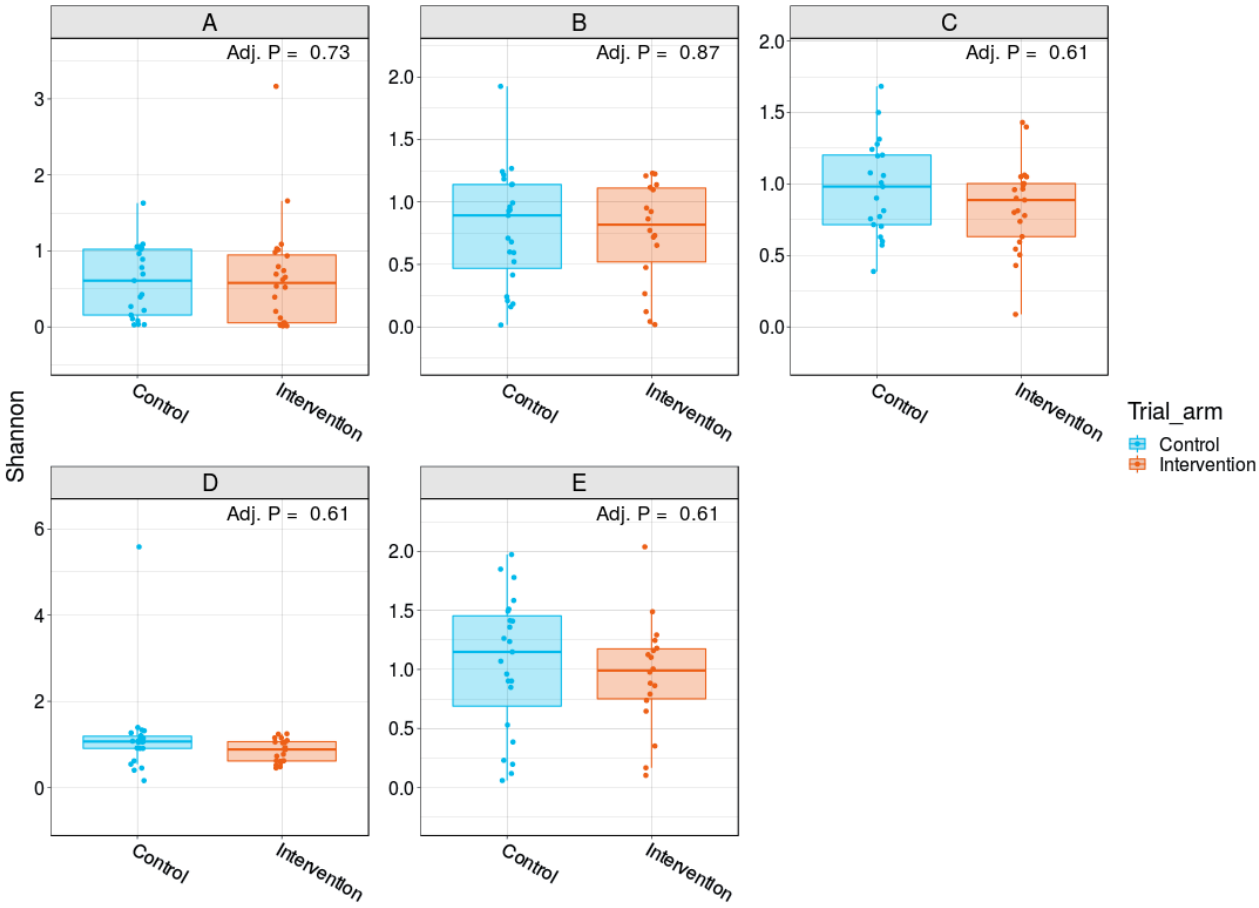


Figure 21 (A-E) - Comparison of Shannon alpha diversity displaying all samples stratified by study group from time-point A-E. All values are displayed on the plots. FDR adjusted p-values using a Mann-Whitney U test are displayed on the plot. The y-axis is scaled for each plot. The line represents the median value whilst the box represents the 25th and 75th centiles, the whiskers are a maximum of 1.5 x IQR

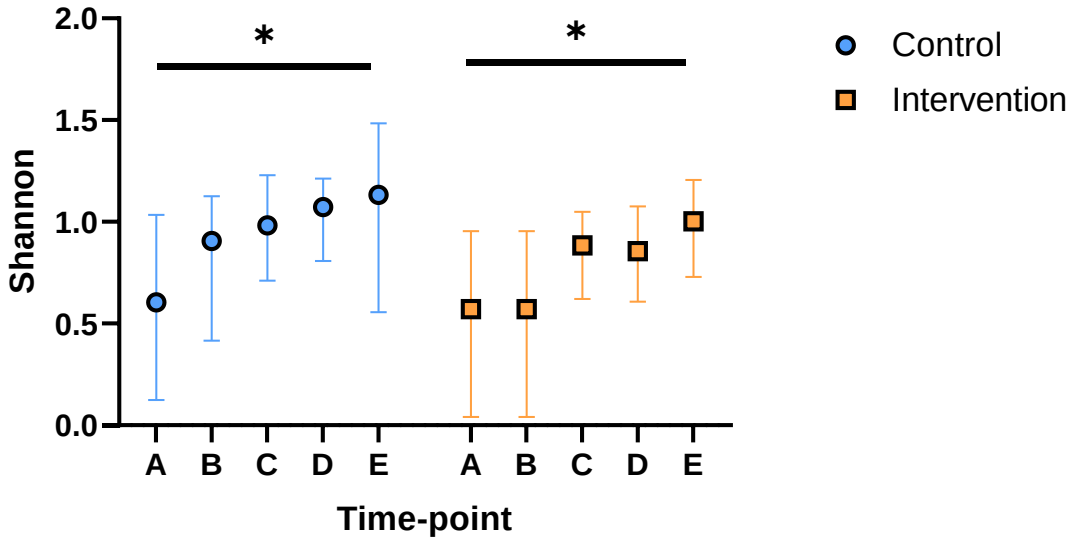


Figure 22 - Comparison of Shannon alpha diversity over time-points, summary data representing all samples stratified by study group and time-point. FDR adjusted p-values comparing time-point A and time-point E, using a Mann-Whitney U test are

shown above the plots (* p-value <0.05). The box represents the median value whilst the whiskers represent the 25th and 75th centiles.

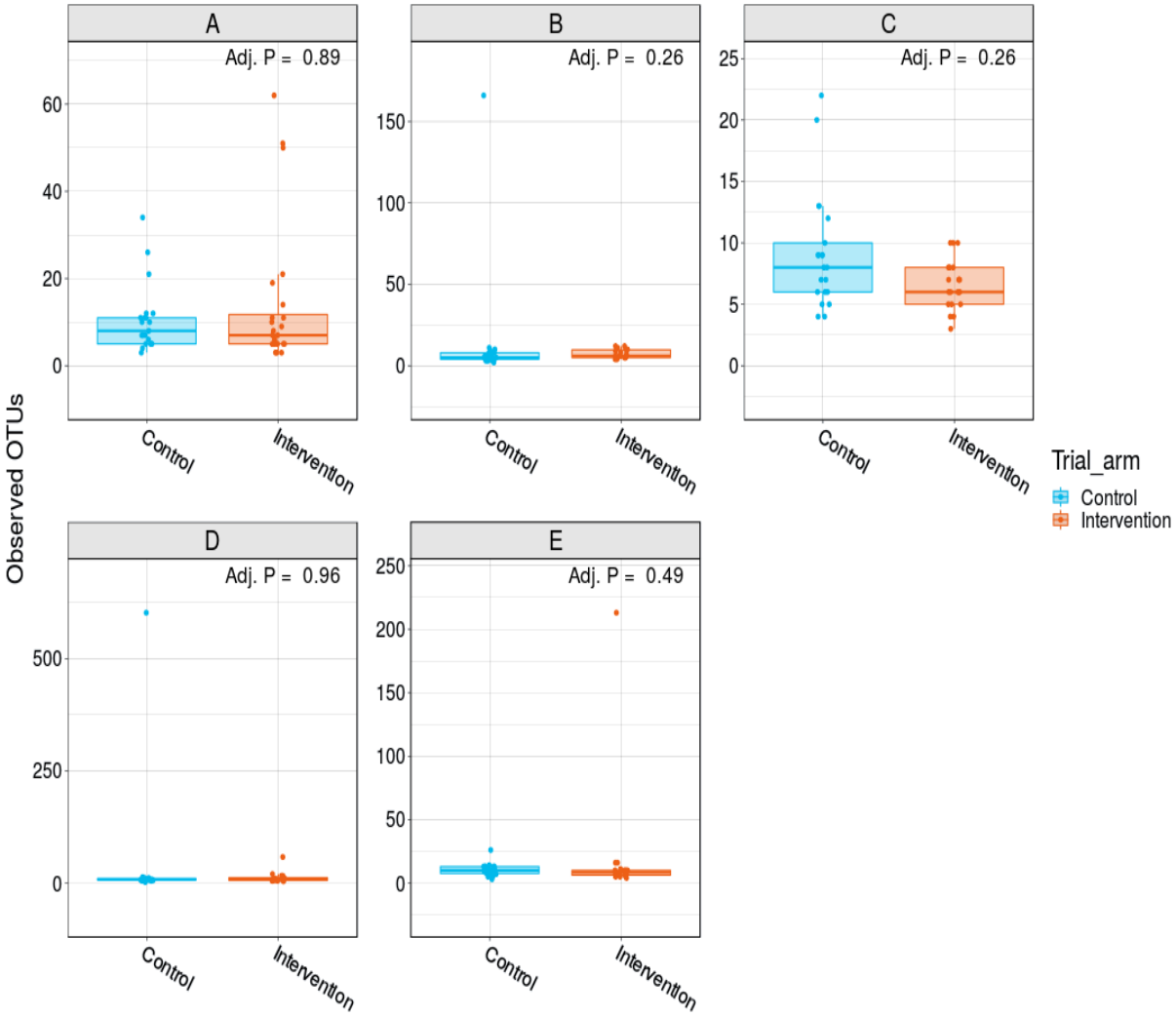


Figure 23(A-E) - Comparison of number of OTU's (richness) in all samples stratified by study groups and time-point (A-E). FDR adjusted p-values are displayed using a Mann-Whitney U test. The y-axis is scaled for each plot. The line represents the median value whilst the box represents the 25th and 75th centiles. The whiskers are a maximum of 1.5 x IQR

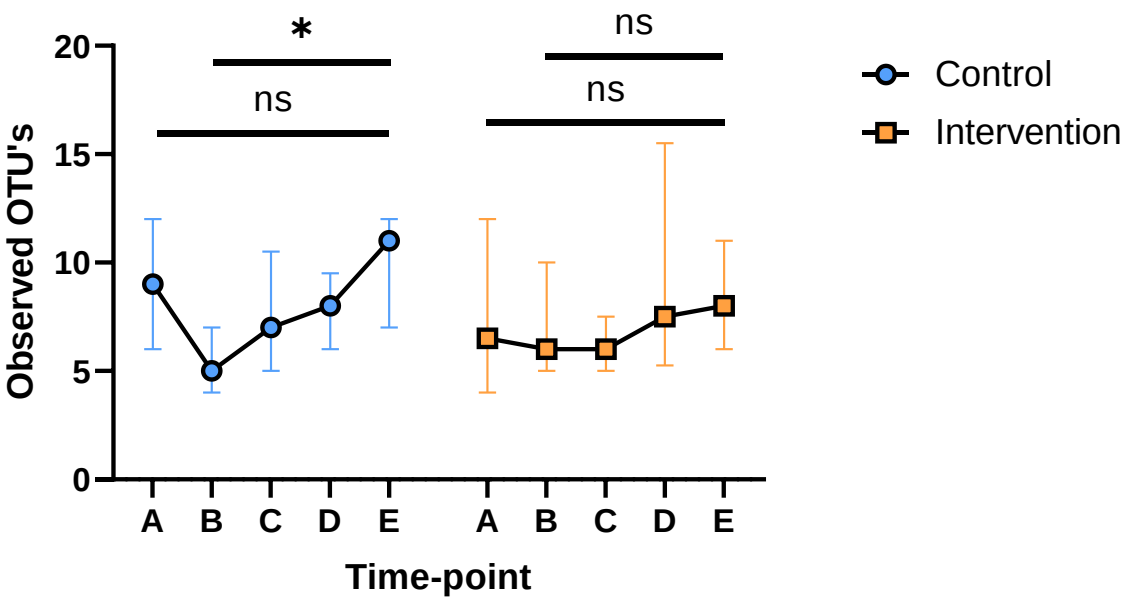


Figure 24 - Comparison of observed OTU's over time, summary data representing all samples stratified by study group and time-point. FDR adjusted p-values comparing time-point A and time-point E, using a Mann-Whitney U test are shown above the plots (* p-value <0.05). The box represents the median value whilst the whiskers represent the 25th and 75th centiles.

I found a significant increase in alpha-diversity (Shannon diversity) between time-point A and E (Figure 22). However, there is no significant difference in alpha-diversity between study groups either in terms of Shannon diversity or richness (number of OTU's) by Mann-Whitney U test, before or after correcting for false discovery rate, as shown in Figures 21 and 23. FDR-corrected p-values are displayed on the plots in Figures 21 and 23.

4.2.2 *Significant difference in unweighted beta-diversity at study end*

A comparison in alpha-diversity can fail to identify true diversity differences between groups of samples. Beta-diversity takes into account alpha-diversity as well as the type of bacteria present. Weighted beta-diversity accounts for the abundance of bacteria whilst unweighted merely accounts for the presence or absense of specific bacteria. Beta-diversity can determine if groups are similar or not but not determine if diversity is increased or decreased.

Figure 25 (A-E) and Figure 26 (A-E) consist of principal component analyses displaying the weighted and unweighted beta-diversity (respectively) stratified by time-point (A-E). Samples are coloured by study group. Distance metric has been calculated using UniFrac.

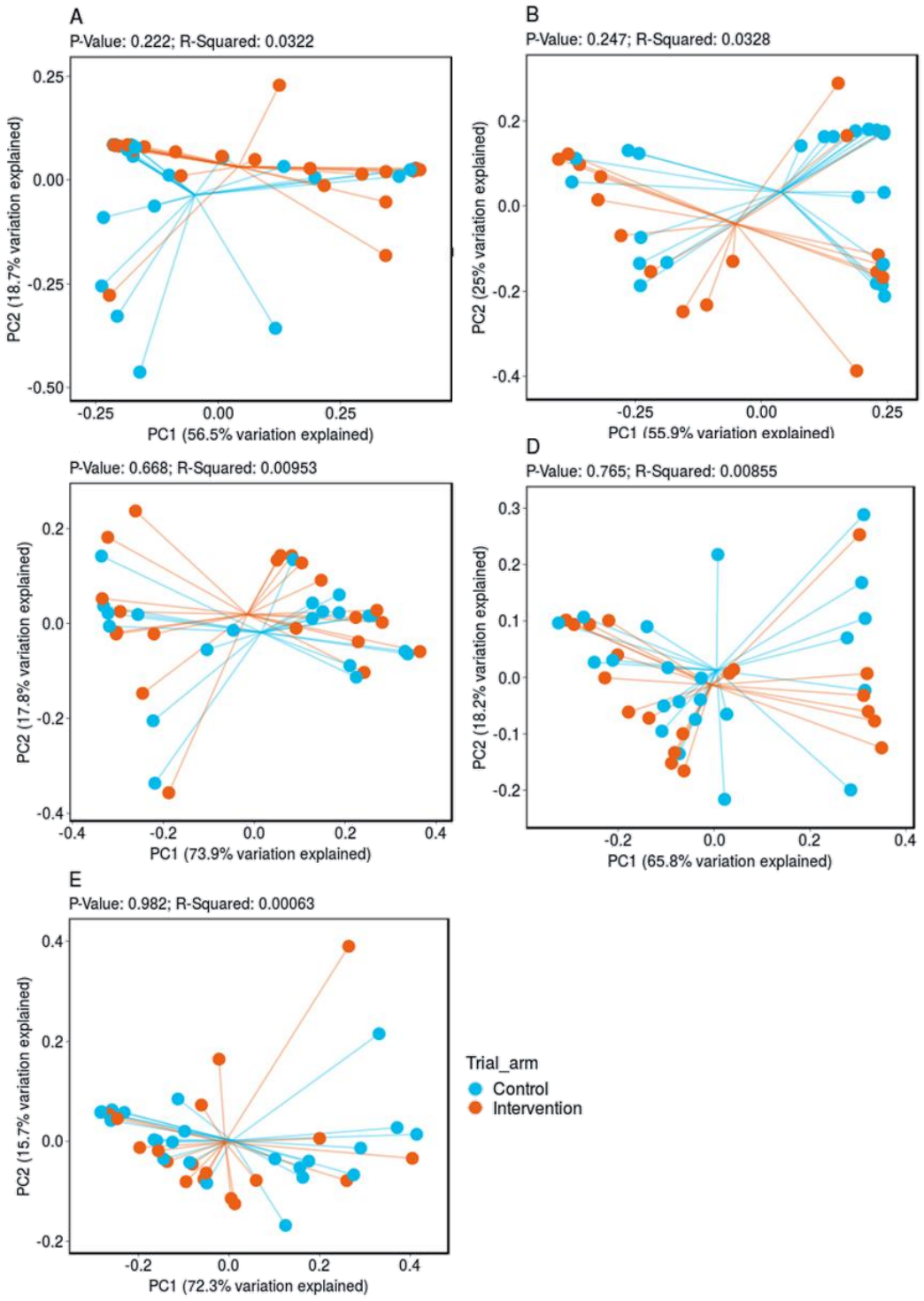


Figure 25(A-E) - Principal Component analyses comparing the beta diversity between study groups stratified by time-point A-E. Distance metrics have been calculated using weighted Unifrac, p-values and R-squared values are displayed on the plot.

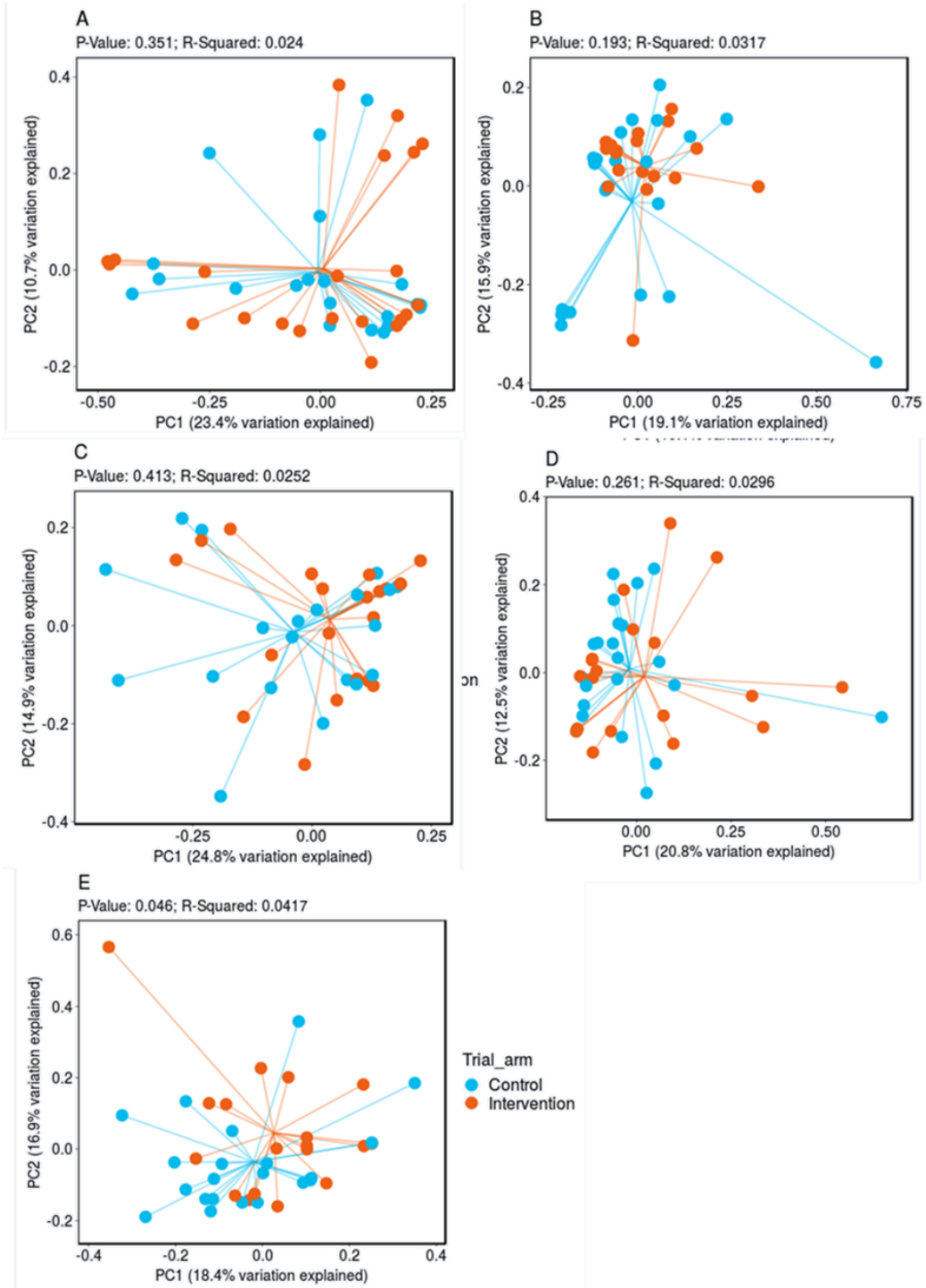


Figure 26(A-E) - Principal Component analyses comparing the beta diversity between study groups stratified by time-point. Distance metrics have been calculated using unweighted Unifrac, p-values and R-squared values are displayed on the plot.

No significant differences between study groups in beta-diversity using weighted Unifrac distance metrics were found at any time-point. Using unweighted Unifrac distance metrics a significant difference was found between study groups at time-point

E. Unweighted (as compared to weighted) beta-diversity does not take account of OTU abundance, merely the presence or absence of a particular OTU.

This finding suggests that there is a difference in the diversity of bacteria between the two study groups, which likely relates to the less abundant bacteria. We can therefore consider the populations dissimilar in this instance, however it is not possible to attribute higher or lower beta-diversity. Given that Figures 21 and 23 demonstrated a non-significant increase in the alpha-diversity (richness and Shannon diversity) in the control group, this suggests that the difference in beta-diversity may be due to an increased diversity of microbials in the control group.

4.2.3 *Significant difference in lactobacillus abundance between study groups*

Figure 27 and Figure 28 display the abundance of OTU's at the family and genus level respectively stratified according to time-point A to E.

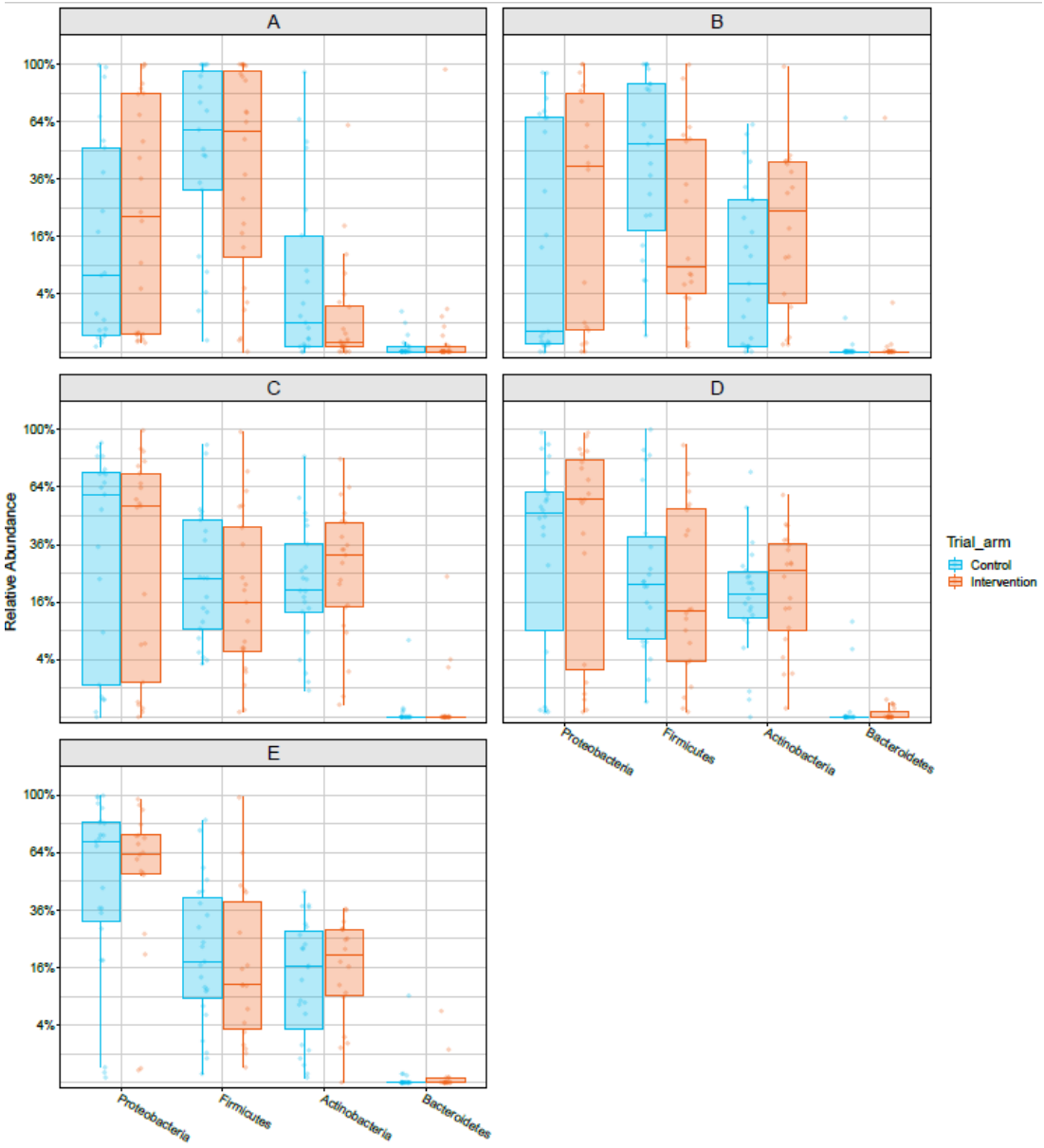


Figure 27 (A-E) - Box plots displaying the relative abundance of the 4 most abundant OTU's (phylum level) for all samples stratified by study group and time-point (A-E). The y-axis has been square root transformed. A Mann-Whitney test could not detect any significant difference between study groups at any time-point.

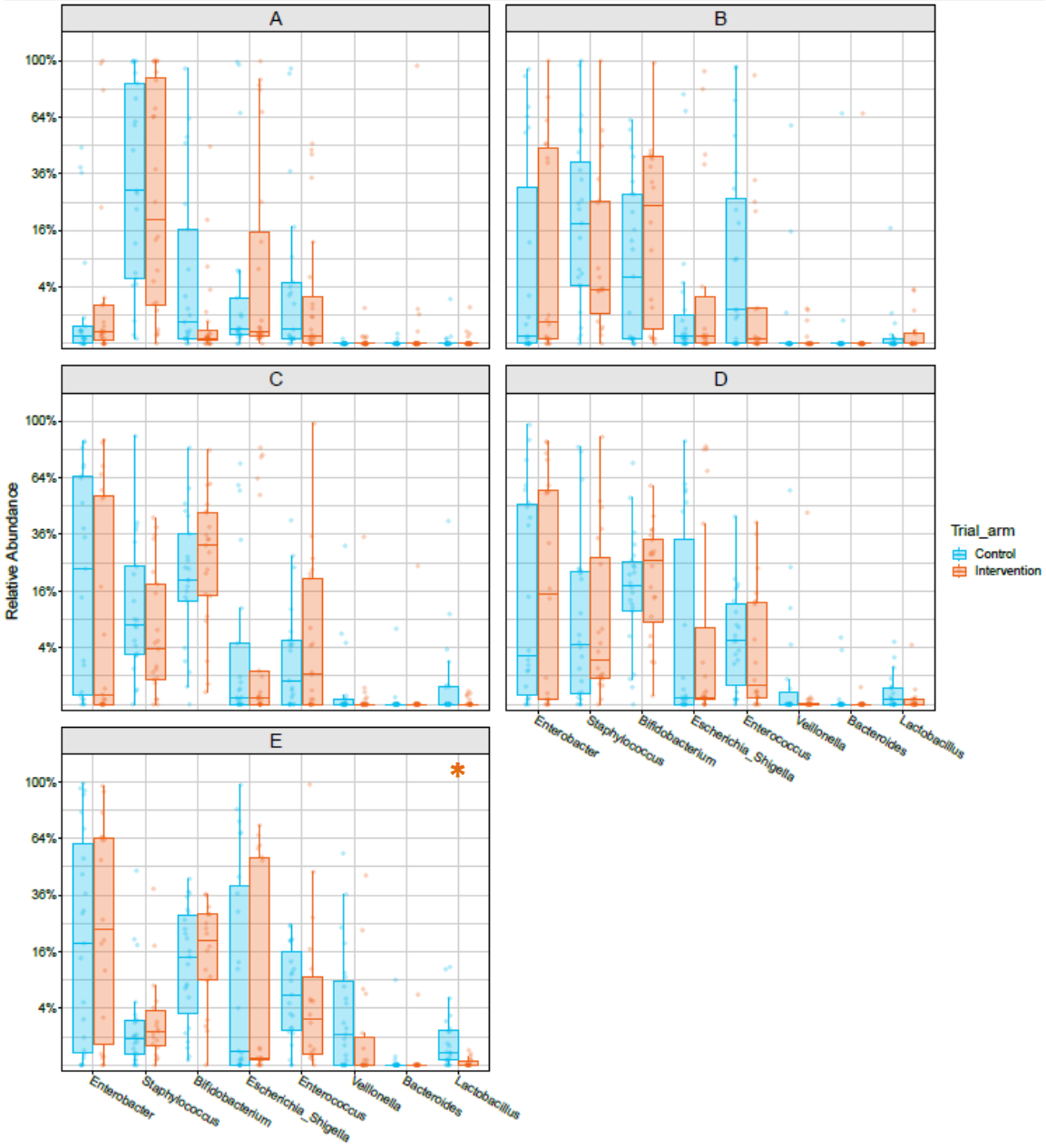


Figure 28(A-E) - Box plots displaying the relative abundance of the 8 most abundant OTU's (genus level) for all samples stratified by study group and time-point A-E. The y-axis has been square root transformed. Comparison between study groups was made using a Mann-Whitney test. Statistically significant (FDR adjusted p <0.05) differences between study groups are highlighted with an asterisk on the plot.

The only significant difference found in OTU abundance between study groups was of *Lactobacillus* at genus level at time-point E (FDR adjusted p value of 0.0119). This is an interesting finding as an increased abundance of Lactobacilli is thought to be beneficial although it has been associated with both CMF and MOM feeding. Probiotics commonly contain *Lactobacillus*, including those used in the RVI NICU. Whilst I demonstrated that

exposure to probiotics was not different between the two groups in the duration of the study (Table 18), I hypothesised that this difference could be due to infants being exposed to probiotics at the time of sampling. Figure 28 (C-E) was repeated limited to samples when infants were exposed to probiotics (Figure 29). This did not reveal a significant difference in *Lactobacillus* between study groups depending on probiotic exposure. Furthermore, when I compared samples from time-point E depending on probiotic exposure irrespective of study arm, this did not reveal any significant difference in the abundance of *Lactobacillus* (Figure 30). Together this suggests that the difference that has been observed in Figure 28 is not due to probiotic exposure.

I next explored whether the increase in *Lactobacillus* was related to a greater exposure to MOM in the control group as highlighted in Figure 17. The amount of MOM received on the day of sample and the preceding two days was divided by the total enteral intake during this time, and each sample was placed in one of three MOM categories (A (0-29%), B (30-69%), C (>=70%)). There was a significant difference in *Lactobacillus* abundance between study groups only when the analysis was limited to samples in category A (0-29% MOM) suggesting that CMF feeding promotes Lactobacilli more than fortified DHM (RTF 26) (Figure 31). Gregory et al. reported increased Lactobacillales at a class bacterial level in CMF-fed compared to MOM-fed infants however conversely Cong et al reported increased *Lactobacillus* at genus level in MOM-fed infants compared to non-MOM fed infants (Cong et al., 2016, Gregory et al., 2016).

In order to ensure that the increased *Lactobacillus* was a true finding, I therefore compared the MOM categories A (0-29% MOM) and C (>=70% MOM) but limited to the Control group (Figure 32). There was no significant difference in *Lactobacillus* (FDR p value 0.25) with a higher mean in category A. Lastly, MOM categories were compared irrespective of study group (Figure 33), revealing no significant difference in *Lactobacillus* abundance between MOM categories.

Taken together these results suggest that CMF-fed infants have an increased abundance of *Lactobacillus* compared to infants fed DHM, and a similar abundance of Lactobacillus compared to infants fed MOM fortified with BMF.

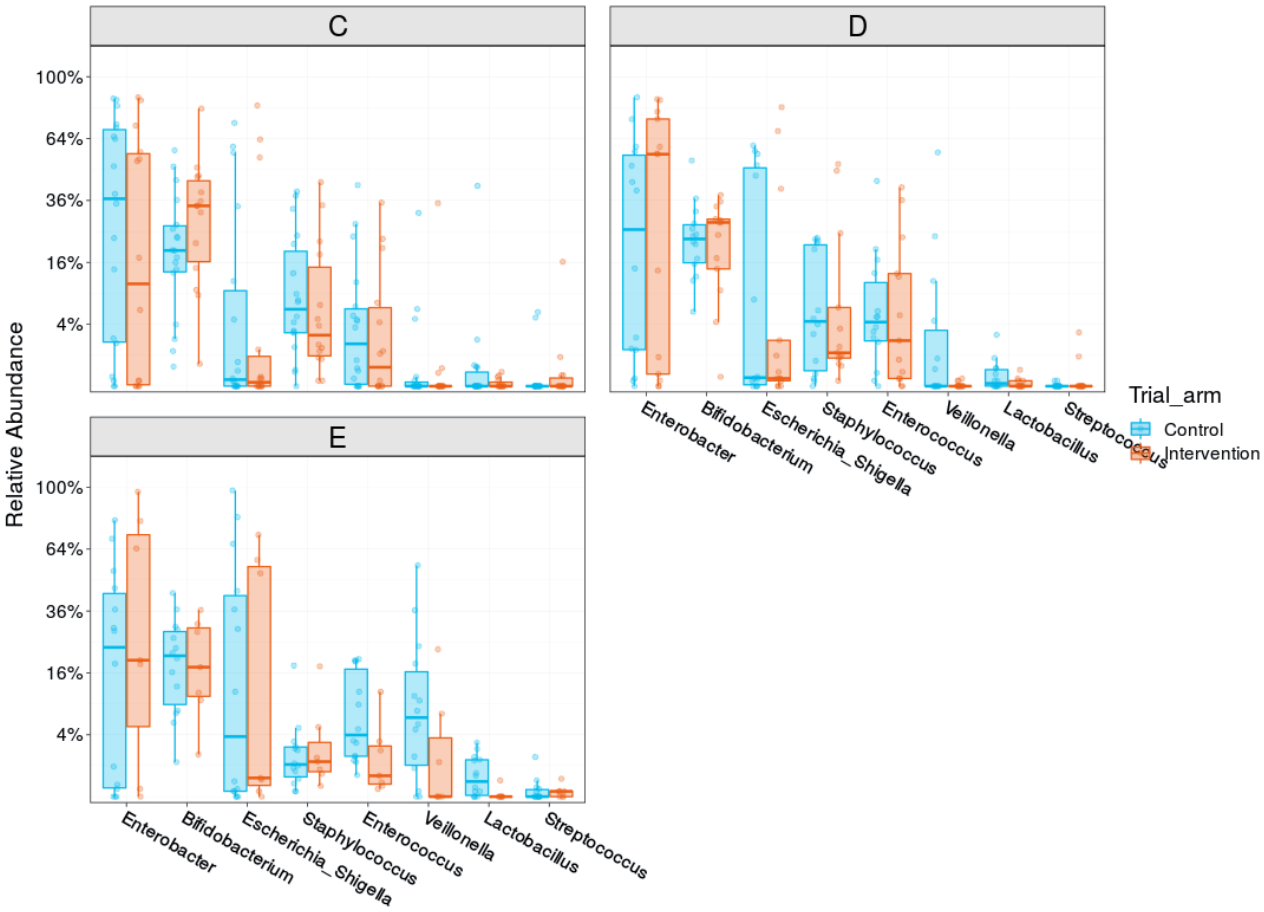


Figure 29 - Box plots displaying the relative abundance of the 8 most abundant OTU's (genus level) for only samples taken when infants were exposed to probiotics stratified by study group and limited to time-point C-E.

The y-axis has been square root transformed. Comparison between study groups was made using a Mann-Whitney test.

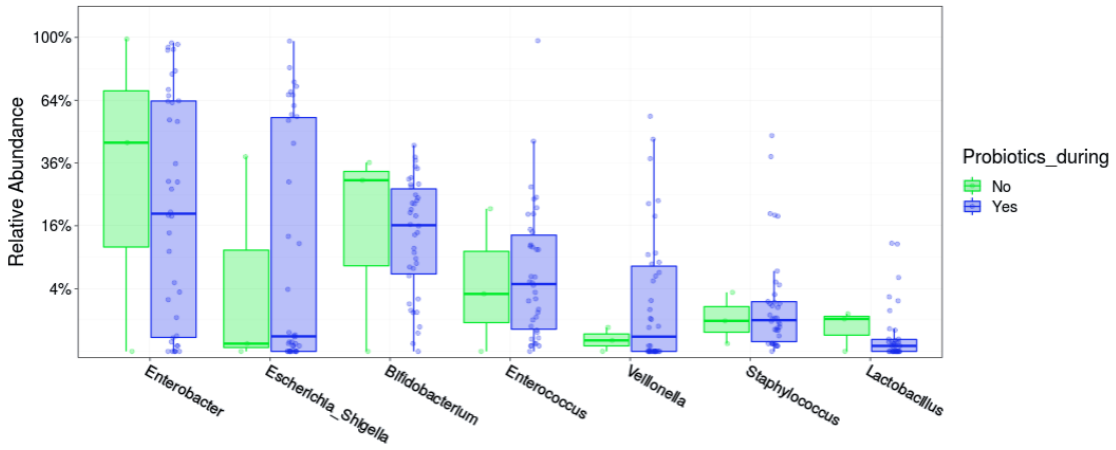


Figure 30 - Box plots displaying the relative abundance of the 8 most abundant OTU's (genus level) for samples stratified by probiotic exposure and limited to time-point E.

The y-axis has been square root transformed. Comparison between study groups was made using a Mann-Whitney U test.

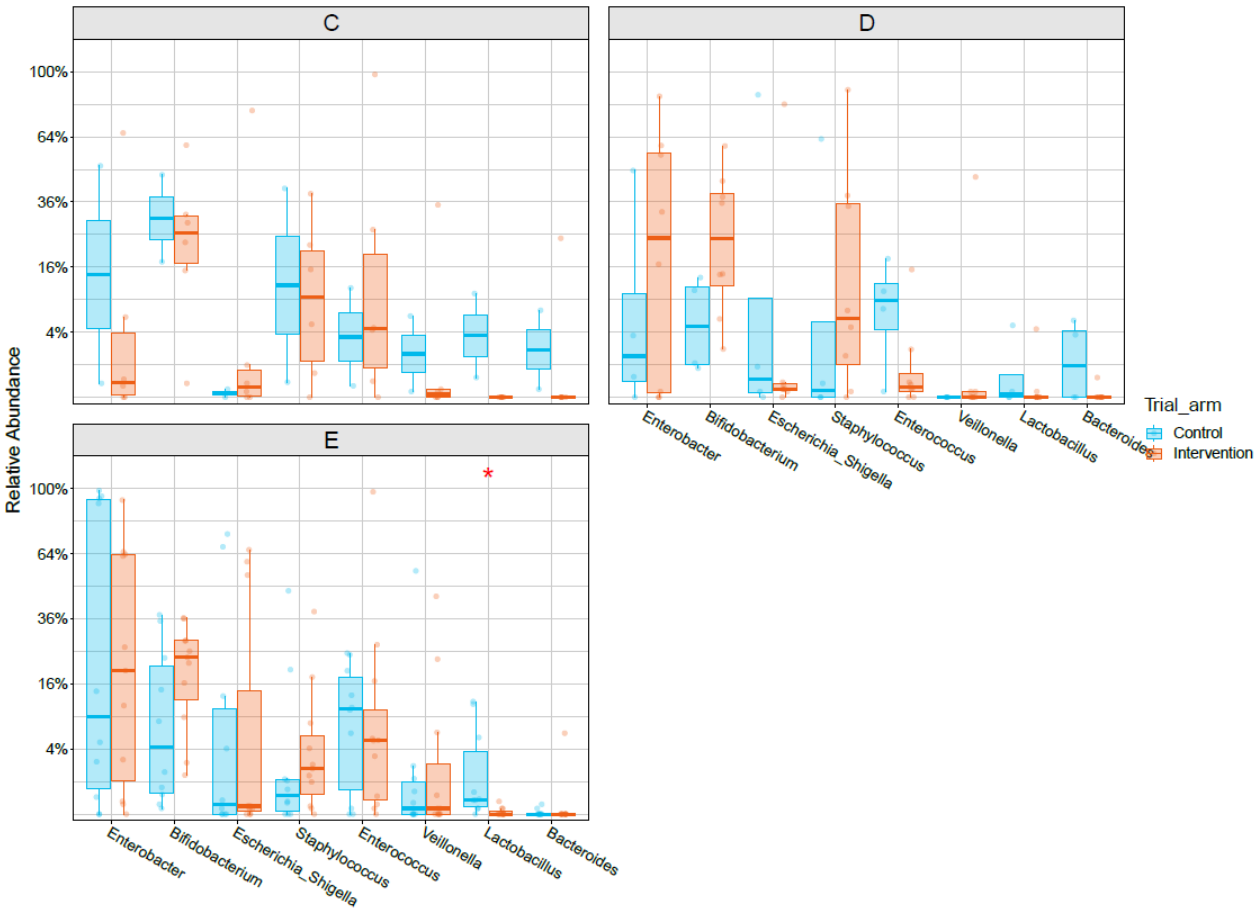


Figure 31 - Comparison of OTU abundance (genus level) between study groups limited to samples when infants fed less than 30% MOM. The y-axis has been square root transformed. Comparison between study groups was made using a Mann-Whitney test. Statistically significant (FDR adjusted p <0.05) differences between study groups are highlighted with an asterisk on the plot.

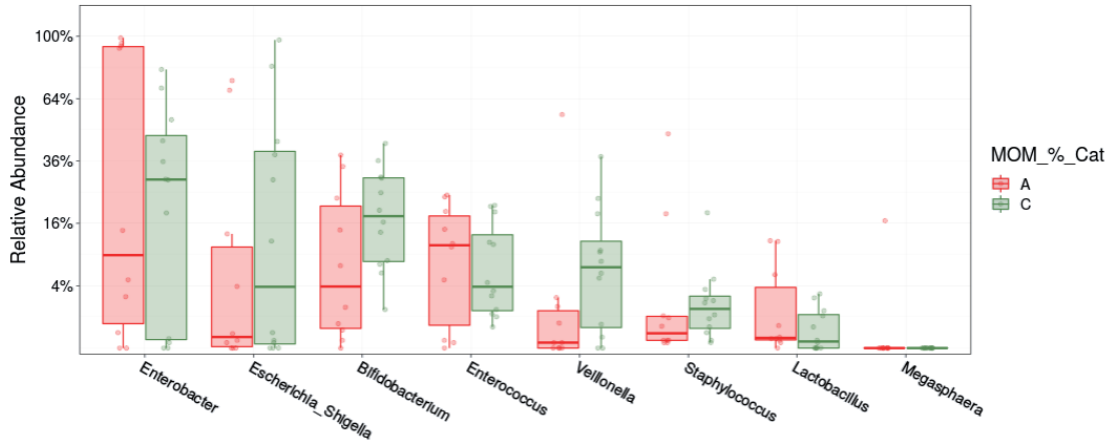


Figure 32 - Comparison of OTU abundance (genus level) between MOM categories A and C limited to control group at time-point E. The y-axis has been square root transformed. Comparison between study groups was made using a Mann-Whitney U test.

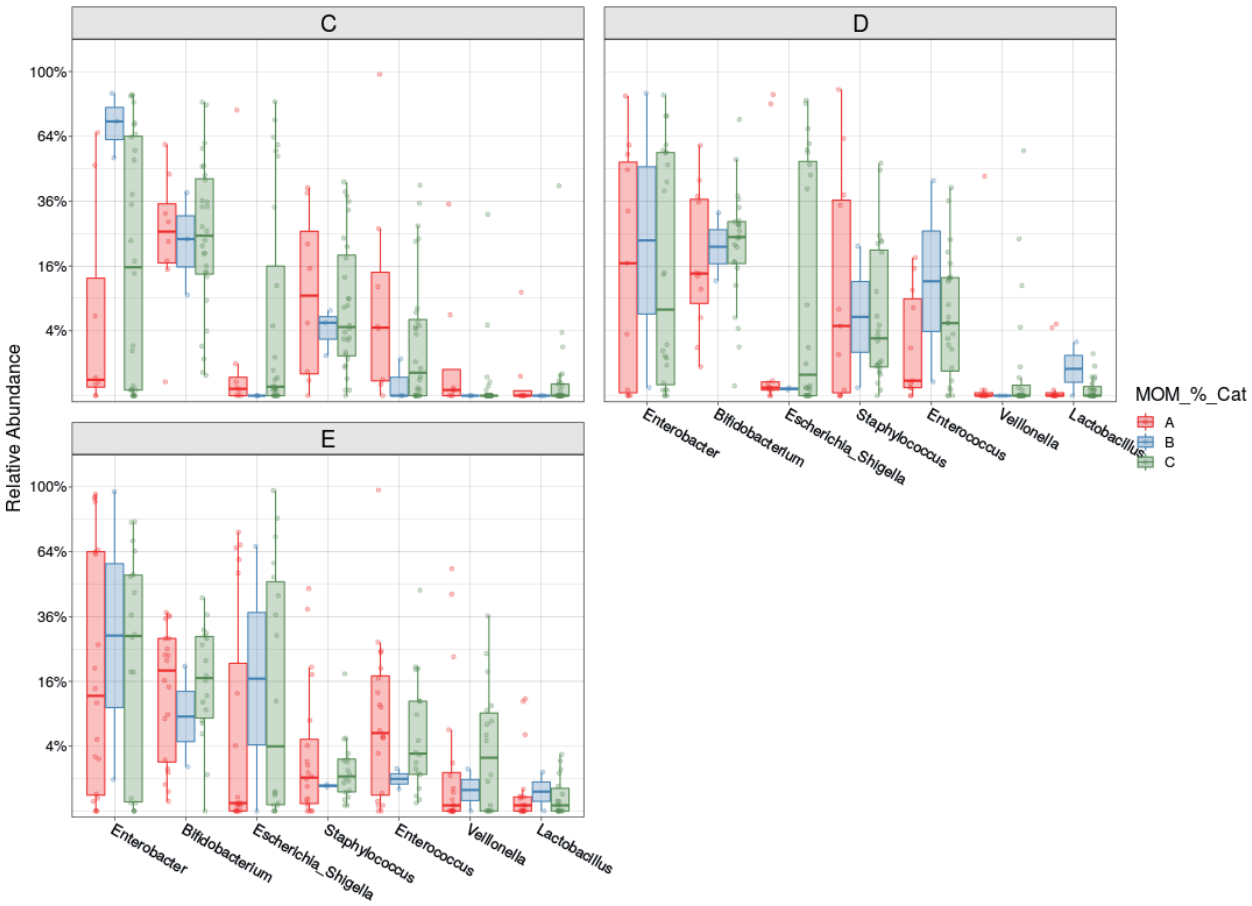


Figure 33 – Box plot comparing OTU abundance (genus level) when samples are stratified by MOM Category (A - 0-29%, B 30-69, C >=70%) limited to when infants were fully enterally fed. MOM Category calculated by dividing total amount of MOM received on day of sample and preceding 2 days by total enteral intake in this time. OTU's with an abundance >1% are shown. The y-axis has been square root transformed. Kruskal-Wallis test did not detect any significant difference between the sample means.

A striking observation when comparing the abundance of OTU's is that *Bifidobacterium* was the 3rd most abundant OTU at the genus level, representing approximately 20% abundance by time-point B (DOL 10). This is in stark contrast to a study undertaken in the same NICU which found the mean abundance of *Bifidobacterium* only reached 20% after DOL 75 (Stewart et al., 2017a). The difference likely relates to exposure to probiotics as they were introduced to the NICU practice in 2013, and reflects that some samples in the study by Stewart et al. were taken before 2013.

4.2.4 *Significant difference in beta diversity depending on type of fortifier*

The dietary intervention was variable in the study. Some infants received only MOM with fortifier added once they were fully enterally fed. I sought to determine if the type of fortifier used had any bearing on the gut microbial composition when infants received mainly MOM. I compared only samples in time-points C-E from infants who had been exposed to fortifier, excluding infants receiving either CMF or RTF-26. This analysis did not reveal any significant difference in alpha-diversity, weighted beta-diversity or OTU abundance between study groups. A plot comparing unweighted Unifrac beta-diversity between the two study groups is shown in Figure 34 which revealed a significant difference. These data suggest that using different fortifiers leads to variance in gut microbials, however does not promote the abundant growth of any individual genera consistent with what was demonstrated in Figure 29. However, this finding may still be important for health.

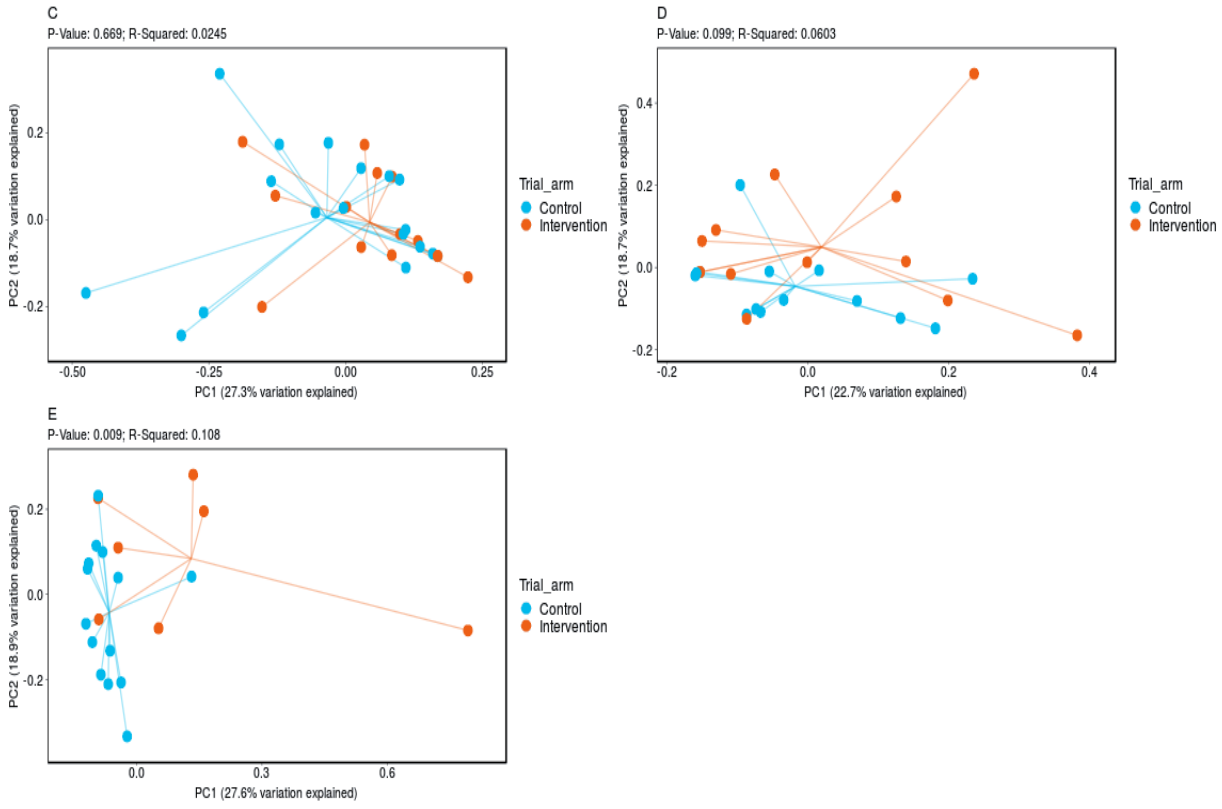


Figure 34 - Comparison of unweighted Unifrac beta-diversity stratified by study group and time-point C to E. Samples are limited to those taken when the infant was receiving fortifier. FDR adjusted p-values and R^2 values are displayed above the plots.

4.2.5 *Microbiome changes over time*

Figure 35 displays stacked bar plots of the OTU abundance over time with plots stratified by study group. Each sample is represented by a stacked bar.

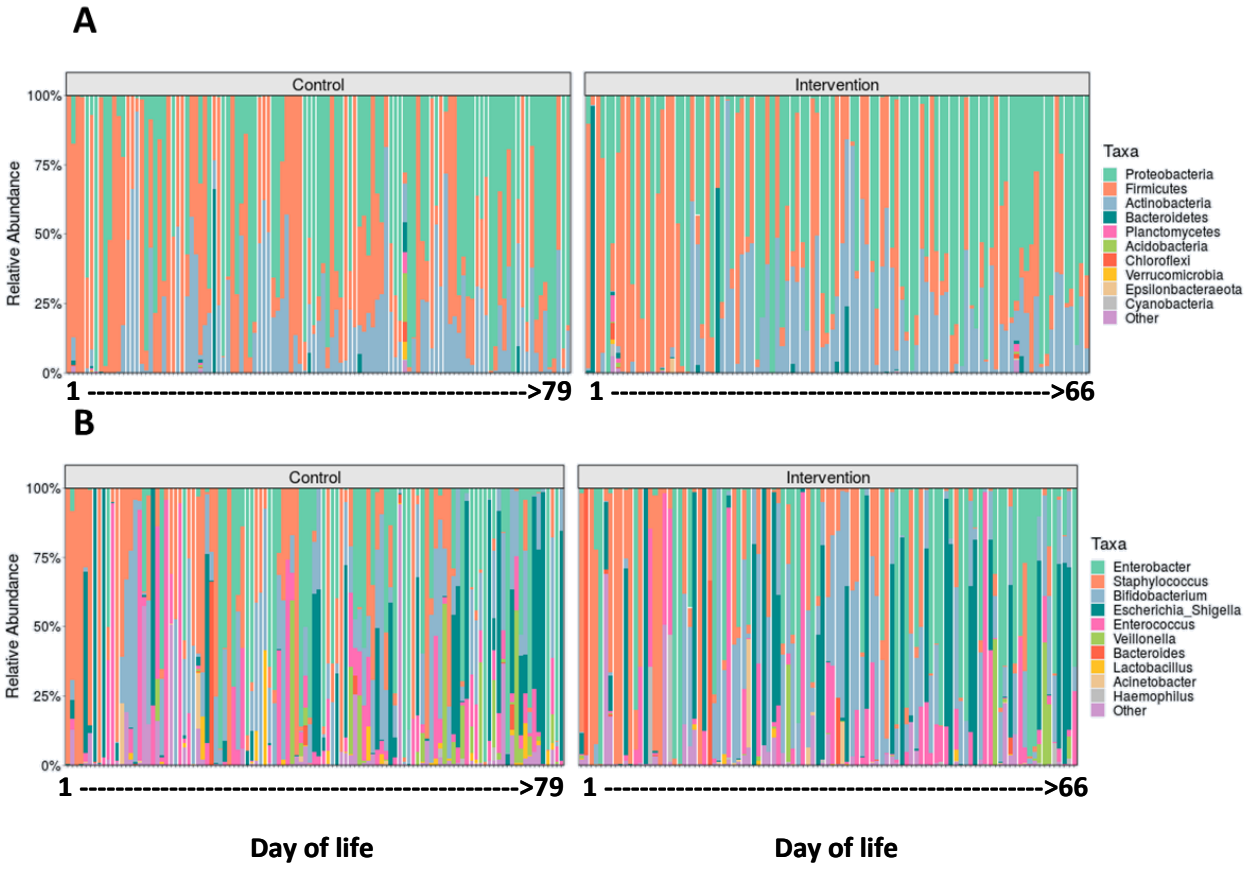


Figure 35 (A-B) - Stacked bar graph displaying the relative abundance of the different taxa over time. All samples are displayed, sorted by DOL on the x-axis with each individual sample represented by a stacked bar. A – OTU are displayed at phylum level. B – OTU are displayed at genus level.

Figure 35 visually demonstrates that the early life microbiota is dominated by Proteobacteria (phylum), Firmicutes (phylum), *Staphylococcus* (genus) and *Escherichia/Shigella* (genus) with an increase in the number of different OTU's. In the later samples, there is an increase of Proteobacteria (phylum), Actinobacteria (phylum) corresponding to *Enterobacter* (genus) and *Bifidobacterium* (genus), whilst Firmicutes (phylum) and correspondingly *Staphylococcus* (genus) decrease.

These visual findings are confirmed using linear regression analysis in Figure 36 and Figure 37. There is a significant increase in Proteobacteria (phylum) and correspondingly *Enterobacter* (genus) over time. There is a significant decrease in Firmicutes (phylum) and correspondingly *Staphylococcus* (genus) over time. These findings are present in both study groups. There is an

increase of Actinobacteria (phylum) and *Bifidobacteria* (genus) however neither is statistically significant.

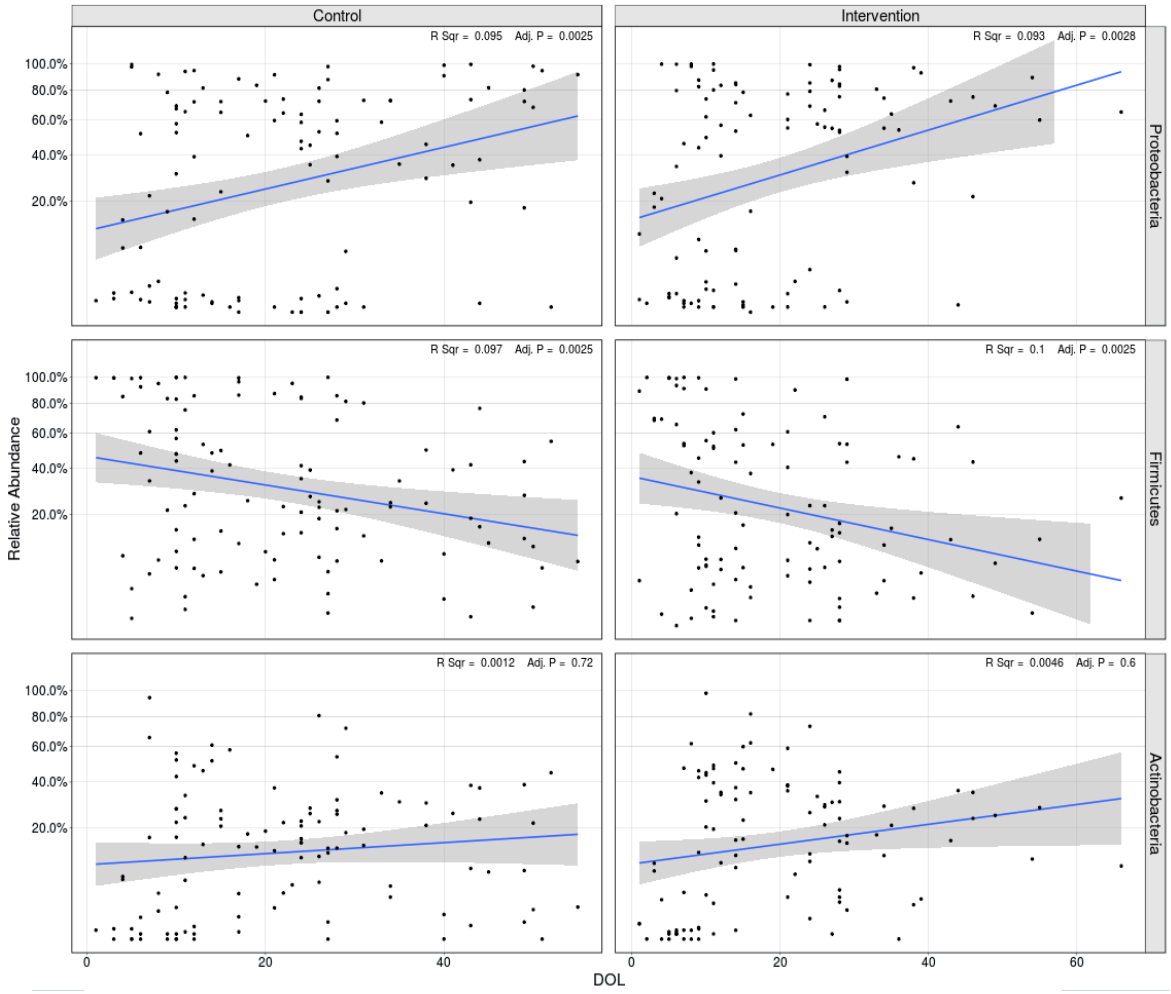


Figure 36 - Linear regression analysis displaying relative abundance of the 3 most abundant OTU over time (labels on right y axis; phylum level), stratified by study group. Line of best fit (blue) and 95% confidence intervals (grey) are displayed. P-value and R-squared values are given based on the change in abundance of the OTU over time.

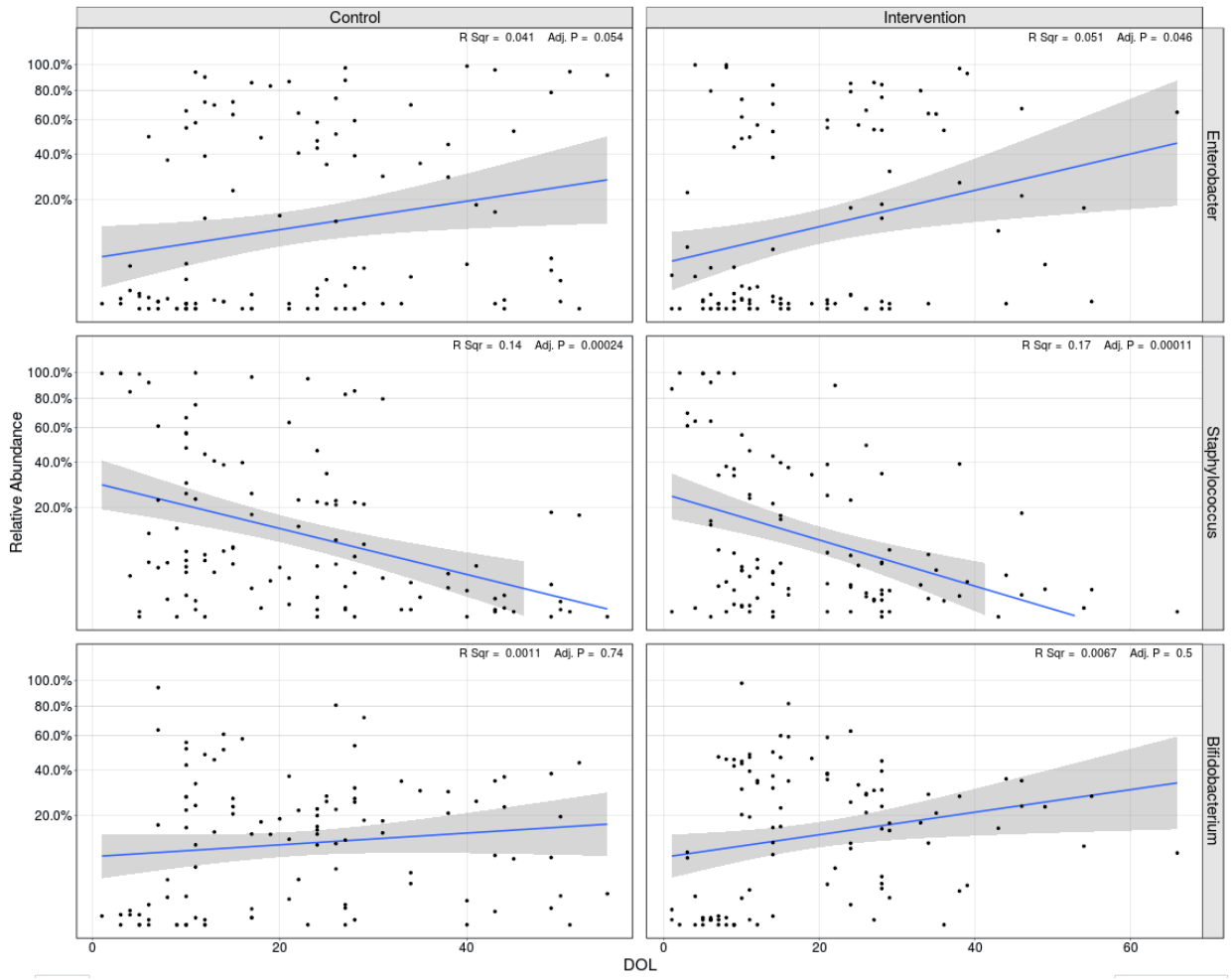


Figure 37 - Linear regression analysis displaying relative abundance of 3 most abundant OTU's over time, (labels on right y axis; genus level), stratified by study group. Line of best fit (blue) and 95% confidence intervals (grey) are displayed. P-value and R-squared values are given for the change in abundance over time

4.3 T cells in preterm infants

In total, 12,387,800 live single cells from 66 samples (23 adult and 43 infant) were analysed using mass cytometry. This included 4,875,847 T cells, of which 3,126,571 T cells were from the infant samples. Red blood cell lysis was performed on 15 infant samples, whilst density gradient centrifugation was performed on 28 infant samples, to isolate leucocytes and lymphocytes respectively. These cells were exposed to a 39-marker antibody panel designed to differentiate T cells in order to identify T_{reg}s, MAIT, iNKT and T_h populations. The combined analysis of these samples is presented below.

A further 590,000 lymphocytes from 59 samples were analysed using flow cytometry. These lymphocytes were exposed to a 6-colour panel designed to differentiate B, natural killer (NK) and T cell (CD4+ and CD8+) populations.

4.3.1 *Comparison of lymphocyte populations between study groups*

First, I sought to identify if there was a difference in the relative abundance of lymphocytes between study groups.

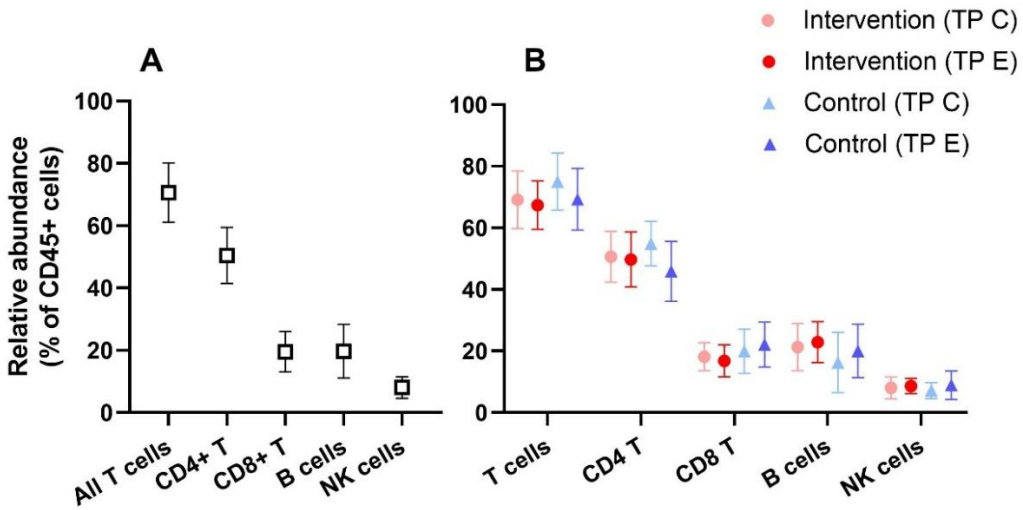


Figure 38 - Comparison of lymphocyte populations using flow cytometry. (A) Box plot displaying relative abundance of each lymphocyte population in 59 samples, box

represents mean and whiskers represent SD. (B) Box and whisker plot showing comparison between study groups and time-points. TP – time-point

Lymphocyte populations were compared between study groups across the time-points. No significant difference was found using unpaired t-tests comparing each cell type between study groups and time-points (Figure 38).

4.3.2 *Mass cytometry reveals comparable T cell proportions to conventional flow cytometry*

As mass cytometry remains a relatively new technology, I planned a comparison between flow and mass cytometry data to ensure I was obtaining a representative sample of cells, using data from blood samples taken at the same time. The gating strategies used for comparison are displayed in Figures 12 and 13 in Section 3.6, namely to identify live cells then single CD45+ cells. We found that flow cytometry identified significantly more lymphocytes as T, B and NK cells. The proportion of T cells identified as CD4 or CD8 positive was not different.

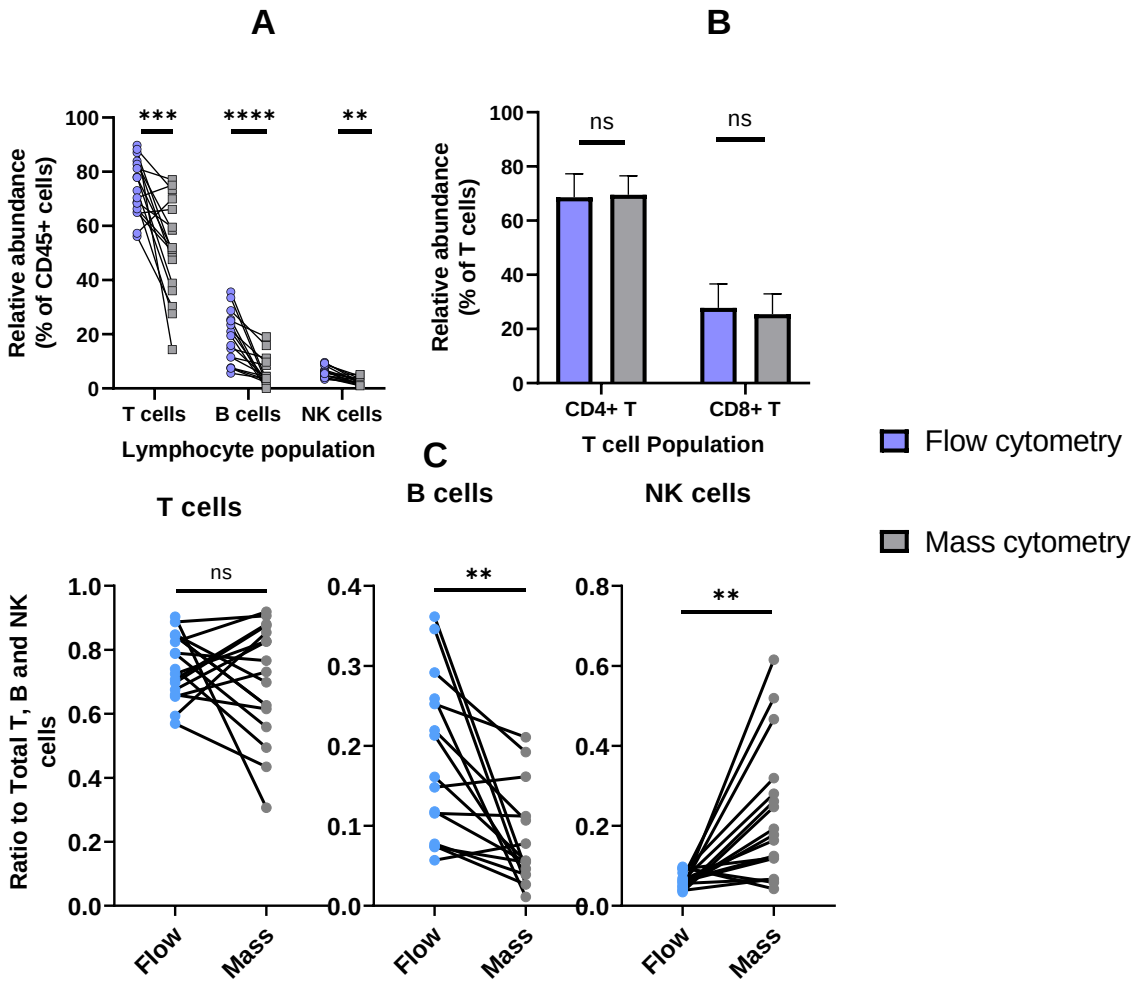


Figure 39 – Comparison of flow and mass cytometry (A) Relative abundance of lymphocyte populations using paired samples. Lines between paired samples are displayed. (B) Box plot displaying relative abundance of CD4 and CD8 T cells. (C) Ratios of lymphocytes are compared. Ratios calculated by dividing percentage of first population by total T, B and NK population. Lines between paired samples are displayed

Asterisks denoting p values using a paired sample Student t-test are displayed above the plots (* <0.05, ** <0.01, *<0.001, **** <0.0001).**

These results demonstrate variation in identified populations using flow and mass cytometry. This could relate to variation in the gating or the cell staining. Figure 39A demonstrates that there was a significantly lower abundance of T, B and NK cells identified using mass cytometry. To investigate this further I analysed the population of CD45+ cells that were not identified as T, B, or NK cells using mass cytometry with the results visible in Supplementary figures A and B.

The heatmap in Supplementary Figure A demonstrates that the expression of CD16 is relatively high in this population. Supplementary Figure B demonstrates that for some samples CD14 and CD66b expression appears higher than expected. This would suggest that there are granulocytes and monocytes within the CD45+ population. A further comparison of samples where cell reduction was completed using either red cell lysis or density gradient centrifugation revealed that there were monocytes and granulocytes within the CD45+ population regardless of cell reduction technique, however a greater proportion of granulocytes (7% vs 2% CD66b+ cells respectively). I was not able to compare these results against flow cytometry as the same lineage markers were not present in the flow cytometry analysis. However, back-gating onto a forward and side scatter plot revealed that the lymphocyte populations were in the expected location (Supplementary Figure C). Reassuringly, Figure 39 demonstrates that the proportions of CD4 and CD8 T cells were not significantly different using either flow or mass cytometry suggesting that staining and gating of these populations within the CD45+CD3+ cells was consistent.

A recent study comparing staining of cells using mass and flow cytometry that handled cells and tissue in the same way found that the two methods were comparable (Gadalla et al., 2019). The study by Gadalla at al. was specifically designed to compare staining and cell identification with mass versus flow cytometry. They therefore used the same clones of antibodies for both techniques, each blood sample was split, the staining technique used was very similar and data acquisition was within a day of staining. As my primary aim was not to compare mass and flow cytometry results, my results possibly reflect that the viability of cells was lower in the mass cytometry group than the flow cytometry, and hence the ratios between cell types were maintained. The median time between acquiring a sample and initial staining was 20 hours for mass cytometry, whilst flow cytometry samples were typically stained and data acquired the same day. Mass cytometry staining required a two-stage process over 3 days as described in the Methods section, before data acquisition. However, this comparison suggests that T cells proportions were comparable between flow and mass cytometry. Mass cytometry offers the advantage of a more in-depth analysis of a cell population using one antibody panel therefore was felt preferable for this analysis.

4.3.3 *Comparison of adult and infant samples demonstrates expected differences in T cell population*

In order to test the sensitivity and reproducibility of the mass cytometry staining method, I first looked for expected differences between infant and adult T cell populations (Figure 40). The heatmaps used for cell population identification used in Figure 40 are shown in

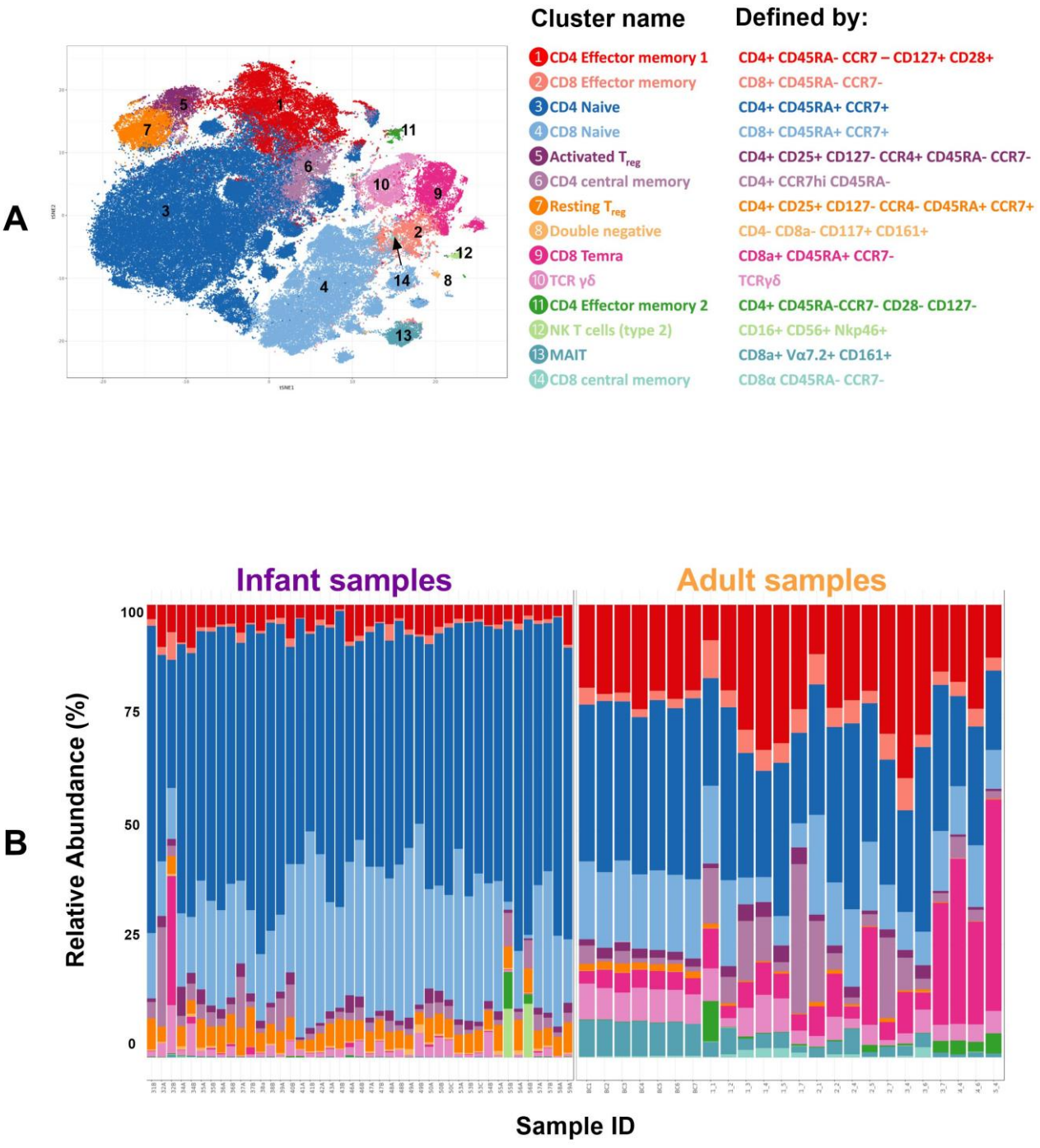
Supplementary Figures D and E. Supplementary Figure D shows the markers used for identification of cell clusters whilst Supplementary Figure E demonstrates how some clusters were merged and the expression of all markers in each cluster. I found a similar median abundance of T_{reg}s (≈ 2%), MAIT (3%) and TCRγδ (5%) cells as % of T cells in adults to those reported (Walker et al., 2014, Jonuleit et al., 2001, Provine et al., 2018). As expected, a clear difference was found comparing infant and adult T cell populations as displayed by the stacked bar plot in Figure 40B and the boxplot in Figure 40C. Namely these differences are:

- a significant decrease in the relative abundance of central, effector and terminal effector memory, and MAIT cells in infants compared to adults
- a significant increase in the relative abundance of naïve CD4 T, DN T and T_{reg} cells in infants.

The differences are known and reflect the relative antigen naïve environment of the preterm infant with an increased population of naïve T cells and decreased population of memory T cells. This is together with a relatively tolerogenic state reflected in the increased population of T_{reg}s. The population of MAIT cells is known to expand over the first years of life.

In comparison to one another, most infant samples appeared similar in terms of the relative abundance of the T cell populations demonstrated in Figure 40. However, three samples are visually different when displayed in the stacked bar plot in Figure 40B. Sample 32B (Infant 32, sample B) appeared clearly different due to an expanded EM CD8 population, and samples 55B and 56B had expanded EM CD4 and NKT cell populations. There was no obvious clinical explanation for this, these infants did not have LOS or NEC and were not unwell at the time of sampling. Retrospective testing of urine for cytomegalovirus (CMV) was however positive for infant 32 at the time of sample 32B (although negative in the first week of life), suggesting that CMV infection likely explains the expanded EM CD8 population. I could not identify any reason for the differences found in samples 55B and 56B, however interestingly infants 55 and 56 were twin siblings. Three infants did develop LOS, two samples were available for two infants (patient 41 and 43) and one for the remaining infant (patient 59). The abundance of T cell populations did not appear different for these infants. This was despite Patient 41 also requiring abdominal surgery for a spontaneous intestinal perforation (SIP).

Figure 40 - Comparison of adult vs infant T cells using mass cytometry



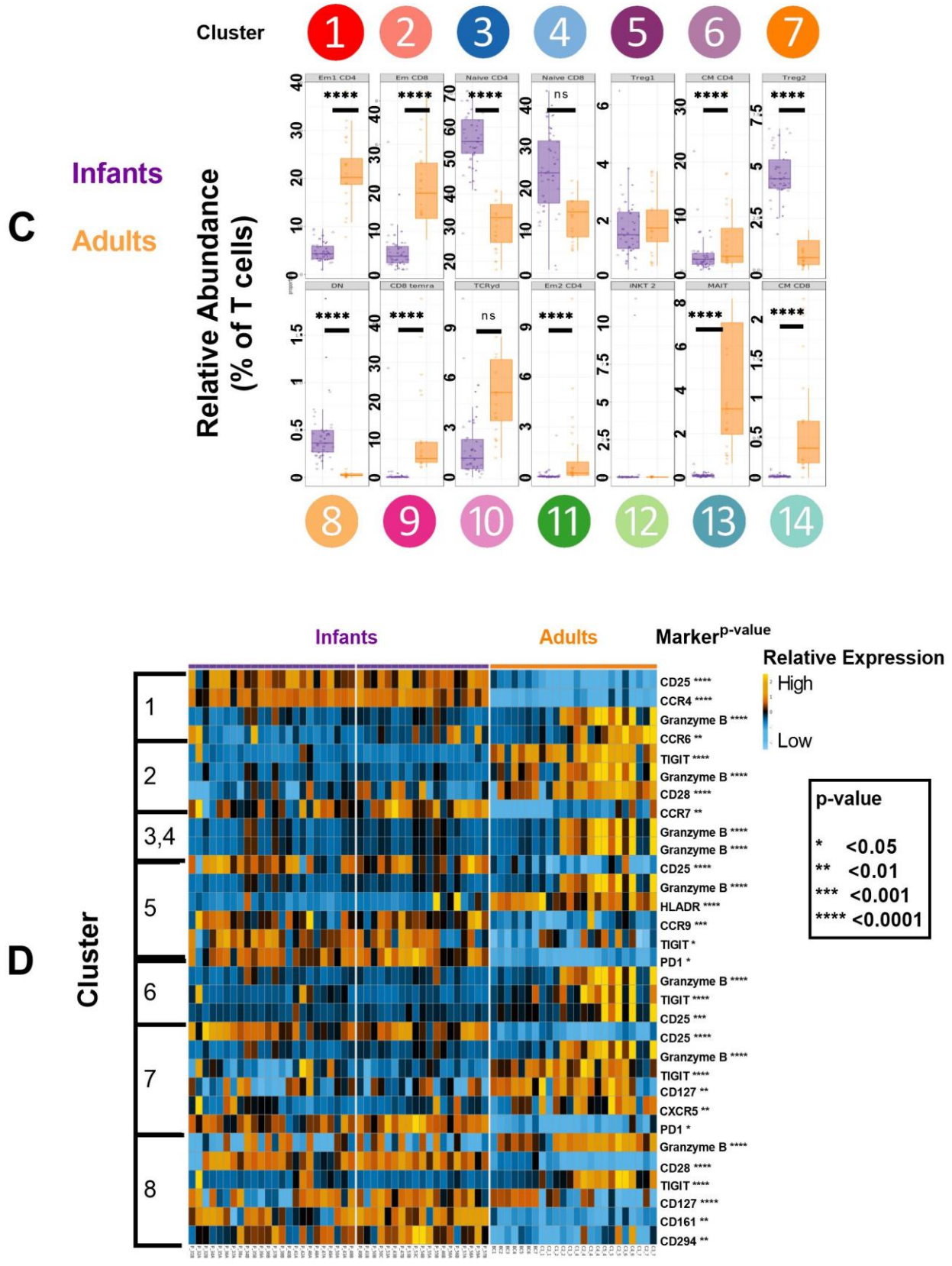

C
Infants
Adults
Cluster
1
2
3
4
5
6
7
Relative Abundance
(% of T cells)
Em1 CD4
Em CD8
Naive CD4
Naive CD8
Treg1
CM CD4
Treg2
DN
CD8 temra
TCRγd
Em2 CD4
iNKT 2
MAIT
CM CD8
****
ns
8
9
10
11
12
13
14
D
Cluster
Infants
Adults
Marker
p-value
Relative Expression
High
Low
1
2
3,4
5
6
7
8
CD25 ****
CCR4 ****
Granzyme B ****
CCR6 **
TIGIT ****
Granzyme B ****
CD28 ****
CCR7 **
Granzyme B ****
Granzyme B ****
CD25 ****
Granzyme B ****
HLADR ****
CCR9 ***
TIGIT *
PD1 *
Granzyme B ****
TIGIT ****
CD25 ***
CD25 ****
Granzyme B ****
TIGIT ****
CD127 **
CXCR5 **
PD1 *
Granzyme B ****
CD28 ****
TIGIT ****
CD127 ****
CD161 **
CD294 **
p-value
* <0.05
** <0.01
*** <0.001
**** <0.0001

Figure 40 - Comparison of adult versus infant T cells using mass cytometry.
(A) T cells were defined as CD45+CD3+lineage- (CD14-CD19-CD66b-) live cells. tSNE plot representing 57 samples, based on the arcsinh transformed expression of 15 markers. 14 clusters have been identified using a heatmap (not shown), and their identity deduced based on expression of key subset markers (right). (B) Stacked bar plot derived from the tSNE plot displaying the relative abundance of the 14 clusters, identified in the key to the right of the plot. Infant samples are displayed on the left, adult samples on the right. (C) Box plot comparing relative abundance of each cluster. Line represents median value, box represents IQR and whiskers (max 1.5 x IQR). Left y axis scales vary to show range of data for each cluster. A generalised linear mixed model (GLMM) was used to compare the relative abundance of each cluster between adults and infants. FDR adjusted p values are displayed above each plot. (D) Heatmap displaying the differential normalised expression of cell markers comparing infant and adult samples, within each cluster (refer to legend in 40(A) for key). Each column represents a sample. A linear mixed model was used to calculate FDR adj p values. The cluster number is given on the left y axis whilst the differentially expressed protein is given on the right y axis. Key to the right of the plot displays meaning of asterisks.

4.3.4 *Comparison of infant T cell subsets between study groups*

The same analysis as shown in Figures 40A and 40C was used to compare infants between study groups and between time-points (C and E). Only 16 preterm infants that provided 2 samples were included in this analysis, with 8 in each study group. The analysis resulted in a t-SNE plot as demonstrated in Figure 41a. The clusters were identified using a heatmap as described in Section 4.3.3. Figure 41(b and c) compare the relative abundance between Control and Intervention groups at both time-points. No significant difference was found between study groups either at time-point C or E, although there was a trend towards increased Resting T_{reg} (cluster 13) and reduced CD8 T_{EMRA} (cluster 8) populations in the intervention group

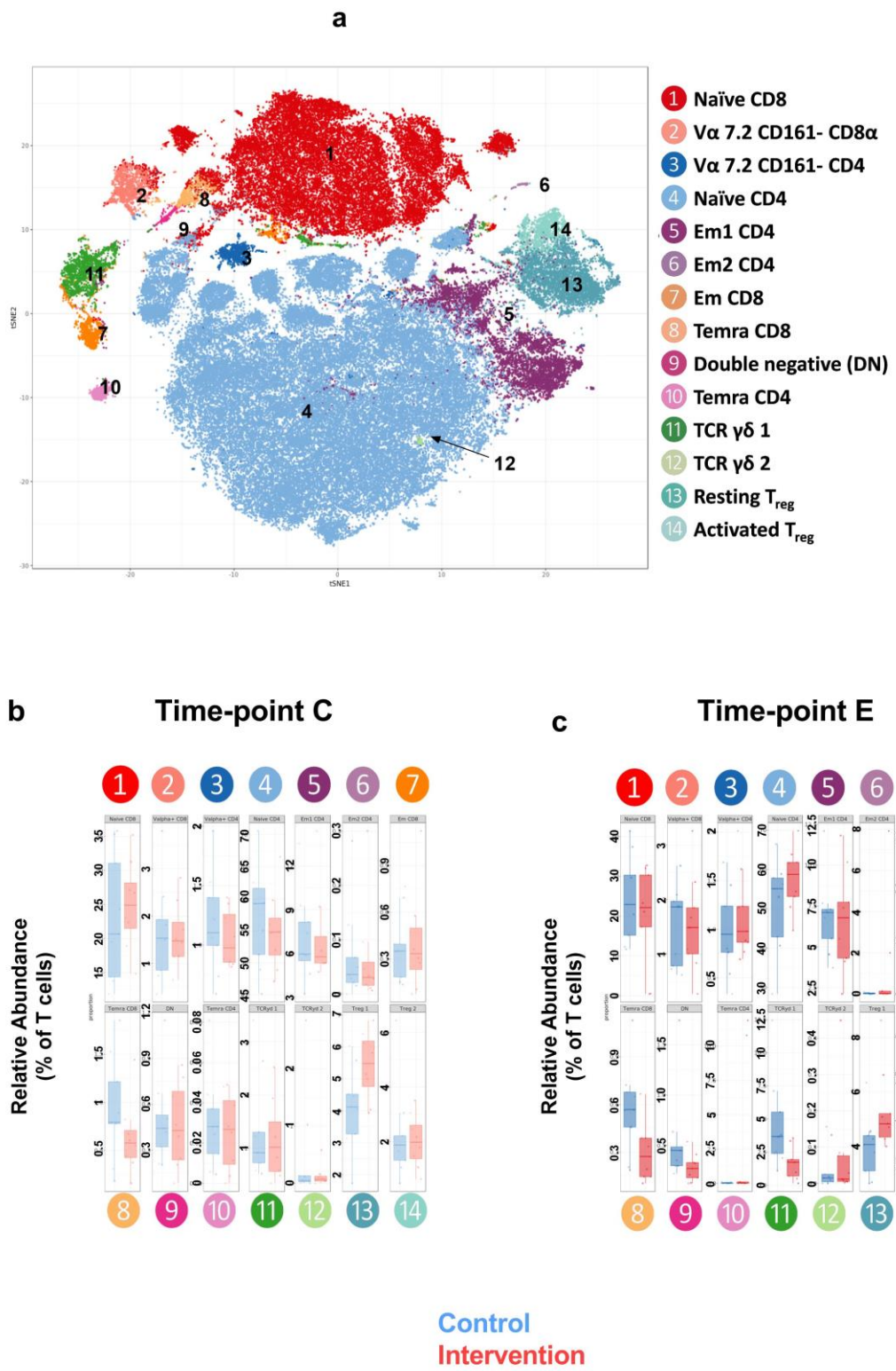


Figure 41 - Comparison of T cells between study groups. (a) tSNE plot created using infant T cells from paired samples. Clusters defined using heatmap (not shown). Cluster key given to right of tSNE plot (b) Comparison of relative abundance of each cluster between study groups at time-point C (c) Comparison of relative abundance of each cluster between study groups at time-point E

In order to further delineate the longitudinal changes in T cell subsets in early neonatal life, irrespective of dietary intervention, we sought to compare all samples at time-points C (full enteral feeds) and time-point E (study end (34 weeks CGA)) respectively. The results are displayed in Figure 42.

This revealed that the relative abundance of CD8 T_{EMRA} and activated T_{reg} cells decreased whilst the abundance of TCR γδ cells increased at the later time point. When comparing the normalised relative expression of markers within each cluster, Granzyme B is increased in every cluster in the heat map (Figure 42C). The polarisation of cells towards a T_h1 or T_h17 phenotype appears increased with time, as shown by increased relative expression of T-bet and ROR-γ in CD4, CD8 and TCRγδ cell clusters.

Together the results demonstrate that the T_{reg} compartment continues to evolve with fewer activated T_{reg}s over time, although this may represent loss of maternal derived T_{reg} s. Over time T cells appear to have an increased ability to express an effector response, by their increased expression of transcription factors as well as Granzyme B.

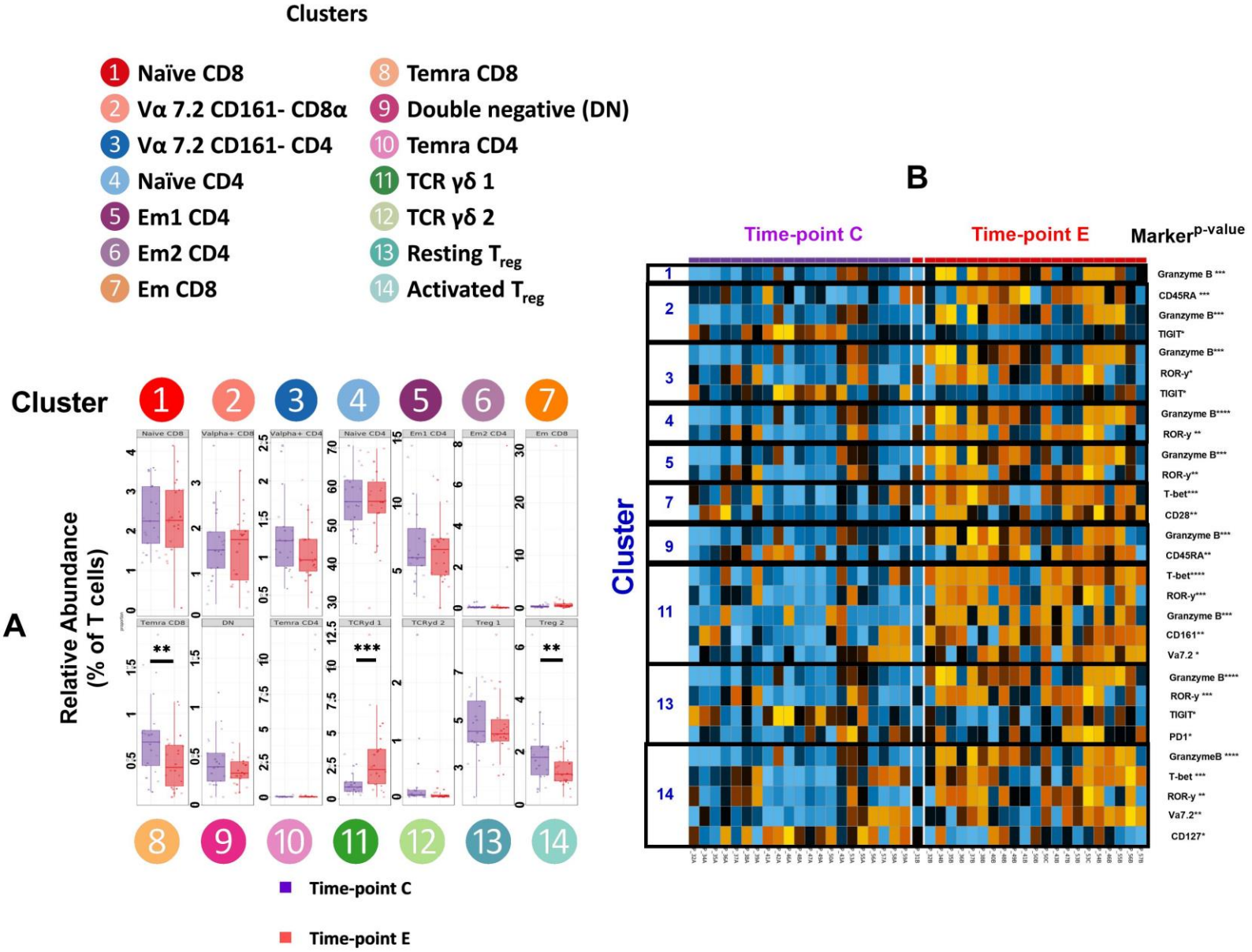


Figure 42 - Comparison of T cells between time-point C and E. Cluster key for t-SNE as generated in Figure 40 (A) Box plot displaying relative abundance of each cluster between time-points. Line represents median, box represents IQR and whiskers 1.5 x IQR. (B) Heatmap demonstrating the normalised relative expression between time-points. Markers that are significantly different are displayed with asterisks indicating p-value (*, **, *, ****; <0.05, <0.01, <0.001, 0.0001)**

4.3.5 *iNKT cells are predominantly of the iNKT2 subset in preterm infants*

iNKT cells were not identified as a separate cluster in the analysis above due to their rarity in peripheral blood. However, a targeted analysis defining iNKT cells as TCR Vα24-Jα18+ T cells, enabled the identification of interesting differences between adults and infants. These cells are rare in the peripheral blood, representing a mean of 0.05% (SD 0.07) and 0.17% (SD 0.11) of all T cells in adults and infants respectively. This represents a significant expansion in the population in infants compared to adults as demonstrated by Figure 43.

iNKT cells, defined as CD45+CD3+TCR Vα24-Jα18+ live cells, were identified in FCS Express, exported and analysed using cytofkit. In keeping with the published literature, five clusters of iNKT cells were identified based on their expression of cell surface and transcription markers, into iNKT1 (CXCR3+CCR4-T-bet+), iNKT2 (CCR4+CXCR3-), iNKT17 (CCR6+ RORγ+), Naïve (CCR7+CD45RA+) and "CCR6+HLADR+". The designation of iNKT1, iNKT2 and iNKT17 is in analogy with the T_h response (Crosby and Kronenberg, 2018, Krovi and Gapin, 2018, Moreira-Teixeira et al., 2011).

Infant iNKT cells appear to be predominantly iNKT2 whilst adult iNKT cells are predominantly iNKT1 cells. No significant difference was found in the relative abundance of iNKT cells (as % of T cells) or iNKT2 cells (as % of iNKT cells) when comparing either study groups (Control vs Intervention), time-points (C vs E) or a combination of study groups and time-points (C vs C and E vs E) using Mann-Whitney U tests.

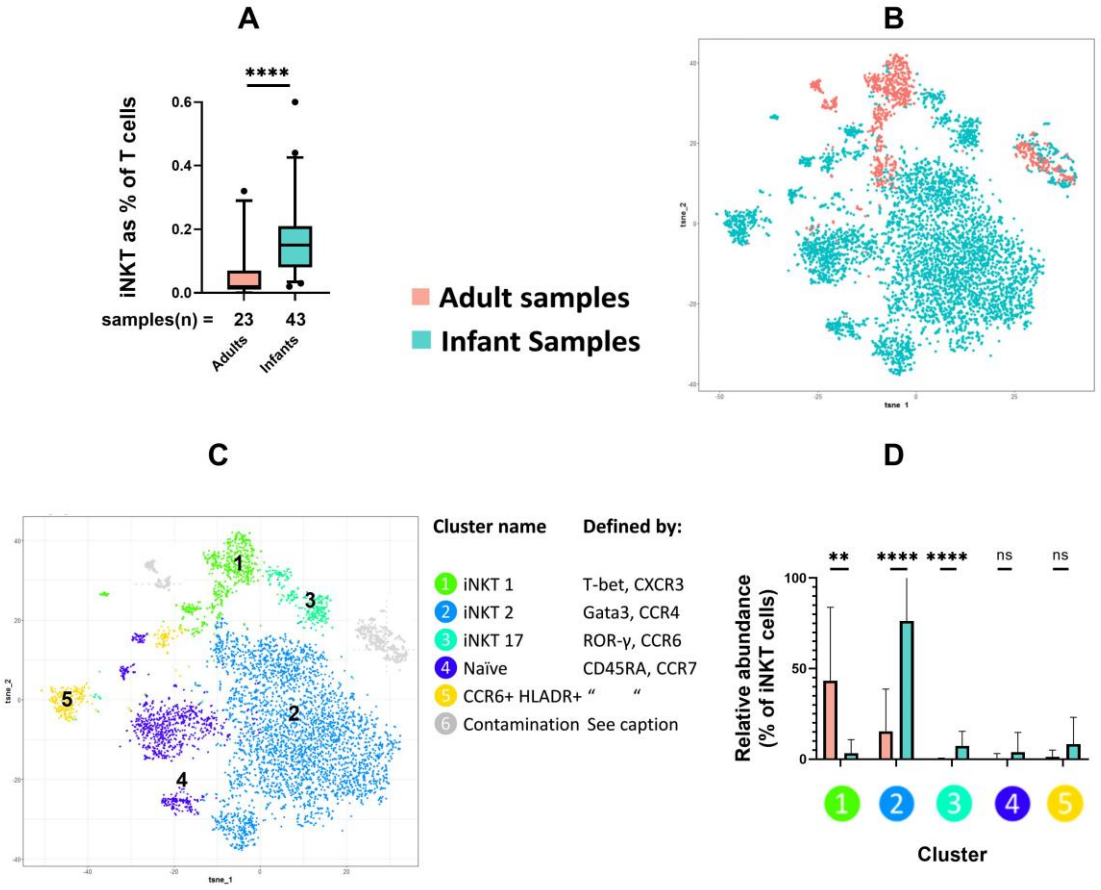


Figure 43 - Comparison of iNKT cells in adults compared to preterm infants. (A) Infants have an expanded iNKT population (% of T cells) in peripheral blood. (B) tSNE plot of iNKT cells based on expression of 33 cell markers, coloured by origin of samples. (C) tSNE plot displaying the identified iNKT subsets with key to the right showing defining markers (D) Relative abundance of each cluster between adult and infant samples. p-values unpaired student t-tests are displayed above the plots (*, **, *, ****; <0.05, <0.01, <0.001, 0.0001). Cells defined as contamination were identified as MAIT cells or expressed multiple antibodies not consistent with a cell type.**

4.3.6 *MAIT cells are less abundant and less mature in infants than adults*

The majority of T cells are described as αβ T cells due to the expression of a TCR that has both an α and β chain. MAIT cells are unusual as they have a restricted TCR repertoire, and can be defined by the expression of the invariant α chain of the TCR, TCR Vα7.2 together with CD161. Interestingly, the normalised median expression of the TCR Vα7.2 in all T cells is similar between infants and adults. Infants were found to have a large Vα7.2 CD161- population, as

shown in Figure 41. We therefore sought to identify if Vα7.2 CD161- cells share antigenic specificity with MAIT cells, which typically utilise the same Vα chain to bind the MR-1 non-classical HLA molecule, yet are CD161+. The tetramer specific for MAIT cells, MR1-5-OP-RU (MR-1), was acquired from the NIH tetramer core facility, USA. MR-1 mimics the natural ligand of MAIT cells, allowing the identification of these cells based on the specificity of their TCR (Le Bourhis et al., 2011). As expected, a significant proportion (30.6%) of adult Va7.2+CD8a+CD161+ T cells (n=3) bound MR-1, consistent with the classical description of MAIT cells. In contrast 9.5% of infant Va7.2+CD8a+ CD161+ T cells bound MR-1 (n=4). Conversely only 0.06% and 0.01% of adult and infant Va7.2+CD8a+CD161- T cells respectively bound MR-1. This confirms that thymic re-arrangement of the TCR Vα7.2 happens equally in preterm infants and adults; however, this is not sufficient to confer the ability to bind MR-1. There is evidence that CD161 plays an important role in fetal T cell activation following TCR activation (Halkias et al., 2019).

Figure 41(A-C) demonstrates that MAIT cells are significantly more abundant in adult than infant blood. We sought to determine whether these MAIT cells shared common properties in preterm infants and adults. CD3+CD8α+Vα7.2+CD161+ cells were identified in FCS express, exported and analysed using cytofkit as described in the Methods section. The results as displayed in Figure 44 reveal distinct clusters of MAIT cells. Adult MAIT cells formed populations distinct from infants. Clusters 2 and 7 were increased in abundance in adults. Cluster 7 has a mature phenotype with expression of the protein required for activation, CD28. Cluster 2 is identified by high expression of granzyme B and T-bet. Infant MAIT cells were predominantly in clusters 4 and 5. Cluster 5 has a naïve phenotype expressing CD45RA and CCR7, however cluster 4 was defined by cells displaying the early activation marker CD69, as well as CD28. Taken together, these results are consistent with the literature that the preterm MAIT cell compartment is both smaller and functionally less mature compared to that of adults (Walker et al., 2014, Ben Youssef et al., 2018).

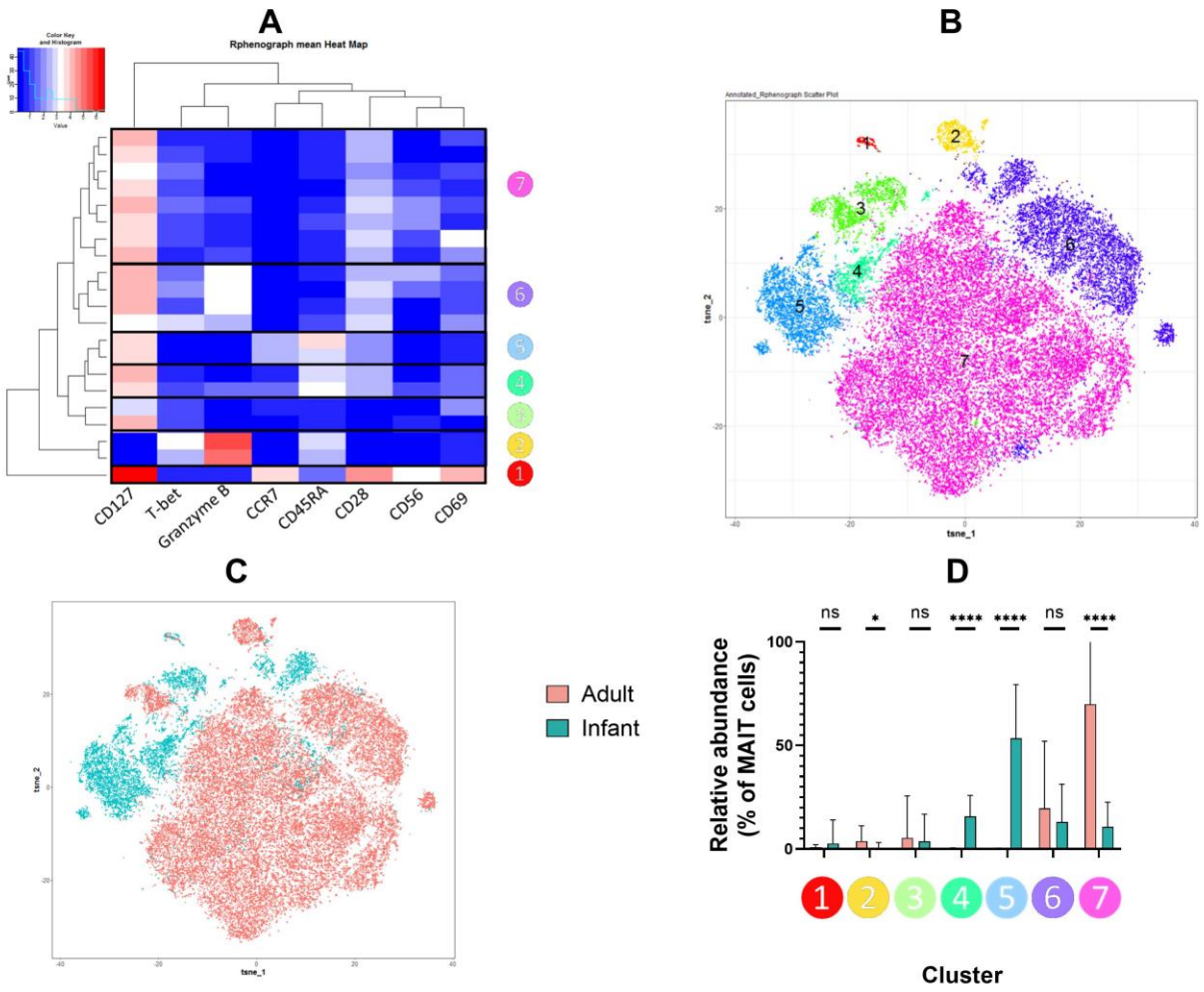


Figure 44 - Comparison of MAIT cells between infants and adults. A - Heatmap used to re-annotate 20 clusters into 7 clusters. B - TSNE plot with 7 clusters labelled. C - T-sne plot displaying whether cells are adult or infant derived. D - Bar graph displaying relative abundance of each cluster between adults and infants. p-values using unpaired student t-tests are displayed above the plots (*, **, *, ****; <0.05, <0.01, <0.001, 0.0001)**

4.4 Integrative analysis of diet, microbiome and T cells

The analyses above showed no clear effect of diet on either microbiome or T cell phenotype but did not address a putative relationship between microbiome and T cells. I therefore integrated the T cell abundance represented in Figure 41 into the sample metadata of each stool sample and performed a linear regression analysis to explore any correlation between abundance of T cell subsets and specific OTU's (genus or phylum level). This identified a negative correlation between the abundance of *Enterobacter* (genus level) and naïve CD8 T

cells. This correlation was stronger in the control group as displayed in Figure 45. This may simply reflect a temporal relationship with the abundance of Enterobacter increasing over time, with an independent decrease in naïve CD8 T cells. However, a similar correlation was not seen with *Staphylococcus* which also showed a significant decrease over time.

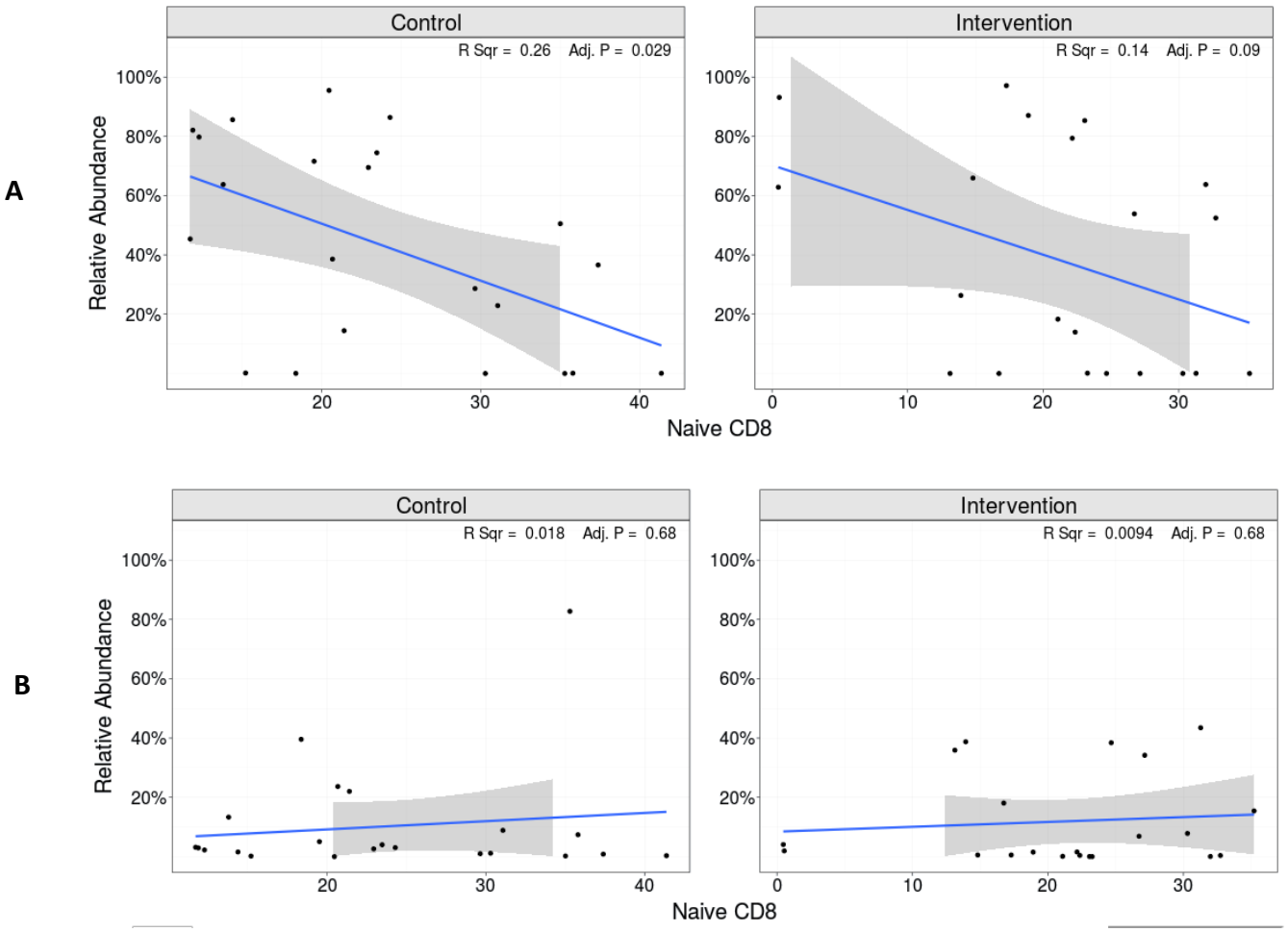


Figure 45 – Linear regression plots comparing the abundance of *Enterobacter* (A) and *Staphylococcus* (B) with the abundance of naive CD8+ T cells, stratified by study group. Relative abundance of *Enterobacter* is shown on the y- axis with relative abundance of naive CD8 T cells on the x-axis.

Chapter 5. Discussion

The focus of the study was on identifying features of the gut microbiome and T cell populations associated with alternative milk diets in extremely preterm infants. However, I uncovered potentially confounding differences between the study groups in relation to growth and receipt of MOM above and beyond the intended study intervention. Before discussing these findings, I present the four null hypotheses described in Section 2.2. The data presented in the results section meant that I was able to reject two of the four null hypotheses, these were the hypotheses numbered 1 and 4. Below I display and discuss the four null hypotheses before exploring other findings from the study, as well as the study's strengths and limitations.

5.1 Null Hypotheses

1. ***There is no association between the dietary intervention and the gut microbiome, either in terms of alpha diversity, beta diversity or composition at genus level, at 34 weeks corrected GA***.

When considering microbial composition in relation to the dietary intervention, I was able to find statistically significant differences between the control and intervention groups. There was a significant difference in unweighted beta diversity and abundance of *Lactobacillus* at genus level at the study end (34 weeks CGA). The increase in *Lactobacillus* was in the control group and did not appear attributable to any variation in probiotics (containing *Lactobacillus)* received or mode of delivery.

Whilst it is extremely difficult to be sure of direction of benefit when exploring gut microbial communities, *Lactobacillus* is often considered a beneficial genus of bacteria. Therefore, from a microbial perspective I could not identify any potential benefit of an exclusive human milk diet over a diet containing bovine products. Conversely, an increased abundance of *Lactobacillus* may be beneficial in the control group, which as demonstrated in Figures 28-33. The abundance of Lactobacillus was similar in infants fed mainly CMF (<30% MOM) or MOM fortified with BMF (>70% MOM).

2. ***There is no association between the dietary intervention and peripheral blood T cell subsets, specifically the abundance of T_h1, T_h2 or T_h17 cells or abundance of iNKT cells, MAIT cells or T_{reg}s***

When comparing dietary intervention and the abundance of T cells I could not detect any difference between study groups confirming the null hypotheses for iNKT, MAIT and T_{reg}s.

However, due to inadequate optimization leading to an inability to detect biological levels of transcription factors I could not be certain about the T_h1, T_h2 or T_h17 response.

3. ***There is no association between gut microbiome composition and peripheral blood T cell subsets, specifically the abundance of T_h1, T_h2 or T_h17 cells or abundance of iNKT cells, MAIT cells or T_{reg}s***

I was able to use linear regression analysis as described in Section 4.4 to compare the abundance of bacteria at genus level with abundance of T cell sub-populations. This only revealed one statistically significant correlation; however, this was most likely due to a temporal relationship between the change in *Enterobacter* and naïve CD8 T cells over time.

I believe this confirms the null hypothesis with no association between gut microbiome composition and T cell subsets, however as mentioned above, I cannot be certain about the T_h1, T_h2 or T_h17 response.

4. ***There are no significant differences in iNKT, MAIT or T_{reg} cell abundance in preterm infants compared to adults***

I have been able to reject this null hypothesis demonstrating clear differences in iNKT, MAIT and T_{reg} cell abundance, together with variable expression of cell and intracellular markers between preterm infants and adults.

Preterm infants have increased populations of naïve and regulatory T cells, reflecting a tolerant state that is yet to encounter many antigens. Innate T cells are systematically altered in abundance, being increased, in the case of iNKT cells, and decreased, in the case of MAIT cells. I have been able to show that these T cells have different characteristics depending on whether they are adult or infant derived, highlighting the complexity of exploring the immunophenotype.

5.2 Increased MOM in the control group

The finding that infants in the control group received an increased volume of MOM after correcting for weight and duration in the study was an interesting but perhaps unsurprising finding given that the fortifier in the intervention group replaced 30% of MOM, compared to a dry powder in the control group. However, the data suggested that the difference may not have been entirely related to the fortifier volume as described in Section 4.1.3, as when comparing the mean volume of MOM received in both study groups, this was less than 70%. The question of how introduction of DHM affects rates of MOM feeding has been addressed

but not in an RCT setting. Studies that have addressed this question have been observational or retrospective studies and have generally found that DHM leads to increased provision of any MOM at discharge, but does not increase exclusive MOM feeding (Williams et al., 2016). Conversely one centre reported a decrease in MOM provision following the introduction of DHM (Esquerra-Zwiers et al., 2014). This suggests there may be variation in how the introduction of DHM affects MOM provision in different geographical locations, potentially influenced by local beliefs about MOM and DHM.

This finding offers a word of caution about the introduction of DHM to any NICU. It would seem sensible for DHM introduction to occur alongside education of staff and parents about DHM, as well as monitoring the effect on MOM feeding rates. Monitoring would need to encompass detail of the amount of MOM received, as simply recording whether MOM was ever received, or received at specific time points would lose the impact on volume received as demonstrated by this study. However, the use of DHM to supplement the diet of preterm infants is not limited to the diet that was used in the INDIGO study. For instance, a diet could consist of DHM if there was a shortfall of MOM, but using BMF rather than HMF as a fortifier. This approach might have reduced the difference in MOM received between groups.

5.3 Decreased weight gain in the intervention group

As I have described in Section 1.1.5, it is not currently possible to define the optimal weight gain for a specific preterm infant. However, I found a disparity in weight gain between the two study groups with less weight gain in the intervention group. Whilst it is unclear if this will affect any meaningful clinical outcome, for example neurodevelopment, it might be a concern. A difficulty with interpreting this finding was that the increased mortality rate in the intervention group led to the skewing of infants completing the study towards increased birth weight and gestation compared to the control group. However, the disparity in weight gain was present when comparing two different time durations (whole study, and DOL14 to study end) so is not fully explained solely by differences in mortality. Furthermore, decreased weight gain has been reported in other studies of similar populations using donor human milk (Schanler et al., 2005, Cristofalo et al., 2013, Sullivan et al., 2010). For these reasons, I consider it likely that the differences in growth rates are a genuine finding, but are not explained by differences in macronutrient intakes which are similar. It is possible therefore, that differences in diet quality i.e. source of protein and lipids, impact on tissue accretion and growth, or there are other differences in metabolic processes that explain these growth effects.

5.4 Significant differences in functional and homing capacity of adult and infant T_{regs} populations

T_{reg}s were defined as a CD25+CD127-FoxP3+ population. I identified two populations of T_{reg}s, Activated T_{reg}s and Resting T_{reg}s based on their expression of CD45RA and CCR7 as has been described (Rosenblum et al., 2016). The activated T_{reg} population would include effector and memory T_{reg}s. The resting T_{reg} population that is expanded in infants compared to adults, expressed CD45RA and CCR7, and describes an antigen naïve group of cells. Conversely, the activated T_{reg} population did not express CD45RA or CCR7 suggesting these are activated cells following antigen exposure. The activated T_{reg} population expressed CCR4 and lacked expression of CXCR3 suggesting either a contaminating population of T_h2 cells or a population of cells expressing markers associated with both T_{reg}s and T_h2 cells as has previously been described (Kunicki et al., 2018) . Figure 40D displays the differential expression of markers within each cluster, between adults and infants. Interestingly, the relative expression of TIGIT and PD-1 varies in infants compared with adults, with increased relative expression of TIGIT and PD-1 on activated infant T_{reg}s. Resting Infant T_{reg}s have an increased expression of PD-1 but a decreased expression of TIGIT relative to adult resting T_{reg}s as shown in Figure 40D. TIGIT is an important co-inhibitory molecule, whilst PD-1 expressing T_{reg}s have similarly been shown to be refractory to stimulation (Kamada et al., 2019). Thus, there may be a difference in the balance of resting or activated T_{reg}s in infant compared to adult life. Furthermore, the infant activated T_{reg} population had an increased relative expression of the gut homing receptor CCR9 compared to adult activated T_{reg} cells. I believe these findings are novel and suggest the possibility that infant T_{reg} subsets defined on the basis of conventional markers may differ substantially from their adult counterparts in terms of functional capacity and homing.

A further possibility is that activated T_{reg}s are of maternal origin. T_{reg}s are thought to be important for tolerance to paternally derived antigens in pregnancy with the population known to expand. It is not definitively known why there are memory T_{reg}s, however it is thought they may confer a benefit in having an enhanced suppressor effect in antigen specific infections which has been demonstrated in mice (Rosenblum et al., 2016). They have been shown to transfer from mother to fetus in mice, and possibly confer a benefit for the next female generation in terms of exposure to paternal antigens (Kinder et al., 2015). The similar population of activated T_{reg} between infants and adults may therefore reflect maternal micro-chimerism, and this population being (at least partially) maternally derived may explain the decrease seen in activated T_{reg}s in the first weeks of life, as shown in Figure 40B.

5.5 The study had several strengths.

5.5.1 *Study population*

The study was designed to be minimally disruptive to normal neonatal care. This was facilitated by using non-invasive biological samples i.e. stool, for microbial analysis and timing blood samples to coincide with blood sampling for clinical reasons. Anthropometric measurements coincided with when the infant was being disturbed and form part of neonatal care.

I described a randomized population with two groups that were similarly matched at study entry. However, an increased mortality rate in the intervention group was observed (Figures 14-16). Three out of 7 of these deaths were in infants who had not been exposed to DHM. Of the remaining 4 infants, 3 ultimately died of respiratory failure (one had severe respiratory failure as a result of a congenital anomaly) and one infant had LOS. The 3 infants who deceased because of respiratory failure did so at 10, 12 and 19 days of age, making it unlikely that the respiratory failure was due to an enteral nutritional deficit. Otherwise it is hard to find a plausible link between these deaths due to respiratory failure and the intervention. The remaining infant who deceased with LOS was slightly unusual as they had *Serratia Marescens* isolated from autopsy samples, whilst the LOS was due to *Enterococcus Faecalis.* This infant died on DOL 6 after receiving 17 ml of RTF 26, out of 35ml total milk. At day of life 6, the gut microbiota is still establishing and it is unlikely that any intervention causing microbial dysbiosis would result in LOS this early in life. Furthermore, poor growth leading to an increased susceptibility to LOS would likely occur later in life. A result of these deaths was that the surviving infants in the intervention group had a later gestational age and higher birth weight in comparison to the control group. More mature Infants with an appropriate weight for gestational age are more likely to have a stable clinical course than immature, growth restricted infants. I believe this advantage to the infants in the intervention group reinforces my findings of no apparent clinical benefit of being in the intervention group. However, conversely it could be argued that any effect of diet may be more pronounced in more immature, growth-restricted infants due to their increased risk of disease and greater time receiving the study intervention.

I recorded a recruitment rate of 64% suggesting that the study was acceptable to a large number of parents, and I conducted the consent process carefully ensuring that parents could decline to participate in the study if they wished. Only one infant was withdrawn from the study, and there were no further discussions regarding study withdrawal, suggesting the study was acceptable to the majority.

I have been able to describe extensive longitudinal data consisting of every millilitre of nutrition the infants received during the study period. Typically, nutrition is recorded in a binary fashion on a daily basis, however detail is lost using this type of assessment. I was able to describe longitudinal sampling from birth until the end of the study period in 80% (41/51) of surviving recruits, enabling the identification of samples that would be most useful in isolating any effect of the dietary intervention. This was possible with minimal disturbance to normal neonatal care.

5.5.2 *Laboratory techniques*

I have described in depth analysis of both immune cells and microbial communities in extremely preterm infants. Whilst 16S sequencing is a relatively established technique, mass cytometry has only developed over the last 5 years and enabled me to explore the T cell populations of preterm infants more extensively than previously possible.

Investigating the immunophenotype of humans using this technique is now common, however I am aware of only one other study using mass cytometry in the postnatal samples of preterm infants (Olin et al., 2018).

5.6 Limitations

5.6.1 *Study population*

This study was designed only to explore a small number of infants, as previous studies suggest differences in gut microbial communities due to dietary interventions can be large. However, for any outcome, greater numbers of study recruits would have enabled more confidence in significant results.

This study was conducted at a NICU that is the regional referral unit for all surgical and cardiac infants. From the study perspective, this meant that some infants were transferred out of the hospital to district general hospitals, earlier than anticipated, accounting for loss to the study group with limited data from 6 infants. Ideally follow-on sites would have been able to continue study interventions, sampling and data collection. As the early transfers were not anticipated, the feasibility of follow-on sites was investigated during the study however due to the large number of potential follow-on sites and the small (possibly none) number of infants they would have cared for, this was not deemed feasible. Likewise waiting until an infant was known to be being transferred and then setting up that specific unit would not allow study intervention to continue quickly enough and was also not deemed feasible.

Unfortunately, when conducting research on extremely preterm infants, high rates of mortality and morbidity are expected. This can be a challenge in interpreting the data both because of loss of patients and also as a result of the introduction of variables (morbidities) that may influence the study outcome. I could not control for these variables as it is not clear what variation in morbidity would have on either immune cells or gut microbiota.

A limitation of the study, which I regard as a positive is that we reported high rates of MOM intake, with over 60% of the enteral nutrition being provided in the study period being MOM. Whilst this is a strength at the individual patient and NICU level, this may limit any effect a dietary intervention may have, assuming quantity is an important factor. A HMF that is 15ml rather than 30ml in volume is now commercially available, which may reduce the amount of MOM that is replaced when using a similar feeding strategy as that in intervention group of the study.

5.6.2 *Laboratory techniques*

5.6.2.1 Samples

A potential limitation in the study design was that I was investigating an interaction that commences in the gut lumen, but had to use biological samples collected at a more remote location i.e. stool (as opposed to small intestinal bacteria) and blood samples (as opposed to immune cells in the gut epithelium or lamina propria). Currently, there are no validated non-invasive methods to investigate interactions directly in the gut mucosa, but ex-vivo studies in the laboratory using organoid models may provide additional insights (Fofanova et al., 2019).

5.6.2.2 Microbiome analysis

5.6.2.2.1 Limitation of 16S rRNA sequencing

When using 16S, bacteria can only be delineated to the genus taxonomic level. The differentiation of different species or strains of bacteria requires alternate approaches. As all infants were exposed to probiotics, I was unable to distinguish whether species or strains of *Bifidobacteria* or *Lactobacillus* originated from the probiotic. It is possible there may be variation at species or strain level in these bacteria between groups that I have not been able to detect. WGS would allow resolution of bacteria to the strain taxonomic level enabling differentiation of bacterial species that are probiotic strains from those that are not. WGS is now being undertaken on some samples from the study however is not available at this time.

5.6.2.3 Identification of T_h response

5.6.2.3.1 Transcription factors

I planned to compare the expression of a number of transcription factors as part of my approach to immunophenotyping. I attempted to optimise the experiment to detect a range of transcription factors, however it became clear during later analysis that these transcription factors were not being measured at biological levels suggesting sub-optimal experiment conditions as highlighted in Section 3.5.8.4.3. This resulted in the inability to definitively compare the T_h1, T_h2 and T_h17 response using transcription factors. I was however able to compare relative expression of the transcription markers to indicate a direction of change between different populations. A limitation of this analysis is that what is detected may simply represent background variation in negatively stained cells. An alternative to this strategy would be the use of chemokine receptors to define the T_h response as mentioned in Figure 4.

As mentioned in Section 3.5.8.4 the reason for the decreased expression of T-bet, Gata3 and ROR-γ was likely due to heparin. Heparin is highly negative charged. Transcription factors have positively charged DNA binding domains, that enable binding to DNA (Liu et al., 2016). It is plausible heparin bound the transcription factors, and that the antibody was not able to bind. This would have resulted in minimal detection of transcription factor. Heparin was added to the mass cytometry staining as there have previously been reports of eosinophils non-specifically binding transcription factors (Rahman et al., 2016). In future, when using PBMC's, it seems pertinent to either not use heparin or compare staining without the addition of heparin. If leucocytes are used then the titration of heparin is important for mass cytometry analysis.

5.7 Conclusion

In conclusion, when comparing two groups of extremely preterm infants randomised either to an exclusive human milk diet or a diet containing bovine products, I was not able to identify any potentially beneficial changes in either gut microbiota, T cell, or amino acid composition. Preterm infants in the intervention group (receiving an exclusive human milk diet) received less MOM and had less weight gain, compared to the control group, which is consistent with previous work that explored a similar exclusive human milk diet (Cristofalo et al., 2013, Sullivan et al., 2010). This work is important as the most likely mechanism of any protective effect of DHM against NEC or LOS, in preterm infants, is through an effect on the gut microbiome.

5.8 Recommendations for future work

Our understanding of why preterm infants are susceptible to NEC and LOS remains limited. The purpose of this study was to compare the effect of two different diets on the gut microbiota and T cell populations. Further analysis of these datasets could try to establish effects of diet or disease on the gut microbiota or T cell populations. As infants were required to be fully enterally fed for blood samples to be taken, no blood samples were available on infants with NEC, and only one infant with SIP. Future work could address the effect of gastro-intestinal surgery on the populations of T cells that are known to preferentially populate the gut mucosa (MAIT, iNKT, T_{reg}s).

As mentioned, expression of transcription factors was not at biological levels, possibly as a result of heparin used in the experiment. Further analysis of the differential expression of chemokine receptors within the CD4 T cells could enable an opportunity to compare the T_h response between study groups.

Outside this study, future work into LOS and NEC should aim to explore both why preterm infants have this increased susceptibilty and potential interventions that can reduce this risk. Alongside this, NEC is likely the end result of a number of different mechanisms, and considering dividing this disease into the likely cause may prove useful to develop our understanding regarding the underlying mechanisms.

This study aimed to explore the diet, as a known risk factor for both NEC and LOS, in preterm infants. Alteration in the gut microbiome as a result of diet, is the most likely mechanism for any protective effect of diet. Considering this mechanism, this pilot study had more than 80% power to detect a 0.5SD difference in gut microbial alpha-diversity between study groups. This is a large difference in alpha diversity and future studies could attempt to exclude smaller differences in gut microbial compostion, in terms of effect on the gut microbiome of the exclusive human milk diet. The INDIGO study will comprise approximately 120 infants enabling greater resolution of any microbial differences. This will include a cohort of infants who have not been exposed to probiotics. It is possible the routine use of probiotics neutralises a major disadvantage of bovine milk products.

The exclusive human milk diet comprised both fortified donor human milk (RTF 26) and donor human milk fortifier(HMF, P+6). Future work should aim to separate these two dietary interventions, aiming to address whether either component affects outcomes. Ideally this would be a large RCT powered for an important clinical outcome such as NEC, as a large RCT has shown no effect on 2 year neuro-development using DHM (albeit using a BMF not a HMF) (O'Connor et al., 2018). Unfortunately, a RCT powered to detect a change is NEC as a primary

outcome is unlikely due to the number of study recruits required and as DHM is being readily accepted as part of neonatal care. A mechanistic study investigating microbial, metabolomic or immunological outcomes to understand the mechanism of action therefore seem extremely important and plausible, particularly as DHM can be so variable in composition. Mechanistic work could then influence further clinical RCT's.

As DHM is being readily accepted as part of neonatal care, developing our understanding of how the bioactive components in human milk are affected by postnatal age, freezing and pasteurisation, will help our understanding of how human milk can differ, as it may be for this reason that some diets using DHM improve outcomes whilst others do not.

As our understanding of the bioactive components improves, the use of organoid models using tissue samples obtained from surgery to explore gut mucosal interactions offer an opportunity to further investigate the influence of dietary components on the gut mucosa, which could then be supplemented with work in animal models. The difficulties with animal models are using a species that is able to survive extreme prematurity and can develop a similar disease to NEC without dramatic environment changes. Pigs offer the best model, as they develop a similar disease to NEC when given different enteral diets, however they are viable to survive at only 90% of normal gestation in comparison to 70% in humans which may be a limitation of this model. Nevertheless, there are many other similarities between preterm piglets and humans including size and the immaturity of their gastro-intestinal tract. They can also be nursed in a similar NICU environment (Sangild et al., 2013).

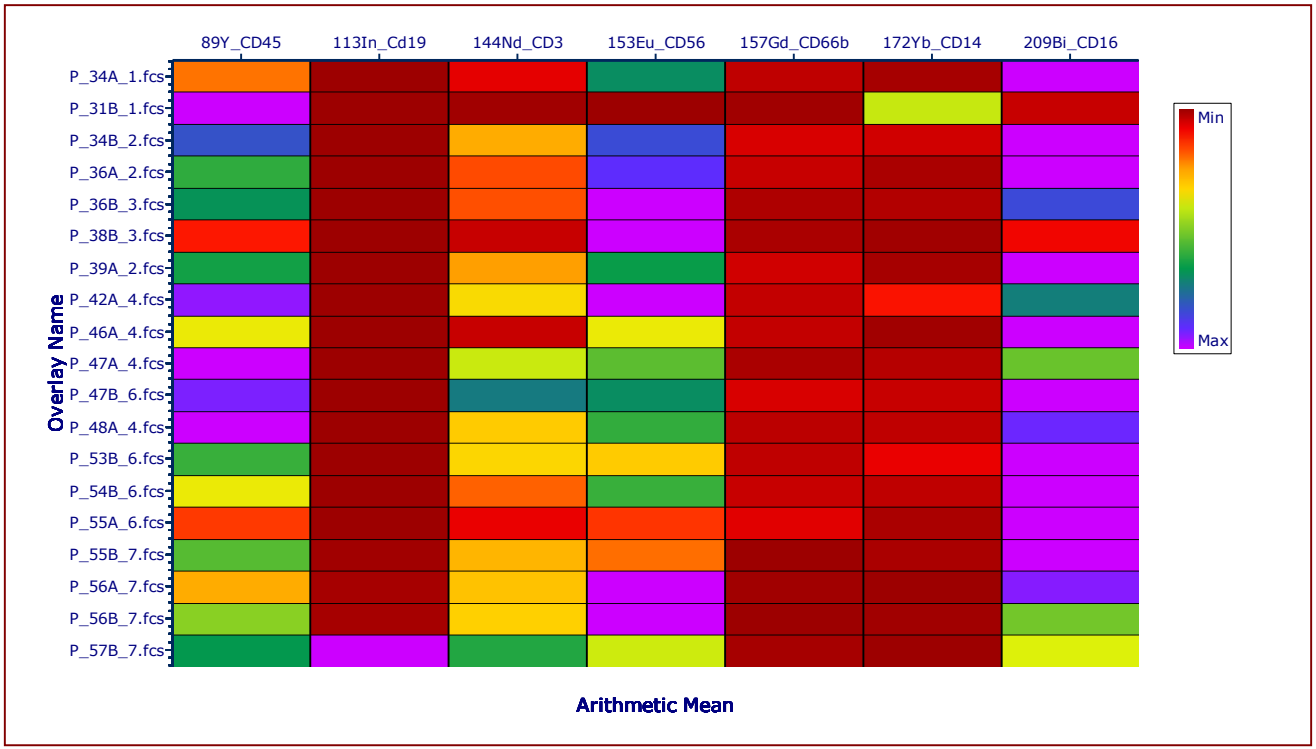
When exploring the gut microbiome in preterm infants in a NICU that uses probiotics, analysis should include either WGS or culturing to identify specific species of bacteria. This will enable the distinguishing of probiotic species from other potentially important species, and as mentioned the INDIGO cohort that was not exposed to probiotics may help our understanding.

There has been little research exploring the preterm immune system. This is despite not only diseases such as NEC and LOS, but chronic inflammatory diseases such as ROP and CLD being mediated by an immune response. In order to identify what effect interventions are having, we need to improve our understanding of the normal composition and function of the preterm infants immune system. Technologies utilising small sample volumes whilst delivering large datasets such as mass cytometry and transcriptomics could be extremely valuable in this area. This understanding could be utilised to develop our knowledge of why some preterm infants are preferentially susceptible to disease, and help understand the effect of interventions.

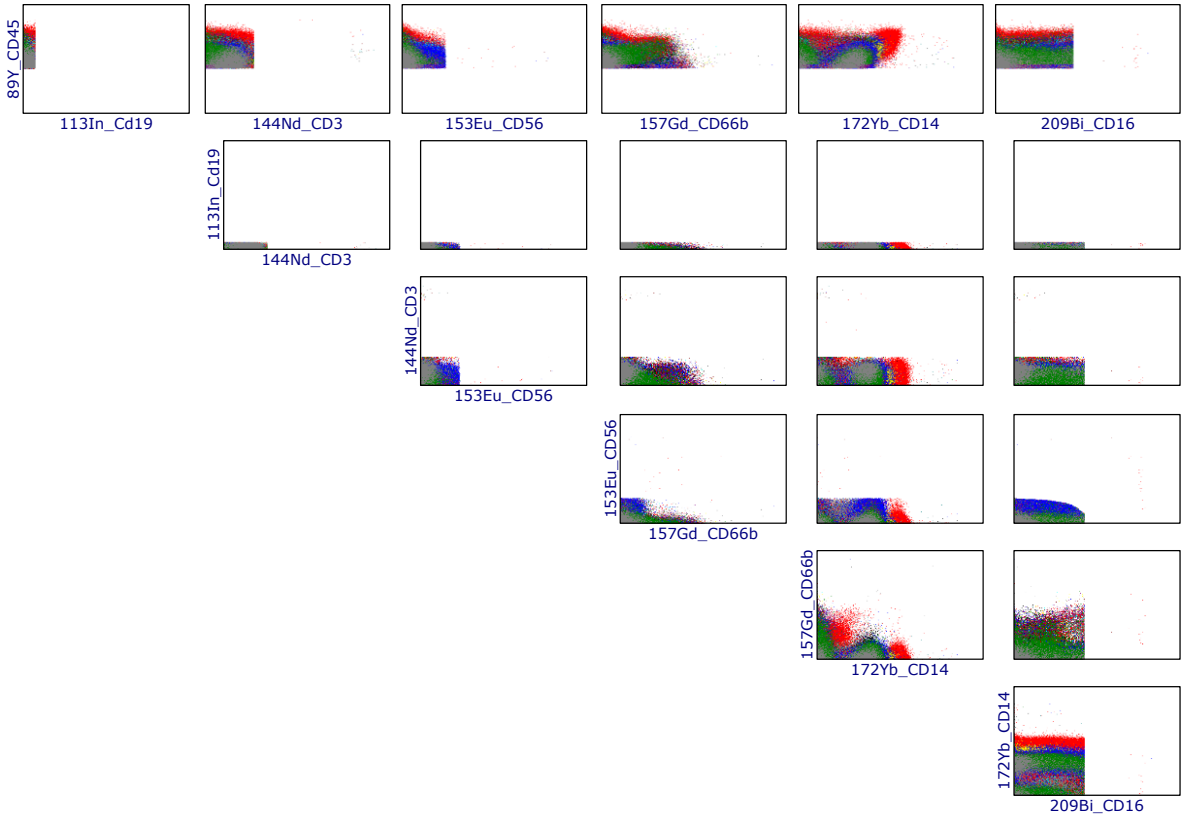
Lastly, but most importantly, the introduction of DHM to a NICU should include work exploring any changes in the provision of MOM to infants following introduction of DHM, as well as the

growth of infants. The introduction of DHM should not be at a cost in the provision of MOM to preterm infants.

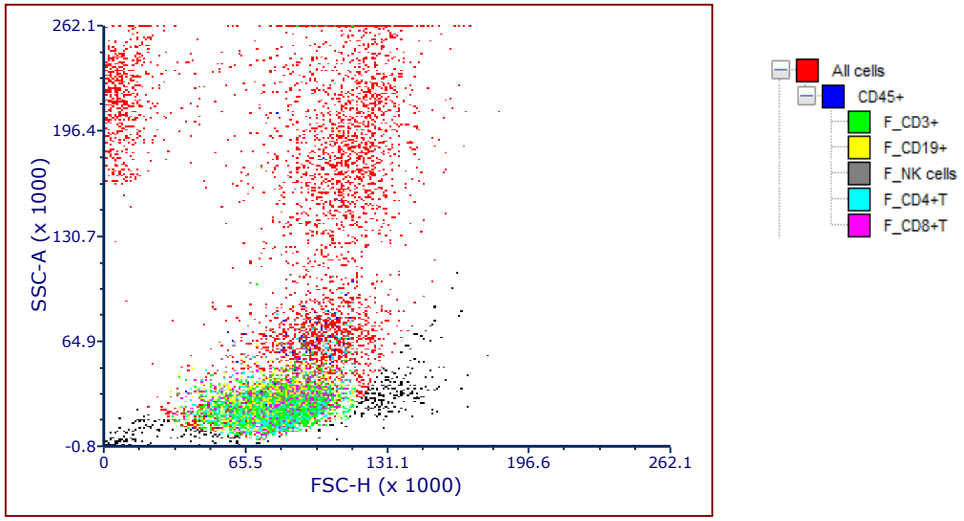
Supplementary Figures



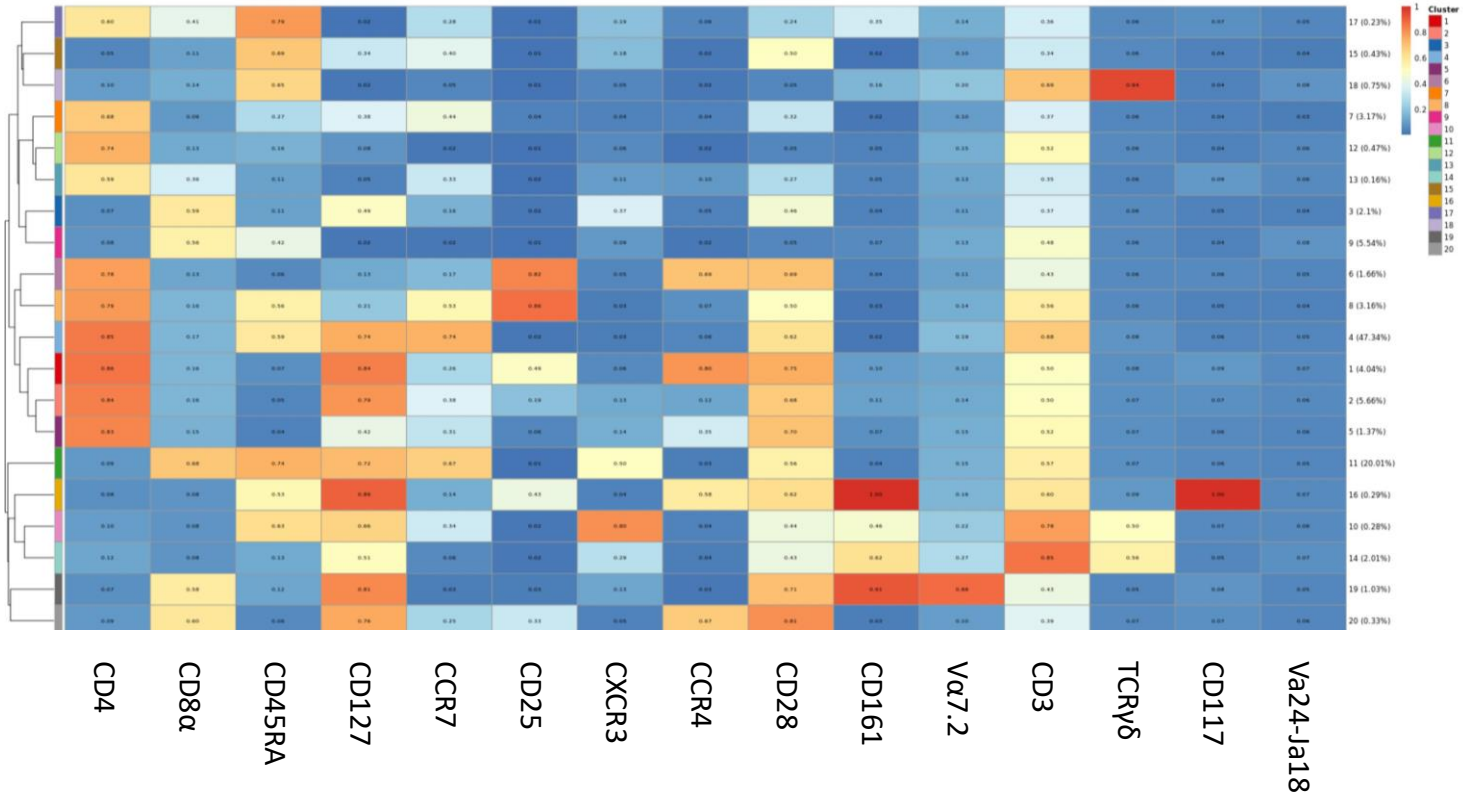
Supplementary Figure A – Heatmap comparing expression of lineage markers of the CD45+ cells not identified as T, B or NK cells between samples. Each sample is represented in a row with the mean arithmetic mean displayed as an expression from high to low (see legend on right of the plot)



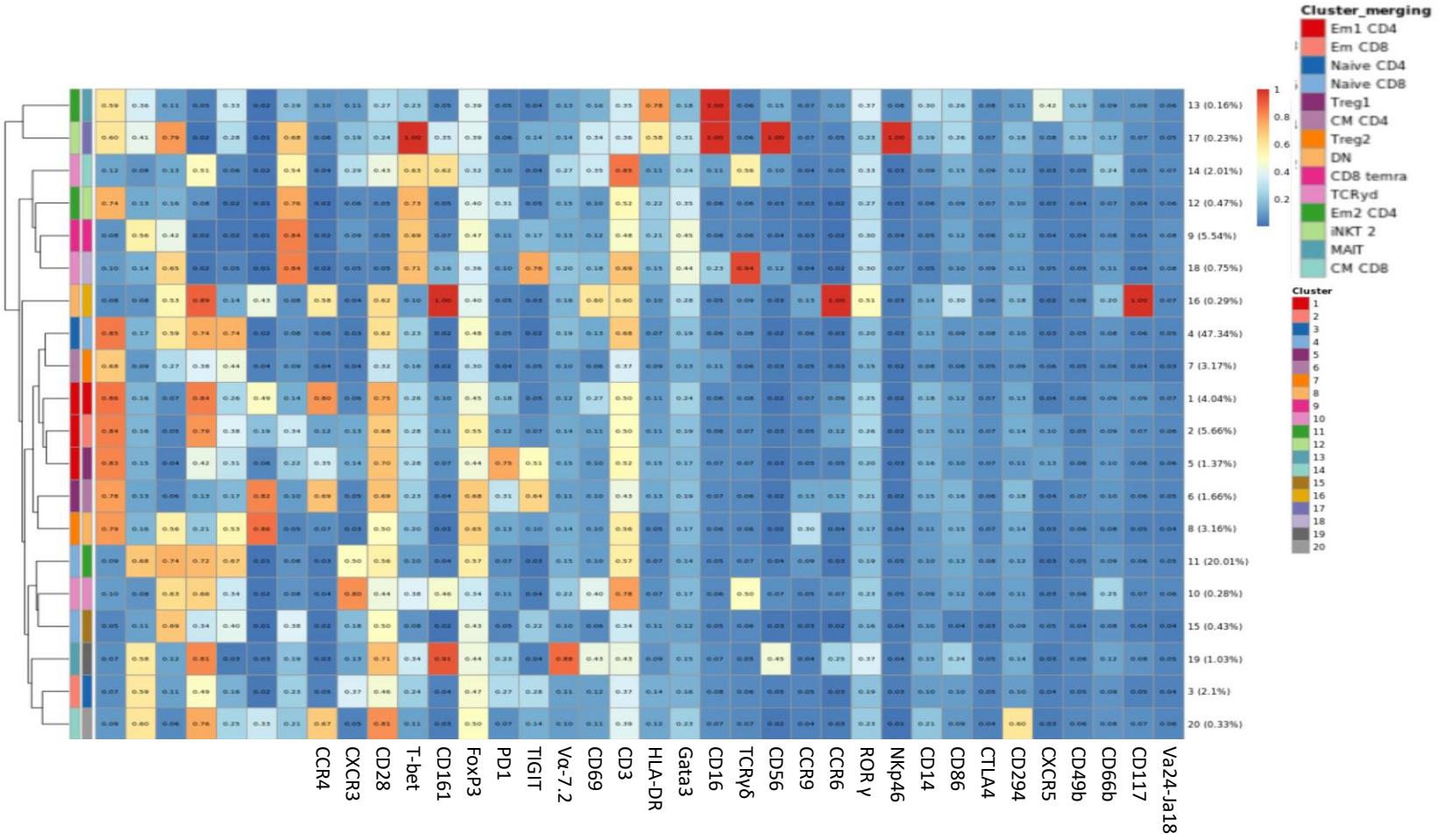
Supplementary Figure B – Colour dot plots displaying CD45+ cells not identified as T, B or NK cells. Each sample has been given a different colour. Each dot plot compares the expression of two lineage markers



Supplementary Figure C – Colour dot plot displaying all events in one representative sample of flow cytometry data. Forward scatter is on the x-axis with side scatter on the y axis. The key displays the cells represented by the different colours. This demonstrates that the lymphocyte population was in the expected location with few CD45+ cells outside this area.



Supplementary Figure D - Heatmap used for the identification of cell clusters in tSNE plot in Figure 40A. Markers used to define clusters are displayed on the x axis. Cluster dendogram is displayed on the left y axis and cluster percentage of total population displayed on right y-axis. Expression bar is shown in top right of plot with relative expression shown in each square.



Supplementary Figure E - Heatmap demonstrating expression of all markers in clusters in tsne plot in Figure 40A. Clusters merged are displayed in top right and colour coded, left y-axis demonstrates how clusters are merged. Right y-axis demonstrates the % of the T cell population the cluster represents. The expression of each marker is colour coded as shown by the key in the top right, whilst the expression of each marker is displayed in each square

Appendix A – INDIGO Patient Information Leaflet

The Newcastle upon Tyne Hospitals NHS
NHS Foundation Trust

Interactions between the diet and gut microbes, metabolism and body composition in preterm infants (INDIGO) IRAS Project ID: **215037**

Parent Information Leaflet (Newcastle Hospitals)
We would like to invite your baby to participate in the INDIGO study. Please read this information carefully; do ask us if you have any questions.

What is the purpose of this study?
The purpose of this study is to find out if there are differences in the type of gut microbes (bacteria), metabolism (how food is digested and used for growth), and the amount of fat and muscle (body composition), in preterm (premature) babies that receive an exclusive diet of human milk and human milk products, and those that receive a diet that includes products made from cow's milk (current standard practice). The information from this study will add to information from other studies to improve understanding of the best way to feed preterm babies.

Background information
Own Mother's Milk is recommended for preterm babies. The neonatal staff will do all they can to help you express enough milk for your baby. However, it can sometimes be difficult to express enough. In this situation either formula specially designed for preterm babies or pasteurised Human Donor Milk from a human milk bank, may be used to top-up any shortfall in Own Mother's Milk. There is wide variation in the use of each option. The choice is largely dependent on the preference of the doctor. Doctors that prefer formula top-ups believe this benefits growth, including brain growth and development, because formula has a consistent composition and a higher amount of nutrients. Human Donor Milk must be pasteurised (heat treated to destroy bacteria that might cause infection) and is therefore different to Own Mother's Milk. However, some doctors prefer donor milk because they believe the risk of a bowel disease in preterm babies called necrotising enterocolitis is lower if they use donor milk. Human Donor Milk use is not widespread in the UK; for example it is not used in Newcastle and other large parts of the UK.

Preterm babies sometimes grow slowly on human milk because it has a variable amount of nutrients. Therefore, a special supplement called 'fortifier' may be added to human milk. The fortifier in standard use is produced from cow's milk. Some doctors believe that cow's milk products may be harmful; others disagree, or feel that there is insufficient evidence either way. There is now a new fortifier available that is made from Human Donor Milk. This is used in some neonatal units in the US but is not currently commercially available in the UK.

Why has my baby been chosen?
Your baby is eligible for this study because s/he was born preterm (before 29 weeks gestation) and is receiving care at one of two hospitals taking part in this study (Newcastle, and Chelsea & Westminster, London). If you are medically unable to provide your own breast milk because of your own health (for example your own medical conditions or drug treatments) your baby is still eligible to join the study if you wish. If you intend to provide your own breast milk, but are not able to keep on doing so, your baby can stay in the study.

The Newcastle upon Tyne Hospitals NHS
NHS Foundation Trust

Does my baby have to take part?
No, it is entirely up to you to decide whether or not you wish your baby to take part. If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form, a copy of which will be given to you. If you decide not to take part this will not affect the care that your baby receives.

What if I change my mind after agreeing for my baby to take part?
If you decide to take part, you are free to withdraw your baby from the study at any time and without giving a reason. Unless you ask us not to, information collected up to that point will be included in the study.

What are the differences in the care my baby will receive depending on whether s/he participates in the INDIGO study?
The only difference between your baby's care and that of a baby not in the study is that your baby will have a fair and equal chance of receiving either formula top-ups or pasteurised human donor milk top-ups if the amount of expressed breast milk available from you is insufficient, and either cow's milk fortifier or fortifier made from pooled Human Donor Milk. This is done through a process called randomisation where information about your baby (gestation, hospital and twin status) are entered into a computer which randomly allocates your baby to one or other diet (donor milk or formula). The computer programme aims to make sure that at the end of the study there are approximately equal numbers of infants receiving each diet in the study (~50 per group). The research team do not know which diet your baby will receive beforehand. There is an equal chance that the baby will receive either of the two study diets (human milk or formula milk top-ups).

In the case of a baby not in the study, this choice of formula or Human Donor Milk is made according to the practices of the staff looking after your baby. Fortifier made from pooled Human Donor Milk is not currently used in the UK and is only available as part of this study.

What will happen to my baby if s/he takes part?
We aim to give your baby your milk within 24 hours of birth as is normal practice in our neonatal unit. Between 24-48 hours of birth your baby will be randomised to receive either formula or Human Donor Milk top-ups if the amount of expressed breast milk available from you is insufficient. When your baby is receiving 150 ml/kg of human milk fortifier will be added to any breast milk. This will either be the standard product we currently use, prepared from cow milk, or the new Prolacta product prepared from pooled human donor milk.

Milk top-ups as part of the INDIGO study will only continue until 34 weeks gestation (that is until 6 weeks before your baby's due date). After this point, your baby will receive breast feeds with top-ups of formula, if needed, as is current standard practice. Your baby will continue to receive the top-ups as part of the study during this time, regardless of how much of your own milk is available. Some babies may never receive any top-ups (if there is plenty of your milk) and some babies may end up mainly receiving the study milks (if you are unable or chose to stop providing your own milk.) After 34 weeks, infants who require top-ups

The Newcastle upon Tyne Hospitals NHS
NHS Foundation Trust

because there is no mother's milk) will receive a formula milk, and if needed a standard breast milk fortifier, according to standard hospital practice.

What tests and measurements will we do?

We will collect a few drops of urine (2-3 mls) and stool from the nappy (using a small plastic spoon), once a day until discharge to measure metabolite levels (waste products) and the pattern of gut bacteria. The human gut contains millions of bacteria. Many of these are 'healthy bacteria'. We want to examine this in preterm babies, and see how different diets affects which types are present. Some bacteria produce chemicals (metabolites). We can detect these in stool and urine by a test called Mass Spectrometry. We want to see if the pattern of chemicals matches changes in bacteria. In the future, we think the pattern of chemicals may tell us which babies are healthy and which are at risk of developing problems. We may also do tests to see if there is any inflammation in the gut. This may help us develop newer treatments or tests in the future.

We will obtain growth measurements, nutrition received, and results of routine blood tests that show the effect of nutrition on the body from your baby's case notes and electronic records.

We will ask the hospital laboratory to save any blood that is leftover after routine tests have been completed. Normally this is a very small volume (0.1mL = about 1-2 drops) and is usually thrown away in the bin. We will also ask your permission to take an extra blood sample (0.5-1mL) on two occasions: when they are tolerating 150ml/kg/day of milk and again before they stop the study at 34 weeks corrected age. We will only take this extra blood if and when they need a blood taking for routine tests and will ask your permission again before we do this closer to the time. This means your baby will never have any extra needle or invasive test just for the purpose of the research. The blood tests will measure proteins affecting growth and the types of white cells present (lymphocytes).

What happens to the samples after the testing is complete?

It will take several months and years to test the stool and urine samples we collect, and we may not analyse all of them. This means we may have some samples remaining at the end of the study. We will ask your permission to store these samples longer term in a biobank. These samples are a precious resource and can be used in further research into gut problems or infections in babies. The research may look at how genes are working in the gut (DNA from gut bacteria or gut cells present in the stool). Research teams may work with commercial partners to develop new techniques or tests, and some research may take place abroad. If you do not want to make these samples available for future research after the study has finished we will destroy them. The biological samples will be stored securely in a Research Tissue Bank at Newcastle University, that has received Ethical approval and in accordance

The Newcastle upon Tyne Hospitals NHS
NHS Foundation Trust

with the Human Tissue Act and national and local NHS Research Governance guidelines. All stored samples are anonymised and can only be used with approval from a steering committee and an ethics committee.

What will happen if my baby does not join the study?
If you choose not to participate, your baby will receive standard care which may include using preterm formula, donor milk or standard fortifier.

What products will be used in the INDIGO study?
We will use standard commercial formula and fortifier made from cow's milk, and new commercially produced Human Donor Milk and fortifier made from Human Donor Milk.
Human Donor Milk from UK Donor Milk Banks has very variable amounts of nutrients. Recently Human Donor Milk prepared from pooled donations has been produced by a US commercial bioscience company (Prolacta Bioscience); unlike Human Donor Milk from UK milk banks, this has a standard amount of nutrition. Prolacta Human Donor Milk is much more expensive than that obtained from a UK Human Milk Bank, and is not currently available in the UK. We will be using Prolacta Human Donor Milk in this study.

In the UK, when fortifier is added to human milk to boost nutrition, we use a commercial product made from cow's milk. Prolacta Bioscience also produces a new fortifer made from pooled human donor milk. However, this is expensive and is not currently available in the UK. We will use Prolacta fortifier in the INDIGO study.

What are the possible disadvantages and risks of taking part?
We do not think there are any risks to taking part in the study because when there isn't enough Own Mother's Milk, formula or Human Donor Milk are normally used as top-ups.
As with all studies, Newcastle Hospitals NHS Foundation Trust that is sponsoring this study holds insurance policies which apply to this study. In the unlikely event of your baby experiencing harm or injury as a result of taking part in the INDIGO study, you may be eligible to claim compensation without having to prove fault. This does not affect your legal rights to seek compensation.

If your baby is harmed due to someone's negligence, then you may have grounds for a legal action. Regardless of this, if you wish to complain, or have any concerns about any aspect of the way you have been treated during the course of this study you should inform a local investigator (Neonatal Consultants Dr Nicholas Embleton and Dr Janet Berrington). Normal NHS complaints mechanisms are also available to you.

What are the possible benefits of taking part?
We do not know whether individual babies will benefit from taking part but we hope the study results will help improve the way in which preterm babies are fed. The results may also help inform future decisions about the costs and possible benefits of different products.

What happens when the research study stops?
Your baby's involvement in the study will end when they are discharged from the neonatal unit. We will write to you to let you know the results of the study.

What if relevant new information becomes available?
Sometimes new information about the treatments being studied becomes available. If this happens, a member of the research team will inform you and discuss any implications.

Will my baby taking part in this study be kept confidential?
All information obtained about your baby during the course of the research will be kept strictly confidential. Study information that leaves the hospital will have your baby's name and address removed so that s/he cannot be recognised. Study information may be looked at by authorised people to check that it is being carried out correctly, all of whom have a duty of confidentiality to your baby.

What will happen to the results of the research study?
The results of this study will be presented in scientific journals, and at scientific meetings, submitted to NHS regulatory authorities, and reported on relevant websites such as ClinicalStudyResults.org and www.clinicaltrials.gov.

Who is organising and funding the research?
The study has been organised by doctors in Newcastle and London. The study is being funded by Prolacta Biosciences, a US company. Prolacta are providing milk products free of charge, and funds to conduct the study.

The overall conduct of the study will be under the review of a Trial Steering Committee which will have an independent chair and will not include any representatives from Prolacta. Prolacta will have no role in analysing the study results, writing study reports, or involvement in decisions about submission for publication.

Prolacta will not have access to identifiable information from any individual baby. At the end of the study the research team will share anonymised data with Prolacta.

Who has reviewed the study?
All research studies in the NHS are reviewed by a Research Ethics Committee, an independent group of people with responsibility to the Health Research Authority to protect your baby's rights, safety, wellbeing and dignity. This study has been approved by the North East Tyne and Wear South Research Ethics Committee.

If you want to raise your concerns with someone not involved in your care, you can contact the Patient Advise and Liaison Service (PALS). This service is confidential and can be contacted on Freephone: 0800 032 0202. Alternatively, if you wish to make a formal complaint you can contact the Patient Relations Department through any of the details below:

The Newcastle upon Tyne Hospitals NHS
NHS Foundation Trust

Telephone: 0191 223 1382 or 0191 223 1454

Email: patient.relations@nuth.nhs.uk

Address: Patient Relations Department, The Newcastle upon Tyne Hospitals NHS Foundation Trust, The Freeman Hospital, Newcastle upon Tyne, NE7 7DN

Further information and contact details
Please ask us about anything that is not clear, now or at any time during this study. Please speak to the doctor or nurse who discussed the study with you and gave you this leaflet. If you wish, you may also ask to speak to a member of the research team, listed below. Thank you for taking the time to read this information leaflet.

Dr Nicholas Embleton Dr Janet Berrington Staff Nurse Julie Groombridge

INDIGO_PIL_Newcastle_V1.3_060617

Appendix B – INDIGO consent form

The Newcastle upon Tyne Hospitals NHS
NHS Foundation Trust

CONSENT FORM

Interactions between the diet and gut microbes, metabolism and body composition in preterm infants (INDIGO)

Chief Investigator: Dr N D Embleton **Participant Number:**

Protocol Version: Version 1.3 dated 06/06/17

Please initial box

1. I have read and understand the information sheet Version 1.3 dated 06/06/2017 for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. I also confirm I have a copy of the information sheet to keep. ☐

2. I understand that my baby's participation is voluntary and that I am free to withdraw my baby without giving any reason, and without my baby's medical care or legal rights being affected. ☐

3. I give permission to regulatory authorities or individuals from NHS Trusts and organisations to look at my baby's health records in respect of the current study and any further research conducted in relation to it ☐

4. I agree that my baby's general practitioner may be informed of their participation in this study. ☐

5. I understand that this consent form does not release the hospital or the research team from their liability. ☐

6. I understand that my baby will be allocated at random to one of two milk diet groups. These diets involve supporting me to provide my breast milk which may be supplemented with additional milk and milk supplements (fortifiers) depending on need and according to standard practices on the neonatal unit. ☐

7. I agree (a) to the collection of samples of urine and stool from my baby's nappy (b) to the saving of any leftover samples of blood from the laboratory after routine tests are completed (c) to the collection of an extra sample of blood (less than 1/2 teaspoon) on 2 occasions taken at the same time as any routine blood tests needed and understand the research team will remind me of this closer to the time. ☐

8. I understand that any samples of stool, urine and blood that are not analysed during the course of the study can be kept longer term (for an indefinite period) in the Great North Neonatal Biobank at Newcastle University HTA licence 12534, ethics approval 15/NE/0334, IRAS 161883. These samples may be used in future research about nutrition or infections in preterm babies. The samples are stored in the UK and subject to NHS regulations, but analysis of some samples may take place outside the UK ☐

Parents/ guardian signature .....Date.....

Printed name .....Date.....

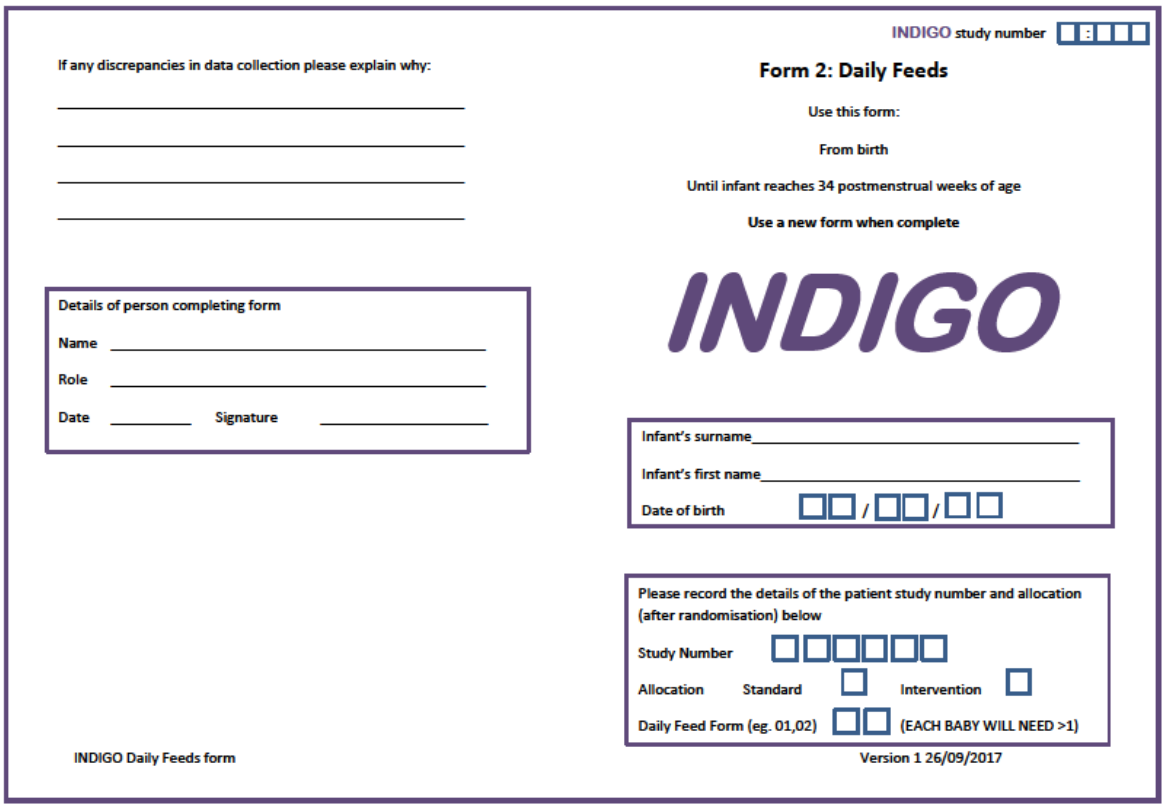
Witness signature .....Date.....

Printed name of witness .....

Copies: 1) Notes 2) Site file 3) parents

INDIGO study consent form Newcastle v1.3 060617

Appendix C – INDIGO data collection forms



INDIGO study number ☐:☐☐☐

If any discrepancies in data collection please explain why:

_____

_____

_____

_____

Form 2: Daily Feeds

Use this form:

From birth

Until infant reaches 34 postmenstrual weeks of age

Use a new form when complete

INDIGO

Details of person completing form

Name _____

Role _____

Date _____ Signature _____

Infant's surname_____

Infant's first name_____

Date of birth ☐☐/☐☐/☐☐

Please record the details of the patient study number and allocation (after randomisation) below

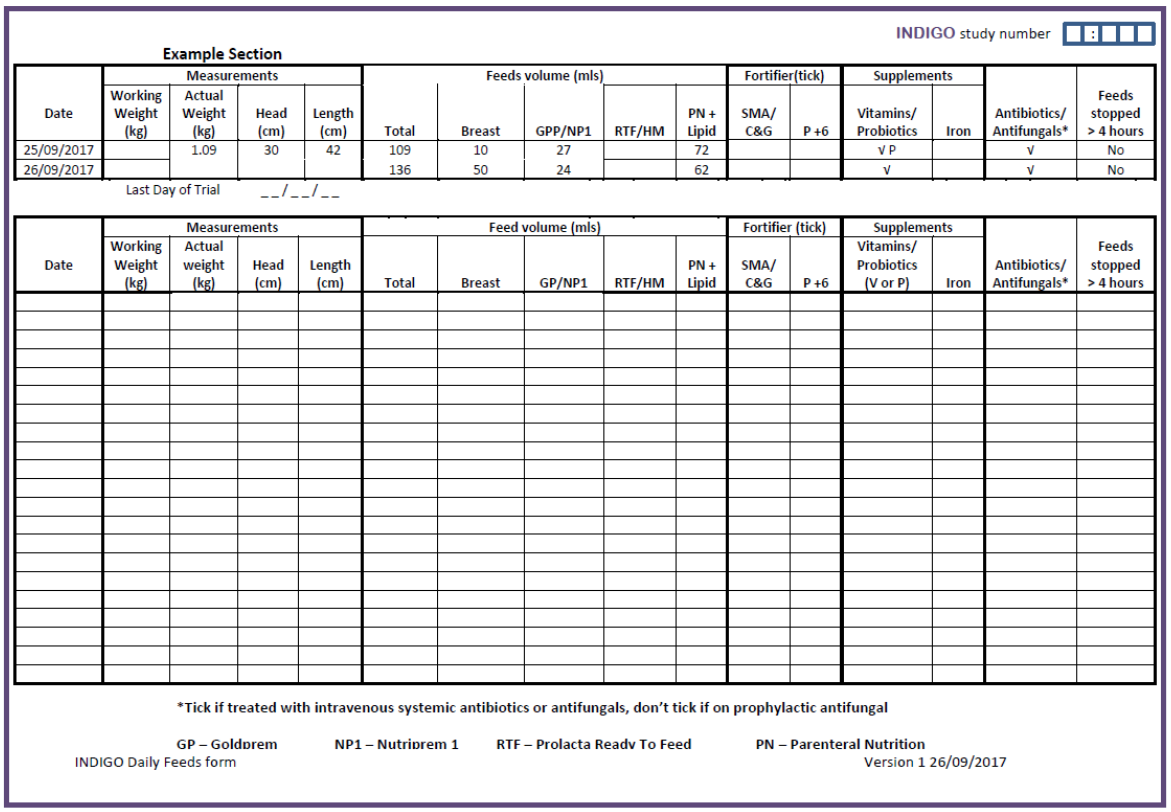
Study Number ☐☐☐☐☐☐

Allocation Standard ☐ Intervention ☐

Daily Feed Form (eg. 01,02) ☐☐ (EACH BABY WILL NEED >1)

INDIGO Daily Feeds form

Version 1 26/09/2017



INDIGO study number ☐:☐☐☐

Example Section

Date	Measurements				Feeds volume (mls)					Fortifier(tick)		Supplements		Antibiotics/ Antifungals*	Feeds stopped > 4 hours
	Working Weight (kg)	Actual Weight (kg)	Head (cm)	Length (cm)	Total	Breast	GPP/NP1	RTF/HM	PN + Lipid	SMA/ C&G	P +6	Vitamins/ Probiotics	Iron		
25/09/2017		1.09	30	42	109	10	27		72			√ P		√	No
26/09/2017					136	50	24		62			√		√	No

Last Day of Trial __/__/__

Date	Measurements				Feed volume (mls)					Fortifier (tick)		Supplements		Antibiotics/ Antifungals*	Feeds stopped > 4 hours
	Working Weight (kg)	Actual weight (kg)	Head (cm)	Length (cm)	Total	Breast	GP/NP1	RTF/HM	PN + Lipid	SMA/ C&G	P +6	Vitamins/ Probiotics (V or P)	Iron		

*Tick if treated with intravenous systemic antibiotics or antifungals, don't tick if on prophylactic antifungal

GP – Goldprem NP1 – Nutriprem 1 RTF – Prolacta Ready To Feed PN – Parenteral Nutrition

INDIGO Daily Feeds form

Version 1 26/09/2017

Form 1: Enrollment

Use this form:

From birth

Until infant reaches 34 postmenstrual weeks of age

Use a new form when complete

If any discrepancies in data collection please explain why:

_____

_____

_____

_____

Details of person completing form

Name _____

Role _____

Date _____ Signature _____

Infant's surname_____

Infant's first name_____

Date of birth ☐☐ / ☐☐ / ☐☐

Please record the details of the patient study number and allocation (after randomisation) below

Study Number ☐☐☐☐☐☐

Allocation Standard ☐ Intervention ☐

Daily Feed Form (eg. 01,02) ☐☐ (EACH BABY WILL NEED >1)

Trial Entry

Version 1 09/10/2017

Section A: Enrolment

A.1 **Is this infant's gestational age at birth less than 30 weeks?** Yes No

A.2 **Is this infant less than 72 hours old?** Yes No

A.3 Infant's expected date of delivery (EDD):
(by maternal scan if available, not LMP) DD/MM/YY

A.4 Infant's birth weight: _____ g

A.5 **Does this infant have a severe congenital anomaly?** Yes No

A.6 **Does this infant have a realistic prospect of survival?** Yes No

A.7 **Has this infant received any milk other that their mothers own milk?** Yes No

A.7 **Has written informed parental consent been obtained?** Yes No

If Yes, please PRINT name of person who obtained consent: _____

If Yes, please PRINT name of hospital where consent obtained: _____

A.8 What is this infant's sex? Male Female Indeterminate

A.9 Is this infant a singleton or multiple fetus? Singleton Multiple

A.10 Infant's NHS number: *(if available)* _____-_____-_____

A.11 **Is this baby likely to be transferred to another hospital before 34 weeks corrected gestation?** Yes No

Section B: Randomisation

B.1 **INDIGO study ID:** _____

Section C: Neonatal details

C.1 Was this infant delivered via caesarean section? Yes No

C.2 Were the membranes ruptured before labour? Yes No

C.3 Were the membranes ruptured >24 hours before delivery? Yes No

C.4 Was this infant's heart rate >100 bpm at 5 minutes of age? Yes No

C.5 Did this infant have absent or reversed end diastolic flow identified on any antenatal ultrasound scan? Yes No

Section D: Maternal details

D.1 Did the mother receive antenatal corticosteroids? Yes No

D.2 Mother's surname: _____

D.3 Mother's first name: _____

D.4 Mother's NHS number: _____

D.5 Mother's date of birth: DD/MM/YY

D.6 Main language: _____

D.8 Mother's mobile number: _____

Trial Entry

Version 1 09/10/2017

INDIGO study number ☐☐:☐☐☐☐

Example Section

Date	Measurements				Feeds volume (mls)					Fortifier(tick)		Supplements		Antibiotics/ Antifungals*	Feeds stopped > 4 hours
	Working Weight (kg)	Actual Weight (kg)	Head (cm)	Length (cm)	Total	Breast	GPP/NP1	RTF/HM	PN + Lipid	SMA/ C&G	P +6	Vitamins/ Probiotics	Iron		
25/09/2017		1.09	30	42	109	10	27		72			√ P		√	No
26/09/2017					136	50	24		62			√		√	No

Last Day of Trial __/__/__

Date	Measurements				Feed volume (mls)					Fortifier (tick)		Supplements		Antibiotics/ Antifungals*	Feeds stopped > 4 hours
	Working Weight (kg)	Actual weight (kg)	Head (cm)	Length (cm)	Total	Breast	GP/NP1	RTF/HM	PN + Lipid	SMA/ C&G	P +6	Vitamins/ Probiotics (V or P)	Iron		

*Tick if treated with intravenous systemic antibiotics or antifungals, don't tick if on prophylactic antifungal

GP – Goldprem NP1 – Nutriprem 1 RTF – Prolacta Ready To Feed PN – Parenteral Nutrition

INDIGO Daily Feeds form Version 1 26/09/2017

If any discrepancies in data collection please explain why:

_____

_____

_____

_____

Form 1: Enrollment

Use this form:

From birth

Until infant reaches 34 postmenstrual weeks of age

Use a new form when complete

INDIGO

Details of person completing form

Name _____

Role _____

Date _____ Signature _____

Infant's surname_____

Infant's first name_____

Date of birth ☐☐ / ☐☐ / ☐☐

Please record the details of the patient study number and allocation (after randomisation) below

Study Number ☐☐☐☐☐☐

Allocation Standard ☐ Intervention ☐

Daily Feed Form (eg. 01,02) ☐☐ (EACH BABY WILL NEED >1)

Trial Entry

Version 1 09/10/2017

Section A: Enrolment

A.1 Is this infant's gestational age at birth less than 30 weeks? Yes No

A.2 Is this infant less than 72 hours old? Yes No

A.3 Infant's expected date of delivery (EDD):

(by maternal scan if available, not LMP) DD/MM/YY

A.4 Infant's birth weight: _____g

A.5 Does this infant have a severe congenital anomaly? Yes No

A.6 Does this infant have a realistic prospect of survival? Yes No

A.7 Has this infant received any milk other that their mothers own milk? Yes No

A.7 Has written informed parental consent been obtained? Yes No

If Yes, please PRINT name of person who obtained consent: _____

If Yes, please PRINT name of hospital where consent obtained: _____

A.8 What is this infant's sex? Male Female Indeterminate

A.9 Is this infant a singleton or multiple fetus? Singleton Multiple

A.10 Infant's NHS number: *(if available)* _____-_____-_____

A.11 Is this baby likely to be transferred to another hospital before 34 weeks corrected gestation? Yes No

Section B: Randomisation

B.1 INDIGO study ID: _____

Section C: Neonatal details

C.1 Was this infant delivered via caesarean section? Yes No

C.2 Were the membranes ruptured before labour? Yes No

C.3 Were the membranes ruptured >24 hours before delivery? Yes No

C.4 Was this infant's heart rate >100 bpm at 5 minutes of age? Yes No

C.5 Did this infant have absent or reversed end diastolic flow identified on any antenatal ultrasound scan? Yes No

Section D: Maternal details

D.1 Did the mother receive antenatal corticosteroids? Yes No

D.2 Mother's surname: _____

D.3 Mother's first name: _____

D.4 Mother's NHS number: _____

D.5 Mother's date of birth: DD/MM/YY

D.6 Main language: _____

D.8 Mother's mobile number: _____

Trial Entry

Version 1 09/10/2017

Bibliography

(CHMP), C. F. T. M. P. F. H. U. & (PDCO), P. C. 2008. Guideline for the Investigation of Medicinal Products in the term and the preterm neonate. *European Medicines Agency*.

ABDULKADIR, B., NELSON, A., SKEATH, T., MARRS, E. C. L., PERRY, J. D. & CUMMINGS, S. P. 2016. Routine use of probiotics in preterm infants: longitudinal impact on the microbiome and metabolome. *Neonatology,* 109.

AFRAZI, A., SODHI, C. P., RICHARDSON, W., NEAL, M., GOOD, M., SIGGERS, R. & HACKAM, D. J. 2011. New Insights Into the Pathogenesis and Treatment of Necrotizing Enterocolitis: Toll-Like Receptors and Beyond. *Pediatric Research*.

AGOSTONI, C., BUONOCORE, G., CARNIELLI, V. P., DE CURTIS, M., DARMAUN, D., DECSI, T., DOMELLOF, M., EMBLETON, N. D., FUSCH, C., GENZEL-BOROVICZENY, O., GOULET, O., KALHAN, S. C., KOLACEK, S., KOLETZKO, B., LAPILLONNE, A., MIHATSCH, W., MORENO, L., NEU, J., POINDEXTER, B., PUNTIS, J., PUTET, G., RIGO, J., RISKIN, A., SALLE, B., SAUER, P., SHAMIR, R., SZAJEWSKA, H., THUREEN, P., TURCK, D., VAN GOUDOEVER, J. B., ZIEGLER, E. E. & NUTRITION, E. C. O. 2010. Enteral nutrient supply for preterm infants: commentary from the European Society of Paediatric Gastroenterology, Hepatology and Nutrition Committee on Nutrition. *J Pediatr Gastroenterol Nutr,* 50**,** 85-91.

ALFALEH, K. & ANABREES, J. 2014. Probiotics for prevention of necrotizing enterocolitis in preterm infants. *Cochrane Database Syst Rev,* 4.

AUTHORITY, H. R. 2017. UK policy framework for health and social care research.

AYDIN, M. S., YIGIT, E. N., VATANDASLAR, E., ERDOGAN, E. & OZTURK, G. 2018. Transfer and Integration of Breast Milk Stem Cells to the Brain of Suckling Pups. *Sci Rep,* 8**,** 14289.

AZAD, M. B., ROBERTSON, B., ATAKORA, F., BECKER, A. B., SUBBARAO, P., MORAES, T. J., MANDHANE, P. J., TURVEY, S. E., LEFEBVRE, D. L., SEARS, M. R. & BODE, L. 2018. Human Milk Oligosaccharide Concentrations Are Associated with Multiple Fixed and Modifiable Maternal Characteristics, Environmental Factors, and Feeding Practices. *J Nutr,* 148**,** 1733-1742.

BALATO, A., UNUTMAZ, D. & GASPARI, A. A. 2009. Natural killer T cells: an unconventional T-cell subset with diverse effector and regulatory functions. *J Invest Dermatol,* 129**,** 1628-42.

BALLARD, O. & MORROW, A. L. 2013. Human milk composition: nutrients and bioactive factors. *Pediatr Clin North Am,* 60**,** 49-74.

BATTERSBY, C., SANTHALINGAM, T., COSTELOE, K. & MODI, N. 2018. Incidence of neonatal necrotising enterocolitis in high-income countries: a systematic review. *Archives of Disease in Childhood - Fetal and Neonatal Edition,* 103**,** F182.

BEHRMAN, R. E., TYSON, J. E., LASKY, R. E., MIZE, C. E., RICHARDS, C. J., BLAIR-SMITH, N., WHYTE, R. & BEER, A. E. 1983. Growth, metabolic response, and development in very-low-birth-weight infants fed banked

human milk or enriched formula. I. Neonatal findings. *The Journal of Pediatrics,* 103**,** 95-104.
BEIN, A., EVENTOV-FRIEDMAN, S., ARBELL, D. & SCHWARTZ, B. 2018. Intestinal tight junctions are severely altered in NEC preterm neonates. *Pediatr Neonatol,* 59**,** 464-473.
BELFORT, M., RIFAS-SHIMAN, S., SULLIVAN, T., COLLINS, C., MCPHEE, A. J., RYAN, P., KLEINMAN, K., GILMAN, M. W., GIBSON, R. A. & MAKRIDES, M. 2011. Infant Growth Before and After Term: Effects on Neurodevelopment in Preterm Infants. *Pediatrics,* 128.
BELL M. J., T. J. L., FEIGIN R. D., KEATING J. P., MARSHALL R., BARTON L., BROTHERTON T. 1978. Neonatal Necrotizing Enterocolitis: Therapeutic Decisions Based upon Clinical Staging. *Annals of Surgery,* 187**,** 1-7.
BEN YOUSSEF, G., TOURRET, M., SALOU, M., GHAZARIAN, L., HOUDOUIN, V., MONDOT, S., MBURU, Y., LAMBERT, M., AZARNOUSH, S., DIANA, J. S., VIRLOUVET, A. L., PEUCHMAUR, M., SCHMITZ, T., DALLE, J. H., LANTZ, O., BIRAN, V. & CAILLAT-ZUCMAN, S. 2018. Ontogeny of human mucosal-associated invariant T cells and related T cell subsets. *J Exp Med,* 215**,** 459-479.
BENDALL, S. C., SIMONDS, E. F., QIU, P., AMIR EL, A. D., KRUTZIK, P. O., FINCK, R., BRUGGNER, R. V., MELAMED, R., TREJO, A., ORNATSKY, O. I., BALDERAS, R. S., PLEVRITIS, S. K., SACHS, K., PE'ER, D., TANNER, S. D. & NOLAN, G. P. 2011. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science,* 332**,** 687-96.
BENJAMINI, Y., KRIEGER, A. M. & YEKUTIELI, D. 2006. Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*.
BERGER, A. M. 2000. Th1 and Th2 responses: What are they? *BMJ*.
BERGMANN, K. R., LIU, S. X., TIAN, R., KUSHNIR, A., TURNER, J. R., LI, H. L., CHOU, P. M., WEBER, C. R. & DE PLAEN, I. G. 2013. Bifidobacteria stabilize claudins at tight junctions and prevent intestinal barrier dysfunction in mouse necrotizing enterocolitis. *Am J Pathol,* 182**,** 1595-606.
BERKHOUT, D. J. C., VAN KEULEN, B. J., NIEMARKT, H. J., BESSEM, J. R., DE BOODE, W. P., COSSEY, V., HOOGENES, N., HULZEBOS, C. V., KLAVER, E., ANDRIESSEN, P., VAN KAAM, A. H., KRAMER, B. W., VAN LINGEN, R. A., SCHOUTEN, A., VAN GOUDOEVER, J. B., VIJLBRIEF, D. C., VAN WEISSENBRUCH, M. M., WICAKSONO, A. N., COVINGTON, J. A., BENNINGA, M. A., DE BOER, N. K. H. & DE MEIJ, T. G. J. 2019. Late-onset Sepsis in Preterm Infants Can Be Detected Preclinically by Fecal Volatile Organic Compound Analysis: A Prospective, Multicenter Cohort Study. *Clin Infect Dis,* 68**,** 70-77.
BERRINGTON, J. E., BARGE, D., FENTON, A. C., CANT, A. J. & SPICKETT, G. P. 2005. Lymphocyte subsets in term and significantly preterm UK infants in the first year of life analysed by single platform flow cytometry. *Clinical and Experimental Immunology,* 140**,** 289-292.
BERRINGTON, J. E., HEARN, R. I., BYTHELL, M., WRIGHT, C. & EMBLETON, N. D. 2012. Deaths in Preterm Infants: Changing Pathology Over 2 Decades. *The Journal of Pediatrics,* 160**,** 49-53.e1.
BERZINS, S. P., SMYTH, M. J. & BAXTER, A. G. 2011. Presumed guilty: natural killer T cell defects and human disease. *Nat Rev Immunol,* 11**,** 131-42.

BIAGI, E., QUERCIA, S., ACETI, A., BEGHETTI, I., RAMPELLI, S., TURRONI, S., FALDELLA, G., CANDELA, M., BRIGIDI, P. & CORVAGLIA, L. 2017. The Bacterial Ecosystem of Mother's Milk and Infant's Mouth and Gut. *Front Microbiol,* 8**,** 1214.

BJORKLUND, K. L., VAHTER, M., PALM, B., GRANDER, M., LIGNELL, S. & BERGLUND, M. 2012. Metals and trace element concentrations in breast milk of first time healthy mothers: a biological monitoring study. *Environ Health,* 11**,** 92.

BODE, L. 2012. Human milk oligosaccharides: every baby needs a sugar mama. *Glycobiology,* 22**,** 1147-62.

BODE, L. 2015. The functional biology of human milk oligosaccharides. *Early Hum Dev,* 91**,** 619-22.

BODE, L. 2018. Human Milk Oligosaccharides in the Prevention of Necrotizing Enterocolitis: A Journey From in vitro and in vivo Models to Mother-Infant Cohort Studies. *Front Pediatr,* 6**,** 385.

BONNECHI, R., BIANCHI, G., BORDIGNON, P. P., D'AMBOSIO, D., LANG, R., BORSATTI, A., SOZZANI, S., ALLAVENA, P., GRAY, P. A., MANTOVANI, A. & SINIGAGLIA, F. 1998. Differential Expression of Chemokine Receptors and Chemotactic Responsiveness of Type 1 T Helper Cells (Th1s) and Th2s. *Journal Experimental Medicine,* 187**,** 129-134.

BRAVI, F., WIENS, F., DECARLI, A., DAL PONT, A., AGOSTONI, C. & FERRARONI, M. 2016. Impact of maternal nutrition on breast-milk composition: a systematic review. *Am J Clin Nutr,* 104**,** 646-62.

BROOKS, B., OLM, M. R., FIREK, B. A., BAKER, R., GELLER-MCGRATH, D., REIMER, S. R., SOENJOYO, K. R., YIP, J. S., DAHAN, D., THOMAS, B. C., MOROWITZ, M. J. & BANFIELD, J. F. 2018. The developing premature infant gut microbiome is a major factor shaping the microbiome of neonatal intensive care unit rooms. *Microbiome,* 6**,** 112.

BROWN, C. T., XIONG, W., OLM, M. R., THOMAS, B. C., BAKER, R., FIREK, B., MOROWITZ, M. J., HETTICH, R. L. & BANFIELD, J. F. 2018. Hospitalized Premature Infants Are Colonized by Related Bacterial Strains with Distinct Proteomic Profiles. *MBio,* 9.

BUISINE, M. P., DEVISME, L., SAVIDGE, T. C., GESPACH, C., GOSSELIN, B., PORCHET, N. & AUBERT, J. P. 1998. Mucin gene expression in human embryonic and fetal intestine. *Gut,* 43**,** 519-24.

BUNESOVA, V., LACROIX, C. & SCHWAB, C. 2018. Mucin Cross-Feeding of Infant Bifidobacteria and Eubacterium hallii. *Microb Ecol,* 75**,** 228-238.

BUTCHER, J., UNGER, S., LI, J., BANDO, N., ROMAIN, G., FRANCIS, J., MOTTAWEA, W., MACK, D., STINTZI, A. & O'CONNOR, D. L. 2018. Independent of Birth Mode or Gestational Age, Very-Low-Birth-Weight Infants Fed Their Mothers' Milk Rapidly Develop Personalized Microbiotas Low in Bifidobacterium. *J Nutr,* 148**,** 326-335.

CABINIAN, A., SINSIMER, D., TANG, M., ZUMBA, O., MEHTA, H., TOMA, A., SANT'ANGELO, D., LAOUAR, Y. & LAOUAR, A. 2016. Transfer of Maternal Immune Cells by Breastfeeding: Maternal Cytotoxic T Lymphocytes Present in Breast Milk Localize in the Peyer's Patches of the Nursed Infant. *PLOS ONE,* 11**,** e0156762.

CACHO, N. T., HARRISON, N. A., PARKER, L. A., PADGETT, K. A., LEMAS, D. J., MARCIAL, G. E., LI, N., CARR, L. E., NEU, J. & LORCA, G. L. 2017. Personalization of the Microbiota of Donor Human Milk with Mother's Own Milk. *Frontiers in Microbiology,* 8.

CANI, P. D. 2018. Human gut microbiome: hopes, threats and promises. *Gut,* 67**,** 1716-1725.
CARL, M. A., NDAO, I. M., SPRINGMAN, A. C., MANNING, S. D., JOHNSON, J. R., JOHNSTON, B. D., BURNHAM, C. A., WEINSTOCK, E. S., WEINSTOCK, G. M., WYLIE, T. N., MITREVA, M., ABUBUCKER, S., ZHOU, Y., STEVENS, H. J., HALL-MOORE, C., JULIAN, S., SHAIKH, N., WARNER, B. B. & TARR, P. I. 2014. Sepsis from the gut: the enteric habitat of bacteria that cause late-onset neonatal bloodstream infections. *Clin Infect Dis,* 58**,** 1211-8.
CHANG, H. Y., CHEN, J. H., CHANG, J. H., LIN, H. C., LIN, C. Y. & PENG, C. C. 2017. Multiple strains probiotics appear to be the most effective probiotics in the prevention of necrotizing enterocolitis and mortality: An updated meta-analysis. *PLoS One,* 12**,** e0171579.
CHEN, H., LAU, M. C., WONG, M. T., NEWELL, E. W., POIDINGER, M. & CHEN, J. 2016. Cytofkit: A Bioconductor Package for an Integrated Mass Cytometry Data Analysis Pipeline. *PLoS Comput Biol,* 12**,** e1005112.
CHEW, S. J., VICTOR, R. S., GOPAGONDANAHALLI, K. R. & CHANDRAN, S. 2018. Pneumatosis intestinalis in a preterm infant: should we treat all intestinal pneumatosis as necrotising enterocolitis? *BMJ Case Rep,* 2018.
CHIBA, A., MURAYAMA, G. & MIYAKE, S. 2018. Mucosal-Associated Invariant T Cells in Autoimmune Diseases. *Front Immunol,* 9**,** 1333.
CLEMINSON, J., AUSTIN, N. & MCGUIRE, W. 2015. Prophylactic systemic antifungal agents to prevent mortality and morbidity in very low birth weight infants. *Cochrane Database Syst Rev***,** CD003850.
COLE, T. J. F., J. V. 1998. British 1990 growth reference centiles for weight, height, body mass index and head circumference fitted by maximum penalised likelihood. *Statistics in Medicine*.
COLONY, P. C. 1983. Successive Phases of Human Fetal Intestinal Development. *Nutritional Adaptation of the Gastrointestinal Tract of the Newborn*.
CONG, X., JUDGE, M., XU, W., DIALLO, A., JANTON, S., BROWNELL, E. A., MAAS, K. & GRAF, J. 2017. Influence of Feeding Type on Gut Microbiome Development in Hospitalized Preterm Infants. *Nursing research,* 66**,** 123-133.
CONG, X., XU, W., JANTON, S., HENDERSON, W. A., MATSON, A., MCGRATH, J. M., MAAS, K. & GRAF, J. 2016. Gut Microbiome Developmental Patterns in Early Life of Preterm Infants: Impacts of Feeding and Gender. *PLOS ONE,* 11**,** e0152751.
CONTADOR, R., DELGADO-ADÁMEZ, J., DELGADO, F. J., CAVA, R. & RAMÍREZ, R. 2013. Effect of thermal pasteurisation or high pressure processing on immunoglobulin and leukocyte contents of human milk. *International Dairy Journal,* 32**,** 1-5.
CORNICK, S., TAWIAH, A. & CHADEE, K. 2015. Roles and regulation of the mucus barrier in the gut. *Tissue Barriers,* 3**,** e982426.
CORPELEIJN, W. E., DE WAARD, M., CHRISTMANN, V. & ET AL. 2016. Effect of donor milk on severe infections and mortality in very low-birth-weight infants: The early nutrition study randomized clinical trial. *JAMA Pediatrics,* 170**,** 654-661.
COSTA, S., MAGGIO, L., ALIGHIERI, G., BARONE, G., COTA, F. & VENTO, G. 2018. Tolerance of preterm formula versus pasteurized donor human

milk in very preterm infants: a randomized non-inferiority trial. *Ital J Pediatr,* 44**,** 96.
COSTELOE, K., HARDY, P., JUSZCZAK, E., WILKS, M. & MILLAR, M. R. 2016. Bifidobacterium breve BBG-001 in very preterm infants: a randomised controlled phase 3 trial. *Lancet,* 387.
CRISTOFALO, E. A., SCHANLER, R. J., BLANCO, C. L., SULLIVAN, S., TRAWOEGER, R., KIECHL-KOHLENDORFER, U., DUDELL, G., RECHTMAN, D. J., LEE, M. L., LUCAS, A. & ABRAMS, S. 2013. Randomized Trial of Exclusive Human Milk versus Preterm Formula Diets in Extremely Premature Infants. *The Journal of Pediatrics,* 163**,** 1592-1595.e1.
CROSBY, C. M. & KRONENBERG, M. 2018. Tissue-specific functions of invariant natural killer T cells. *Nat Rev Immunol,* 18**,** 559-574.
DAMACENO, Q. S., SOUZA, J. P., NICOLI, J. R., PAULA, R. L., ASSIS, G. B., FIGUEIREDO, H. C., AZEVEDO, V. & MARTINS, F. S. 2017. Evaluation of Potential Probiotics Isolated from Human Milk and Colostrum. *Probiotics Antimicrob Proteins,* 9**,** 371-379.
DANIELS, B., SCHMIDT, S., KING, T., ISRAEL-BALLARD, K., AMUNDSON MANSEN, K. & COUTSOUDIS, A. 2017. The Effect of Simulated Flash-Heat Pasteurization on Immune Components of Human Milk. *Nutrients,* 9.
DARRASSE-JEZE, G., MARODON, G., SALOMON, B. L., CATALA, M. & KLATZMANN, D. 2005. Ontogeny of CD4+CD25+ regulatory/suppressor T cells in human fetuses. *Blood,* 105**,** 4715-21.
DE MEIJ, T. G., VAN DER SCHEE, M. P., BERKHOUT, D. J., VAN DE VELDE, M. E., JANSEN, A. E., KRAMER, B. W., VAN WEISSENBRUCH, M. M., VAN KAAM, A. H., ANDRIESSEN, P., VAN GOUDOEVER, J. B., NIEMARKT, H. J. & DE BOER, N. K. 2015. Early Detection of Necrotizing Enterocolitis by Fecal Volatile Organic Compounds Analysis. *J Pediatr,* 167**,** 562-7 e1.
DERMYSHI, E., WANG, Y., YAN, C., HONG, W., QIU, G., GONG, X. & ZHANG, T. 2017. The "Golden Age" of Probiotics: A Systematic Review and Meta-Analysis of Randomized and Observational Studies in Preterm Infants. *Neonatology,* 112**,** 9-23.
DIRIX, V., VERMEULEN, F. & MASCART, F. 2013. Maturation of CD4+ Regulatory T Lymphocytes and of Cytokine Secretions in Infants Born Prematurely. *Journal of Clinical Immunology,* 33**,** 1126-1133.
DOGRA, S., SAKWINSKA, O., SOH, S.-E., NGOM-BRU, C., BRÜCK, W. M., BERGER, B., BRÜSSOW, H., LEE, Y. S., YAP, F., CHONG, Y.-S., GODFREY, K. M., HOLBROOK, J. D., DE VOS, W. M. & DOMINGUEZ BELLO, M. G. 2015. Dynamics of Infant Gut Microbiota Are Influenced by Delivery Mode and Gestational Duration and Are Associated with Subsequent Adiposity. *mBio,* 6.
DOMINGUEZ-BELLO, M. G., COSTELLO, E. K., CONTRERAS, M., MAGRIS, M., HIDALGO, G., FIERER, N. & KNIGHT, R. 2010. Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci U S A,* 107**,** 11971-5.
DUTTA, S., SINGH, B., CHESSELL, L., WILSON, J., JANES, M., MCDONALD, K., SHAHID, S., GARDNER, V. A., HJARTARSON, A., PURCHA, M., WATSON, J., DE BOER, C., GAAL, B. & FUSCH, C. 2015. Guidelines for feeding very low birth weight infants. *Nutrients,* 7**,** 423-42.

E CORPELEIJN, W., M P KOUWENHOVEN, S., PAAP, M., VLIET, I., SCHEERDER, I., MUIZER, Y., HELDER, O., GOUDOEVER, J. & VERMEULEN, M. 2012. *Intake of Own Mother's Milk during the First Days of Life Is Associated with Decreased Morbidity and Mortality in Very Low Birth Weight Infants during the First 60 Days of Life*.

EGAN, C. E., SODHI, C. P., GOOD, M., LIN, J., JIA, H., YAMAGUCHI, Y., LU, P., MA, C., BRANCA, M. F., WEYANDT, S., FULTON, W. B., NINO, D. F., PRINDLE, T., JR., OZOLEK, J. A. & HACKAM, D. J. 2016. Toll-like receptor 4-mediated lymphocyte influx induces neonatal necrotizing enterocolitis. *J Clin Invest,* 126**,** 495-508.

EL MANOUNI EL HASSANI, S., NIEMARKT, H. J., SAID, H., BERKHOUT, D. J. C., VAN KAAM, A. H., VAN LINGEN, R. A., BENNINGA, M. A., DE BOER, N. K. H. & DE MEIJ, T. G. J. 2018. Fecal Volatile Organic Compounds in Preterm Infants Are Influenced by Enteral Feeding Composition. *Sensors (Basel),* 18.

EMBLETON, N. D., KORADA, M., WOOD, C. L., PEARCE, M. S., SWAMY, R. & CHEETHAM, T. D. 2016. Catch-up growth and metabolic outcomes in adolescents born preterm. *Arch Dis Child,* 101**,** 1026-1031.

ESQUERRA-ZWIERS, A., WICKS, J., ROGERS, L. M., CHEN, Y., SCALA, C. M., JANES, J., MEIER, P. P., SILVESTRI, J. M. & PATEL, A. L. Impact of donor human milk in a high mother's own milk feeding neonatal intensive care unit. 17rg ISRHML Conference, 2014 South Carolina, USA.

EXCELLENCE, N. I. F. H. A. C. 2014a. Maternal and child nutrition 2008. Updated 2014 (Public Health Guideline 11).

EXCELLENCE, N. I. F. H. A. C. 2014b. Vitamin D: supplement use in specific population groups (Public Health Guideline 56).

EYERICH, S. & ZIELINSKI, C. E. 2014. Defining Th-cell subsets in a classical and tissue-specific manner: Examples from the skin. *Eur J Immunol,* 44**,** 3475-83.

FAST, C. & ROSEGGER, H. 1994. Necrotizing enterocolitis prophylaxis: oral antibiotics and lyophilized enterobacteria vs oral immunoglobulins. *Acta Paediatr Suppl,* 396**,** 86-90.

FENTON, A. C. 2017. Newcastle Neonatal Service: Annual Report 2017. *Accessed 26th July 2019,* https://nornet.org.uk/wp-content/uploads/2018/09/RVIReport2017.pdf.

FENTON, T. R. & KIM, J. H. 2013. A systematic review and meta-analysis to revise the Fenton growth chart for preterm infants. *BMC Pediatr,* 13**,** 59.

FLUIDIGM. 2016. *Maxpar Panel Designer* [Online]. [Accessed 20th August 2019].

FOFANOVA, T. Y., STEWART, C. J., AUCHTUNG, J. M., WILSON, R. L., BRITTON, R. A., GRANDE-ALLEN, K. J., ESTES, M. K. & PETROSINO, J. F. 2019. A novel human enteroid-anaerobe co-culture system to study microbial-host interaction under physiological hypoxia. *BioRxiv***,** 555755.

FOSTER, J. P., SETH, R. & COLE, M. J. 2016. Oral immunoglobulin for preventing necrotizing enterocolitis in preterm and low birth weight neonates. *Cochrane Database Syst Rev,* 4**,** CD001816.

FOTEINI, H., ADRIANA, B., ELLEN, C., ALISON, M. S., ALECIA‐JANE, T., PHILIPP, M., NAOMI, T., CHING TAT, L., LUIS, F., PILAR, B. & PETER, E. H. 2012. Breastmilk Is a Novel Source of Stem Cells with Multilineage Differentiation Potential. *STEM CELLS,* 30**,** 2164-2174.

FUKUDA, S., TOH, H., TAYLOR, T. D., OHNO, H. & HATTORI, M. 2012. Acetate-producing bifidobacteria protect the host from enteropathogenic infection via carbohydrate transporters. *Gut Microbes,* 3**,** 449-54.

GADALLA, R., NOAMANI, B., MACLEOD, B. L., DICKSON, R. J., GUO, M., XU, W., LUKHELE, S., ELSAESSER, H. J., RAZAK, A. R. A., HIRANO, N., MCGAHA, T. L., WANG, B., BUTLER, M., GUIDOS, C. J., OHASHI, P. S., SIU, L. L. & BROOKS, D. G. 2019. Validation of CyTOF Against Flow Cytometry for Immunological Studies and Monitoring of Human Cancer Clinical Trials. *Front Oncol,* 9**,** 415.

GALINDO-ALBARRÁN, ARIEL O., LÓPEZ-PORTALES, OSCAR H., GUTIÉRREZ-REYNA, DARELY Y., RODRÍGUEZ-JORGE, O., SÁNCHEZ-VILLANUEVA, JOSÉ A., RAMÍREZ-PLIEGO, O., BERGON, A., LORIOD, B., HOLOTA, H., IMBERT, J., HERNÁNDEZ-MENDOZA, A., FERRIER, P., CARRILLO-DE SANTA PAU, E., VALENCIA, A., SPICUGLIA, S. & SANTANA, M. A. 2016. CD8+ T Cells from Human Neonates Are Biased toward an Innate Immune Response. *Cell Reports,* 17**,** 2151-2160.

GHAZARIAN, L., CAILLAT-ZUCMAN, S. & HOUDOUIN, V. 2017. Mucosal-Associated Invariant T Cell Interactions with Commensal and Pathogenic Bacteria: Potential Role in Antimicrobial Immunity in the Child. *Front Immunol,* 8**,** 1837.

GIBBONS, D., FLEMING, P., VIRASAMI, A., MICHEL, M. L., SEBIRE, N. J., COSTELOE, K., CARR, R., KLEIN, N. & HAYDAY, A. 2014. Interleukin-8 (CXCL8) production is a signatory T cell effector function of human newborn infants. *Nat Med,* 20**,** 1206-10.

GIBBONS, D. L., HAQUE, S. F., SILBERZAHN, T., HAMILTON, K., LANGFORD, C., ELLIS, P., CARR, R. & HAYDAY, A. C. 2009. Neonates harbour highly active gammadelta T cells with selective impairments in preterm infants. *Eur J Immunol,* 39**,** 1794-806.

GIBSON, M. K., WANG, B., AHMADI, S., BURNHAM, C. A., TARR, P. I., WARNER, B. B. & DANTAS, G. 2016. Developmental dynamics of the preterm infant gut microbiota and antibiotic resistome. *Nat Microbiol,* 1**,** 16024.

GKOUSKOU, K. K., DELIGIANNI, C., TSATSANIS, C. & ELIOPOULOS, A. G. 2014. The gut microbiota in mouse models of inflammatory bowel disease. *Front Cell Infect Microbiol,* 4**,** 28.

GLADSTONE, M., OLIVER, C. & VAN DEN BROEK, N. 2015. Survival, morbidity, growth and developmental delay for babies born preterm in low and middle income countries - a systematic review of outcomes measured. *PLoS One,* 10**,** e0120566.

GLOCKER, E. O., KOTLARZ, D., BOZTUG, K., GERTZ, E. M., SCHAFFER, A. A., NOYAN, F., PERRO, M., DIESTELHORST, J., ALLROTH, A., MURUGAN, D., HATSCHER, N., PFEIFER, D., SYKORA, K. W., SAUER, M., KREIPE, H., LACHER, M., NUSTEDE, R., WOELLNER, C., BAUMANN, U., SALZER, U., KOLETZKO, S., SHAH, N., SEGAL, A. W., SAUERBREY, A., BUDERUS, S., SNAPPER, S. B., GRIMBACHER, B. & KLEIN, C. 2009. Inflammatory bowel disease and mutations affecting the interleukin-10 receptor. *N Engl J Med,* 361**,** 2033-45.

GOMEZ-GALLEGO, C., GARCIA-MANTRANA, I., SALMINEN, S. & COLLADO, M. C. 2016. The human milk microbiome and factors influencing its composition and activity. *Semin Fetal Neonatal Med,* 21**,** 400-405.

GONIA, S., TUEPKER, M., HEISEL, T., AUTRAN, C., BODE, L. & GALE, C. A. 2015. Human milk oligosaccharides inhibit Candida albicans invasion of human premature intestinal epithelial cells. *Journal of Nutrition,* 145**,** 1992-1998.

GOPALAKRISHNA, K. P., MACADANGDANG, B. R., ROGERS, M. B., TOMETICH, J. T., FIREK, B. A., BAKER, R., JI, J., BURR, A. H. P., MA, C., GOOD, M., MOROWITZ, M. J. & HAND, T. W. 2019. Maternal IgA protects against the development of necrotizing enterocolitis in preterm infants. *Nat Med,* 25**,** 1110-1115.

GREGORY, K. E., SAMUEL, B. S., HOUGHTELING, P., SHAN, G., AUSUBEL, F. M., SADREYEV, R. I. & WALKER, W. A. 2016. Influence of maternal breast milk ingestion on acquisition of the intestinal microbiome in preterm infants. *Microbiome,* 4**,** 68.

GRIFFITHS, J., JENKINS, P., VARGOVA, M., BOWLER, U., JUSZCZAK, E., KING, A., LINSELL, L., MURRAY, D., PARTLETT, C., PATEL, M., BERRINGTON, J., DORLING, J., EMBLETON, N. D., HEATH, P. T., ODDIE, S., MCGUIRE, W., AINSWORTH, S., BOYLE, E., CLARKE, P., CRAIG, S., JOHNSON, K., MACTIER, H., SCORRER, T., LEDWIDGE, M., STORY, I., HOLDER, G., OHADIKE, P., ELLIS, S., VAIKUTE, R., GOWDA, G., YATES, H., GARG, S., PILLING, E., ROEHR, C., BATRA, D., GIBSON, D., JOHNSON, M., KUMAR, Y., BARTLE, D., PETERS, C., QUINE, D., GUPTA, R., MATTHES, J., KENNEA, N., REYNOLDS, P., GEETHANATH, R., JANAKIRAMAN, S., VASU, V. & MANJUNATHA, C. M. 2019. Enteral lactoferrin supplementation for very preterm infants: a randomised placebo-controlled trial. *The Lancet,* 393**,** 423-433.

GROUP, E. T. I. 2013. Lactoferrin immunoprophylaxis for very preterm infants. *Arch Dis Child Fetal Neonatal Ed,* 98**,** F2-4.

HACKAM, D. J. & SODHI, C. P. 2018. Toll-Like Receptor-Mediated Intestinal Inflammatory Imbalance in the Pathogenesis of Necrotizing Enterocolitis. *Cell Mol Gastroenterol Hepatol,* 6**,** 229-238 e1.

HAHN, W. H., KIM, J., SONG, S., PARK, S. & KANG, N. M. 2017. The human milk oligosaccharides are not affected by pasteurization and freeze-drying. *J Matern Fetal Neonatal Med***,** 1-7.

HALKIAS, J., RACKAITYTE, E., HILLMAN, S. L., ARAN, D., MENDOZA, V. F., MARSHALL, L. R., MACKENZIE, T. C. & BURT, T. D. 2019. CD161 contributes to prenatal immune suppression of IFNgamma-producing PLZF+ T cells. *J Clin Invest,* 130**,** 3562-3577.

HAMPEL, D., SHAHAB-FERDOWS, S., ISLAM, M. M., PEERSON, J. M. & ALLEN, L. H. 2017. Vitamin Concentrations in Human Milk Vary with Time within Feed, Circadian Rhythm, and Single-Dose Supplementation. *J Nutr,* 147**,** 603-611.

HARRIS, P. A., TAYLOR, R., MINOR, B. L., ELLIOTT, V., FERNANDEZ, M., O'NEAL, L., MCLEOD, L., DELACQUA, G., DELACQUA, F., KIRBY, J., DUDA, S. N. & CONSORTIUM, R. E. 2019. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform,* 95**,** 103208.

HASSIOTOU, F., HEPWORTH, A. R., METZGER, P., TAT LAI, C., TRENGOVE, N., HARTMANN, P. E. & FILGUEIRA, L. 2013. Maternal and infant infections stimulate a rapid leukocyte response in breastmilk. *Clin Transl Immunology,* 2**,** e3.

HASSIOTOU, F., MOBLEY, A., GEDDES, D., HARTMANN, P. & WILKIE, T. 2015. Breastmilk Imparts the Mother's Stem Cells to the Infant. *The FASEB Journal,* 29**,** 876.8.
HAYNES, B. F. & HEINLY, C. S. 1995. Early human T cell development: analysis of the human thymus at the time of initial entry of hematopoietic stem cells into the fetal thymic microenvironment. *J Exp Med,* 181**,** 1445-58.
HE, Y., LAWLOR, N. T. & NEWBURG, D. S. 2016. Human Milk Components Modulate Toll-Like Receptor-Mediated Inflammation. *Adv Nutr,* 7**,** 102-11.
HE, Y., LIU, S., LEONE, S. & NEWBURG, D. S. 2014. Human colostrum oligosaccharides modulate major immunologic pathways of immature human intestine. *Mucosal Immunol,* 7**,** 1326-39.
HEALTH, D. O. 2013. DH Guidance Notes on the Infant Formula and Follow-on Formula Regulations 2007 (as amended).
HEALTH, R. C. O. P. A. C. 2017. National Neonatal Audit Programme: A guide to the 2018 measures. *Royal College of Paediatrics and Child Health*.
HELLER, F., FUSS, I. J., NIEUWENHUIS, E. E., BLUMBERG, R. S. & STROBER, W. 2002. Oxazolone colitis, a Th2 colitis model resembling ulcerative colitis, is mediated by IL-13-producing NK-T cells. *Immunity,* 17**,** 629-638.
HICKEY, M., GEORGIEFF, M. & RAMEL, S. 2018. Neurodevelopmental outcomes following necrotizing enterocolitis. *Semin Fetal Neonatal Med,* 23**,** 426-432.
HO, T. T. B., GROER, M. W., KANE, B., YEE, A. L., TORRES, B. A., GILBERT, J. A. & MAHESHWARI, A. 2018. Dichotomous development of the gut microbiome in preterm infants. *Microbiome,* 6**,** 157.
HOOPER, L. V., LITTMAN, D. R. & MACPHERSON, A. J. 2012. Interactions between the microbiota and the immune system. *Science (New York, N.Y.),* 336**,** 1268-1273.
HUMPHREY, E. & CAUD, E. 2018. *Mechanisms underlying host-microbe interactions in pathophysiology of diseases,* books.google.co.uk, Springer Nature.
JAEGGI, T., KORTMAN, G. A. M., MORETTI, D., CHASSARD, C., HOLDING, P., DOSTAL, A., BOEKHORST, J., TIMMERMAN, H. M., SWINKELS, D. W., TJALSMA, H., NJENGA, J., MWANGI, A., KVALSVIG, J., LACROIX, C. & ZIMMERMANN, M. B. 2015. Iron fortification adversely affects the gut microbiome, increases pathogen abundance and induces intestinal inflammation in Kenyan infants. *Gut,* 64**,** 731.
JOBE, A. H. & BANCALARI, E. 2001. Bronchopulmonary Dysplasia. *American Journal Respiratory and Critical Care Medicine,* 163**,** 1723-1729.
JONULEIT, H., SCHMITT, E., STASSEN, M., TUETTENBERG, A., KNOP, J. & ENK, A. H. 2001. Identification and functional characterization of human CD4(+)CD25(+) T cells with regulatory properties isolated from peripheral blood. *J Exp Med,* 193**,** 1285-94.
JOOSTEN, K., EMBLETON, N., YAN, W., SENTERRE, T. & NUTRITION, E. E. E. C. W. G. O. P. P. 2018. ESPGHAN/ESPEN/ESPR/CSPEN guidelines on pediatric parenteral nutrition: Energy. *Clin Nutr,* 37**,** 2309-2314.
JUHL, S. M., YE, X., ZHOU, P., LI, Y., IYORE, E. O., ZHANG, L., JIANG, P., VAN GOUDOEVER, J. B., GREISEN, G. & SANGILD, P. T. 2018. Bovine Colostrum for Preterm Infants in the First Days of Life: A

Randomized Controlled Pilot Trial. *J Pediatr Gastroenterol Nutr,* 66**,** 471-478.
K., A., TANOUE, T., SHIMA, T., IMAOKA, A., KUWAHARA, T., MOMOSE, Y. & CHENG, G. 2011. Induction of Colonic Regulatory T Cells by Indigenous Clostridium Species - murine model of Treg and Th population being changed by microbiota. *Science*.
KAMADA, T., TOGASHI, Y., TAY, C., HA, D., SASAKI, A., NAKAMURA, Y., SATO, E., FUKUOKA, S., TADA, Y., TANAKA, A., MORIKAWA, H., KAWAZOE, A., KINOSHITA, T., SHITARA, K., SAKAGUCHI, S. & NISHIKAWA, H. 2019. PD-1(+) regulatory T cells amplified by PD-1 blockade promote hyperprogression of cancer. *Proc Natl Acad Sci U S A,* 116**,** 9999-10008.
KATZ, J., LEE, A. C. C., KOZUKI, N., LAWN, J. E., COUSENS, S., BLENCOWE, H., EZZATI, M., BHUTTA, Z. A., MARCHANT, T., WILLEY, B. A., ADAIR, L., BARROS, F., BAQUI, A. H., CHRISTIAN, P., FAWZI, W., GONZALEZ, R., HUMPHREY, J., HUYBREGTS, L., KOLSTEREN, P., MONGKOLCHATI, A., MULLANY, L. C., NDYOMUGYENYI, R., NIEN, J. K., OSRIN, D., ROBERFROID, D., SANIA, A., SCHMIEGELOW, C., SILVEIRA, M. F., TIELSCH, J., VAIDYA, A., VELAPHI, S. C., VICTORA, C. G., WATSON-JONES, D. & BLACK, R. E. 2013. Mortality risk in preterm and small-for-gestational-age infants in low-income and middle-income countries: a pooled country analysis. *The Lancet,* 382**,** 417-425.
KEAG, O. E., NORMAN, J. E. & STOCK, S. J. 2018. Long-term risks and benefits associated with cesarean delivery for mother, baby, and subsequent pregnancies: Systematic review and meta-analysis. *PLoS Med,* 15**,** e1002494.
KERKHOF, G. F., WILLEMSEN, R. H., LEUNISSEN, R. W., BREUKHOVEN, P. E. & HOKKEN-KOELEGA, A. C. 2012. Health profile of young adults born preterm: negative effects of rapid weight gain in early life. *J Clin Endocrinol Metab,* 97**,** 4498-506.
KINDER, J. M., JIANG, T. T., ERTELT, J. M., XIN, L., STRONG, B. S., SHAABAN, A. F. & WAY, S. S. 2015. Cross-Generational Reproductive Fitness Enforced by Microchimeric Maternal Cells. *Cell,* 162**,** 505-15.
KING, C. & DOGRA, H. 2011. Evaluation of the leicester incubator measure(lim) for obtaining length measurements in preterm babies. *Arch Dis Child Fetal Neonatal Ed*.
KORPELA, K., BLAKSTAD, E. W., MOLTU, S. J., STRØMMEN, K., NAKSTAD, B., RØNNESTAD, A. E., BRÆKKE, K., IVERSEN, P. O., DREVON, C. A. & DE VOS, W. 2018. Intestinal microbiota development and gestational age in preterm neonates. *Scientific Reports,* 8**,** 2453.
KOZICH, J. & SCHLOSS, P. D. 2014. 16S Sequencing with the Illumina MiSeq Personal Sequencer. *Library Generation, QC, Sequencing, and Data Quality Analysis Procedures.* University of Michigan: University of Michigan Health System.
KROVI, S. H. & GAPIN, L. 2018. Invariant Natural Killer T Cell Subsets-More Than Just Developmental Intermediates. *Front Immunol,* 9**,** 1393.
KRUEGER, A., ZIETARA, N. & LYSZKIEWICZ, M. 2017. T Cell Development by the Numbers. *Trends Immunol,* 38**,** 128-139.
KUBINAK, J. L., PETERSEN, C., STEPHENS, W. Z., SOTO, R., BAKE, E., O'CONNELL, R. M. & ROUND, J. L. 2015. MyD88 signaling in T cells

directs IgA-mediated control of the microbiota to promote health. *Cell host & microbe,* 17**,** 153-163.
KUNICKI, M. A., AMAYA HERNANDEZ, L. C., DAVIS, K. L., BACCHETTA, R. & RONCAROLO, M.-G. 2018. Identity and Diversity of Human Peripheral Th and T Regulatory Cells Defined by Single-Cell Mass Cytometry. *The Journal of Immunology,* 200**,** 336.
KURIOKA, A., USSHER, J. E., COSGROVE, C., CLOUGH, C., FERGUSSON, J. R., SMITH, K., KANG, Y. H., WALKER, L. J., HANSEN, T. H., WILLBERG, C. B. & KLENERMAN, P. 2015. MAIT cells are licensed through granzyme exchange to kill bacterially sensitized targets. *Mucosal Immunol,* 8**,** 429-40.
LADD, M., SHARMA, A., HUANG, Q., WANG ADELE, Y., XU, L., GENOWATI, I., LEVINGS MEGAN, K. & LAVOIE PASCAL, M. 2010. Natural killer T cells constitutively expressing the interleukin‐2 receptor α chain early in life are primed to respond to lower antigenic stimulation. *Immunology,* 131**,** 289-299.
LE BOURHIS, L., GUERRI, L., DUSSEAUX, M., MARTIN, E., SOUDAIS, C. & LANTZ, O. 2011. Mucosal-associated invariant T cells: unconventional development and function. *Trends Immunol,* 32**,** 212-8.
LEBLANC, J. G., MILANI, C., DE GIORI, G. S., SESMA, F., VAN SINDEREN, D. & VENTURA, M. 2013. Bacteria as vitamin suppliers to their host: a gut microbiota perspective. *Curr Opin Biotechnol,* 24**,** 160-8.
LEEANSYAH, E., LOH, L., NIXON, D. F. & SANDBERG, J. K. 2014. Acquisition of innate-like microbial reactivity in mucosal tissues during human fetal MAIT-cell development. *Nature Communications,* 5**,** 3143.
LEGRAND, D. 2016. Overview of Lactoferrin as a Natural Immune Modulator. *J Pediatr,* 173 Suppl**,** S10-5.
LI, J., LI, H., MAO, H., YU, M., FENG, T., YANG, F., FAN, Y., LU, Q., SHEN, C., YIN, Z., TU, W. & MAO, M. 2013. Vgamma9Vdelta2-T lymphocytes have impaired antiviral function in small-for-gestational-age and preterm neonates. *Cellular & Molecular Immunology,* 10**,** 253-60.
LI, N., VAN UNEN, V., ABDELAAL, T., GUO, N., KASATSKAYA, S. A., LADELL, K., MCLAREN, J. E., EGOROV, E. S., IZRAELSON, M., CHUVA DE SOUSA LOPES, S. M., HOLLT, T., BRITANOVA, O. V., EGGERMONT, J., DE MIRANDA, N., CHUDAKOV, D. M., PRICE, D. A., LELIEVELDT, B. P. F. & KONING, F. 2019. Memory CD4(+) T cells are generated in the human fetal intestine. *Nat Immunol,* 20**,** 301-312.
LIMA, H. K., WAGNER-GILLESPIE, M., PERRIN, M. T. & FOGLEMAN, A. D. 2017. Bacteria and Bioactivity in Holder Pasteurized and Shelf-Stable Human Milk Products. *Curr Dev Nutr,* 1**,** e001438.
LING, X., LINGLONG, P., WEIXIA, D. & HONG, W. 2016. Protective Effects of Bifidobacterium on Intestinal Barrier Function in LPS-Induced Enterocyte Barrier Injury of Caco-2 Monolayers and in a Rat NEC Model. *PLoS One,* 11**,** e0161635.
LIU, C. F., BRANDT, G. S., HOANG, Q. Q., NAUMOVA, N., LAZAREVIC, V., HWANG, E. S., DEKKER, J., GLIMCHER, L. H., RINGE, D. & PETSKO, G. A. 2016. Crystal structure of the DNA binding domain of the transcription factor T-bet suggests simultaneous recognition of distant genome sites. *Proc Natl Acad Sci U S A,* 113**,** E6572-E6581.
LUCAS, A. & COLE, T. J. 1990. Breast milk and neonatal necrotising enterocolitis. *Lancet,* 336**,** 1519-23.

LUCAS, A., GORE, S. M., COLE, T. J., BAMFORD, M. F., DOSSETOR, J. F., BARR, I., DICARLO, L., CORK, S. & LUCAS, P. J. 1984. Multicentre trial on feeding low birthweight infants: effects of diet on early growth. *Arch Dis Child,* 59**,** 722-30.

LUCAS, A. & HUDSON, G. J. 1984. Preterm milk as a source of protein for low birthweight infants. *Arch Dis Child,* 59**,** 831-836.

MA, L. J., WALTER, B., DEGUZMAN, A., MULLER, H. K. & WALKER, A. M. 2008. Trans-epithelial immune cell transfer during suckling modulates delayed-type hypersensitivity in recipients as a function of gender. *PLoS One,* 3**,** e3562.

MACK, D. R., MICHAIL, S., WEI, S., MCDOUGALL, L. & HOLLINGSWORTH, M. A. 1999. Probiotics inhibit enteropathogenic E. coli adherence in vitro by inducing intestinal mucin gene expression. *Am J Physiol,* 276**,** G941-50.

MACPHERSON, A. J., GATTO, D., SAINSBURY, E., HARRIMAN, G. R., HENGARTNER, H. & ZINKERNAGEL, R. M. 2000. A Primitive T Cell–Independent Mechanism of Intestinal Mucosal IgA Responses to Commensal Bacteria. *Science*.

MACPHERSON, A. J. & UHR, T. 2004. Induction of protective IgA by intestinal dendritic cells carrying commensal bacteria. *Science,* 303**,** 1662-5.

MANTIS, N. J., ROL, N. & CORTHESY, B. 2011. Secretory IgA's complex roles in immunity and mucosal homeostasis in the gut. *Mucosal Immunol,* 4**,** 603-11.

MANZONI, P., RINALDI, M., CATTANI, S., PUGNI, L., ROMEO, M. G., MESSNER, H., STOLFI, I., DECEMBRINO, L., LAFORGIA, N., VAGNARELLI, F., MEMO, L., BORDIGNON, L., SAIA, O. S., MAULE, M., GALLO, E., MOSTERT, M., MAGNANI, C., QUERCIA, M., BOLLANI, L., PEDICINO, R., RENZULLO, L., BETTA, P., MOSCA, F., FERRARI, F., MAGALDI, R., STRONATI, M., FARINA, D., ITALIAN TASK FORCE FOR THE, S. & PREVENTION OF NEONATAL FUNGAL INFECTIONS, I. S. O. N. 2009. Bovine lactoferrin supplementation for prevention of late-onset sepsis in very low-birth-weight neonates: a randomized trial. *JAMA,* 302**,** 1421-8.

MARA, M. A., GOOD, M. & WEITKAMP, J. H. 2018. Innate and adaptive immunity in necrotizing enterocolitis. *Semin Fetal Neonatal Med,* 23**,** 394-399.

MARCHANT, A., APPAY, V., VAN DER SANDE, M., DULPHY, N., LIESNARD, C., KIDD, M., KAYE, S., OJUOLA, O., GILLESPIE, G. M. A., VARGAS CUERO, A. L., CERUNDOLO, V., CALLAN, M., MCADAM, K. P. W. J., ROWLAND-JONES, S. L., DONNER, C., MCMICHAEL, A. J. & WHITTLE, H. 2003. Mature CD8+ T lymphocyte response to viral infection during fetal life. *Journal of Clinical Investigation,* 111**,** 1747-1755.

MCGOVERN, N., SHIN, A., LOW, G., LOW, D., DUAN, K., YAO, L. J., MSALLAM, R., LOW, I., SHADAN, N. B., SUMATOH, H. R., SOON, E., LUM, J., MOK, E., HUBERT, S., SEE, P., KUNXIANG, E. H., LEE, Y. H., JANELA, B., CHOOLANI, M., MATTAR, C. N. Z., FAN, Y., LIM, T. K. H., CHAN, D. K. H., TAN, K. K., TAM, J. K. C., SCHUSTER, C., ELBE-BURGER, A., WANG, X. N., BIGLEY, V., COLLIN, M., HANIFFA, M., SCHLITZER, A., POIDINGER, M., ALBANI, S., LARBI, A., NEWELL, E. W., CHAN, J. K. Y. & GINHOUX, F. 2017. Human fetal dendritic cells

promote prenatal T-cell immune suppression through arginase-2. *Nature,* 546**,** 662-666.
MCGUIRE, W. & ANTHONY, M. Y. 2003. Donor human milk versus formula for preventing necrotising enterocolitis in preterm infants: systematic review. *Arch Dis Child Fetal Neonatal Ed,* 88**,** F11-4.
MEINZEN-DERR, J., POINDEXTER, B., WRAGE, L., MORROW, A. L., STOLL, B. & DONOVAN, E. F. 2009. Role of human milk in extremely low birth weight infants' risk of necrotizing enterocolitis or death. *J Perinatol,* 29**,** 57-62.
MELVILLE, J. M. & MOSS, T. J. 2013. The immune consequences of preterm birth. *Front Neurosci,* 7**,** 79.
MENKES, J. H., CHERNICK, V. & RINGEL, B. 1966. Effect of elevated blood tyrosine on subsequent intellectual development of premature infants. *The Journal of Pediatrics,* 69**,** 583-588.
MEYERS, J. M., TAN, S., BELL, E. F., DUNCAN, A. F., GUILLET, R., STOLL, B. J., D'ANGIO, C. T., EUNICE KENNEDY SHRIVER NATIONAL INSTITUTE OF CHILD, H. & HUMAN DEVELOPMENT NEONATAL RESEARCH, N. 2019. Neurodevelopmental outcomes among extremely premature infants with linear growth restriction. *J Perinatol,* 39**,** 193-202.
MICHAELSSON, J., MOLD, J. E., MCCUNE, J. M. & NIXON, D. F. 2006. Regulation of T cell responses in the developing human fetus. *J Immunol,* 176**,** 5741-8.
MINCHEVA-NILSSON, L., HAMMARSTROM, S. & HAMMARSTROM, M. 1997. Activated Human yd T Lymphocytes Express Functional Lactoferrin Receptors. *Scand J Immunol,* 46**,** 609-618.
MIRAGAIA, R. J., GOMES, T., CHOMKA, A., JARDINE, L., RIEDEL, A., HEGAZY, A. N., WHIBLEY, N., TUCCI, A., CHEN, X., LINDEMAN, I., EMERTON, G., KRAUSGRUBER, T., SHIELDS, J., HANIFFA, M., POWRIE, F. & TEICHMANN, S. A. 2019. Single-Cell Transcriptomics of Regulatory T Cells Reveals Trajectories of Tissue Adaptation. *Immunity,* 50**,** 493-504 e7.
MONTGOMERY, R. K., MULBERG, A. E. & GRAND, R. J. 1999. Development of the human gastrointestinal tract: twenty years of progress. *Gastroenterology,* 116**,** 702-31.
MOOSSAVI, S., SEPEHRI, S., ROBERTSON, B., BODE, L., GORUK, S., FIELD, C. J., LIX, L. M., DE SOUZA, R. J., BECKER, A. B., MANDHANE, P. J., TURVEY, S. E., SUBBARAO, P., MORAES, T. J., LEFEBVRE, D. L., SEARS, M. R., KHAFIPOUR, E. & AZAD, M. B. 2019. Composition and Variation of the Human Milk Microbiota Are Influenced by Maternal and Early-Life Factors. *Cell Host Microbe,* 25**,** 324-335 e4.
MOREIRA-TEIXEIRA, L., RESENDE, M., COFFRE, M., DEVERGNE, O., HERBEUVAL, J. P., HERMINE, O., SCHNEIDER, E., ROGGE, L., RUEMMELE, F. M., DY, M., CORDEIRO-DA-SILVA, A. & LEITE-DE-MORAES, M. C. 2011. Proinflammatory environment dictates the IL-17-producing capacity of human invariant NKT cells. *J Immunol,* 186**,** 5758-65.
MORRIS, S. M., JR. 2010. Arginine: master and commander in innate immune responses. *Sci Signal,* 3**,** pe27.
MORROW, A. L., LAGOMARCINO, A. J., SCHIBLER, K. R., TAFT, D. H., YU, Z. & WANG, B. 2013. Early microbial and metabolomic signatures predict later onset of necrotizing enterocolitis in preterm infants. *Microbiome,* 1.

MOTULSKY, H. J. & BROWN, R. E. 2006. Detecting outliers when fitting data with nonlinear regression - a new method based on robust nonlinear regression and the false discovery rate. *BMC Bioinformatics,* 7**,** 123.
NEAL, M. D., JIA, H., EYER, B., GOOD, M., GUERRIERO, C. J., SODHI, C. P., AFRAZI, A., PRINDLE, T., JR., MA, C., BRANCA, M., OZOLEK, J., BRODSKY, J. L., WIPF, P. & HACKAM, D. J. 2013. Discovery and validation of a new class of small molecule Toll-like receptor 4 (TLR4) inhibitors. *PLoS One,* 8**,** e65779.
NIELSEN, S. M., HANSEN, G. H. & DANIELSEN, E. M. 2010. Lactoferrin targets T cells in the small intestine. *J Gastroenterol,* 45**,** 1121-8.
NOWICKA, M., KRIEG, C., CROWELL, H. L., WEBER, L. M., HARTMANN, F. J., GUGLIETTA, S., BECHER, B., LEVESQUE, M. P. & ROBINSON, M. D. 2017. CyTOF workflow: differential discovery in high-throughput high-dimensional cytometry datasets. *F1000Res,* 6**,** 748.
O'CONNOR, D. L., KISS, A., TOMLINSON, C., BANDO, N., BAYLISS, A., CAMPBELL, D. M., DANEMAN, A., FRANCIS, J., KOTSOPOULOS, K., SHAH, P. S., VAZ, S., WILLIAMS, B., UNGER, S. & OPTIMO, M. F. G. 2018. Nutrient enrichment of human milk with human and bovine milk-based fortifiers for infants born weighing <1250 g: a randomized clinical trial. *Am J Clin Nutr,* 108**,** 108-116.
O'CONNOR, D. L., GIBBINS, S., KISS, A. & ET AL. 2016. Effect of supplemental donor human milk compared with preterm formula on neurodevelopment of very low-birth-weight infants at 18 months: A randomized clinical trial. *JAMA,* 316**,** 1897-1905.
OLIN, A., HENCKEL, E., CHEN, Y., LAKSHMIKANTH, T., POU, C., MIKES, J., GUSTAFSSON, A., BERNHARDSSON, A. K., ZHANG, C., BOHLIN, K. & BRODIN, P. 2018. Stereotypic Immune System Development in Newborn Children. *Cell,* 174**,** 1277-1292 e14.
OLSEN, L. R., LEIPOLD, M. D., PEDERSEN, C. B. & MAECKER, H. T. 2019. The anatomy of single cell mass cytometry data. *Cytometry A,* 95**,** 156-172.
OLSZAK, T., AN, D., ZEISSIG, S., VERA, M. P., RICHTER, J., FRANKE, A., GLICKMAN, J. N., SIEBERT, R., BARON, R. M., KASPER, D. L. & BLUMBERG, R. S. 2012. Microbial Exposure During Early Life Has Persistent Effects on Natural Killer T Cell Function. *Science,* 336**,** 489.
ORAM, J. D. & REITER, B. 1968. Inhibition of bacteria by lactoferrin and other iron-chelating agents. *Biochim Biophys Acta,* 170**,** 351-65.
ORGANISATION, W. H. 2004. ICD-10: International statistical classification of disease and related health problems : tenth revision, 2nd edition. *World Health Organisation*.
PAGEL, J., HARTZ, A., FIGGE, J., GILLE, C., ESCHWEILER, S., PETERSEN, K., SCHREITER, L., HAMMER, J., KARSTEN, C. M., FRIEDRICH, D., HERTING, E., GOPEL, W., RUPP, J. & HARTEL, C. 2016. Regulatory T cell frequencies are increased in preterm infants with clinical early-onset sepsis. *Clinical & Experimental Immunology,* 185**,** 219-27.
PANDYA, J. M., LUNDELL, A. C., HALLSTROM, M., ANDERSSON, K., NORDSTROM, I. & RUDIN, A. 2016. Circulating T helper and T regulatory subsets in untreated early rheumatoid arthritis and healthy control subjects. *J Leukoc Biol,* 100**,** 823-833.
PARRA-LLORCA, A., GORMAZ, M., ALCÁNTARA, C., CERNADA, M., NUÑEZ-RAMIRO, A., VENTO, M. & COLLADO, M. C. 2018. Preterm Gut

Microbiome Depending on Feeding Type: Significance of Donor Human Milk. *Frontiers in Microbiology,* 9.

PATEL, A. L., JOHNSON, T. J., ENGSTROM, J. L., FOGG, L. F., JEGIER, B. J., BIGGER, H. R. & MEIER, P. P. 2013. Impact of early human milk on sepsis and health-care costs in very low birth weight infants. *Journal Of Perinatology,* 33**,** 514.

PEKALSKI, M. L., GARCIA, A. R., FERREIRA, R. C., RAINBOW, D. B., SMYTH, D. J., MASHAR, M., BRADY, J., SAVINYKH, N., DOPICO, X. C., MAHMOOD, S., DULEY, S., STEVENS, H. E., WALKER, N. M., CUTLER, A. J., WALDRON-LYNCH, F., DUNGER, D. B., SHANNON-LOWE, C., COLES, A. J., JONES, J. L., WALLACE, C., TODD, J. A. & WICKER, L. S. 2017. Neonatal and adult recent thymic emigrants produce IL-8 and express complement receptors CR1 and CR2. *JCI Insight,* 2.

PETERSEN, S. M., GREISEN, G. & KROGFELT, K. A. 2016. Nasogastric feeding tubes from a neonatal department yield high concentrations of potentially pathogenic bacteria- even 1 d after insertion. *Pediatr Res,* 80**,** 395-400.

PETSCHOW, B. W., TALBOTT, R. D. & BATEMA, R. P. 1999. Ability of lactoferrin to promote the growth of Bifidobacterium spp. in vitro is independent of receptor binding capacity and iron saturation level. *Journal of Medical Microbiology,* 48**,** 541-549.

PFISTER, K. M., ZHANG, L., MILLER, N. C., INGOLFSLAND, E. C., DEMERATH, E. W. & RAMEL, S. E. 2018. Early body composition changes are associated with neurodevelopmental and metabolic outcomes at 4 years of age in very preterm infants. *Pediatr Res,* 84**,** 713-718.

POLLOCK, J., GLENDINNING, L., WISEDCHANWET, T. & WATSON, M. 2018. The Madness of Microbiome: Attempting To Find Consensus "Best Practice" for 16S Microbiome Studies. *Appl Environ Microbiol,* 84.

PROVINE, N. M., BINDER, B., FITZPATRICK, M. E. B., SCHUCH, A., GARNER, L. C., WILLIAMSON, K. D., VAN WILGENBURG, B., THIMME, R., KLENERMAN, P. & HOFMANN, M. 2018. Unique and Common Features of Innate-Like Human Vdelta2(+) gammadeltaT Cells and Mucosal-Associated Invariant T Cells. *Front Immunol,* 9**,** 756.

PUCCIO, G., ALLIET, P., CAJOZZO, C., JANSSENS, E., CORSELLO, G., SPRENGER, N., WERNIMONT, S., EGLI, D., GOSONIU, L. & STEENHOUT, P. 2017. Effects of Infant Formula With Human Milk Oligosaccharides on Growth and Morbidity: A Randomized Multicenter Trial. *J Pediatr Gastroenterol Nutr,* 64**,** 624-631.

QUIGLEY, M., EMBLETON, N. D. & MCGUIRE, W. 2018. Formula versus donor breast milk for feeding preterm or low birth weight infants. *Cochrane Database Syst Rev,* 6**,** CD002971.

QUINELLO, C., SILVEIRA‐LESSA, A. L., CECCON, M. E. J. R., CIANCIARULLO, M. A., CARNEIRO‐SAMPAIO, M. & PALMEIRA, P. 2014. Phenotypic Differences in Leucocyte Populations among Healthy Preterm and Full‐Term Newborns. *Scandinavian Journal of Immunology,* 80**,** 57-70.

RAHMAN, A. H., TORDESILLAS, L. & BERIN, M. C. 2016. Heparin reduces nonspecific eosinophil staining artifacts in mass cytometry experiments. *Cytometry A,* 89**,** 601-7.

RAMAKRISHNA, C., KUJAWSKI, M., CHU, H., LI, L., MAZMANIAN, S. K. & CANTIN, E. M. 2019. Bacteroides fragilis polysaccharide A induces IL-10 secreting B and T cells that prevent viral encephalitis. *Nat Commun,* 10**,** 2153.

RAMEL, S. E., DEMERATH, E. W., GRAY, H. L., YOUNGE, N., BOYS, C. & GEORGIEFF, M. K. 2012. The relationship of poor linear growth velocity with neonatal illness and two-year neurodevelopment in preterm infants. *Neonatology,* 102**,** 19-24.

RAPHAEL, I., NALAWADE, S., EAGAR, T. N. & FORSTHUBER, T. G. 2015. T cell subsets and their signature cytokines in autoimmune and inflammatory diseases. *Cytokine,* 74**,** 5-17.

RCPCH 2017. National Neonatal Audit Programme 2016 and 2017 Audit Measures Overview of data derivation, data entry and outputs. RCPCH website.

REGEV, R. H., ARNON, S., LITMANOVITZ, I., BAUER-RUSEK, S., BOYKO, V., LERNER-GEVA, L., REICHMAN, B. & ISRAEL NEONATAL, N. 2016. Association between neonatal morbidities and head growth from birth until discharge in very-low-birthweight infants born preterm: a population-based study. *Dev Med Child Neurol,* 58**,** 1159-1166.

REMON, J. I., AMIN, S. C., MEHENDALE, S. R., RAO, R., LUCIANO, A. A., GARZON, S. A. & MAHESHWARI, A. 2015. Depth of bacterial invasion in resected intestinal tissue predicts mortality in surgical necrotizing enterocolitis. *J Perinatol,* 35**,** 755-62.

REN, S., HUI, Y., GOERICKE-PESCH, S., PANKRATOVA, S., KOT, W., PAN, X., THYMANN, T., SANGILD, P. T. & NGUYEN, D. N. 2019. Gut and immune effects of bioactive milk factors in preterm pigs exposed to prenatal inflammation. *Am J Physiol Gastrointest Liver Physiol,* 317**,** G67-G77.

RICE, D. N., HOUSTON, I. B., LYON, I. C. T., MACARTHUR, B. A., MULLINS, P. R., VEALE, A. M. O. & GUTHRIE, R. 1989. Transient neonatal tyrosinaemia. *Journal of Inherited Metabolic Disease,* 12**,** 13-22.

RIEZZO, G., INDRIO, F., RAIMONDI, F., MONTAGNA, O., SALVIA, G., MASSIMO, B., POLIMENO, L., CAVALLO, L. & FRANCAVILLA, R. 2009. Maturation of gastric electrical activity, gastric emptying and intestinal permeability in preterm newborns during the first month of life. *Ital J Pediatr,* 35**,** 6.

ROGNUM, T. O., THRANE, S., STOLTENBERG, L., VEGE, A. & BRANDTZAEG, P. 1992. Development of intestinal mucosal immunity in fetal life and the first postnatal months. *Pediatr Res,* 32**,** 145-9.

ROMERO, P., ZIPPELIUS, A., KURTH, I., PITTET, M. J., TOUVREY, C., IANCU, E. M., CORTHESY, P., DEVEVRE, E., SPEISER, D. E. & RUFER, N. 2007. Four functionally distinct populations of human effector-memory CD8+ T lymphocytes. *J Immunol,* 178**,** 4112-9.

ROSE, A. T. & PATEL, R. M. 2018. A critical analysis of risk factors for necrotizing enterocolitis. *Semin Fetal Neonatal Med,* 23**,** 374-379.

ROSENBLUM, M. D., WAY, S. S. & ABBAS, A. K. 2016. Regulatory T cell memory. *Nat Rev Immunol,* 16**,** 90-101.

ROUND, J. L., LEE, S. M., LI, J., TRAN, G., JABRI, B., CHATILA, T. A. & MAZMANIAN, S. K. 2011. The Toll-Like Receptor 2 Pathway Establishes Colonization by a Commensal of the Human Microbiota. *Science,* 332**,** 974.

ROZE, J. C., ANCEL, P. Y. & LEPAGE 2017. Nutritional strategies and gut microbiota composition as risk factors for necrotising enterocolitis in very-preterm infants.
RUAS-MADIEDO, P., GUEIMONDE, M., FERNANDEZ-GARCIA, M., DE LOS REYES-GAVILAN, C. G. & MARGOLLES, A. 2008. Mucin degradation by Bifidobacterium strains isolated from the human intestinal microbiota. *Appl Environ Microbiol,* 74**,** 1936-40.
RUHAAK, L. R., STROBLE, C., UNDERWOOD, M. A. & LEBRILLA, C. B. 2014. Detection of milk oligosaccharides in plasma of infants. *Anal Bioanal Chem,* 406**,** 5775-84.
SAARELA, T., KOKKONEN, J. & KOIVISTO, M. 2005. Macronutrient and energy contents of human milk fractions during the first six months of lactation. *Acta Paediatr,* 94**,** 1176-81.
SAKWINSKA, O., MOINE, D., DELLEY, M., COMBREMONT, S., REZZONICO, E., DESCOMBES, P., VINYES-PARES, G., ZHANG, Y., WANG, P. & THAKKAR, S. K. 2016. Microbiota in Breast Milk of Chinese Lactating Mothers. *PLoS One,* 11**,** e0160856.
SALLUSTO, F., LENIG, D., FORSTER, R., LIPP, M. & LANZAVECCHIA, A. 1999. Two subsets of memory T lymphocytes with distinct homing potentials and effector functions. *Nature,* 401**,** 708-12.
SANGILD, P. T., THYMANN, T., SCHMIDT, M., STOLL, B., BURRIN, D. G. & BUDDINGTON, R. K. 2013. Invited review: the preterm pig as a model in pediatric gastroenterology. *J Anim Sci,* 91**,** 4713-29.
SCHANLER, R. J., LAU, C., HURST, N. M. & SMITH, E. O. 2005. Randomized trial of donor human milk versus preterm formula as substitutes for mothers' own milk in the feeding of extremely premature infants. *Pediatrics,* 116**,** 400-6.
SCHEIBLE, K. M., EMO, J., LANIEWSKI, N., BARAN, A. M., PETERSON, D. R., HOLDEN-WILTSE, J., BANDYOPADHYAY, S., STRAW, A. G., HUYCK, H., ASHTON, J. M., TRIPI, K. S., ARUL, K., WERNER, E., SCALISE, T., MAFFETT, D., CASERTA, M., RYAN, R. M., REYNOLDS, A. M., REN, C. L., TOPHAM, D. J., MARIANI, T. J. & PRYHUBER, G. S. 2018. T cell developmental arrest in former premature infants increases risk of respiratory morbidity later in infancy. *JCI Insight,* 3.
SCHEIBLE, K. M., EMO, J., YANG, H., HOLDEN-WILTSE, J., STRAW, A., HUYCK, H., MISRA, S., TOPHAM, D. J., RYAN, R. M., REYNOLDS, A. M., MARIANI, T. J. & PRYHUBER, G. S. 2015. Developmentally determined reduction in CD31 during gestation is associated with CD8+ T cell effector differentiation in preterm infants. *Clin Immunol,* 161**,** 65-74.
SCHREURS, R., BAUMDICK, M. E., SAGEBIEL, A. F., KAUFMANN, M., MOKRY, M., KLARENBEEK, P. L., SCHALTENBERG, N., STEINERT, F. L., VAN RIJN, J. M., DREWNIAK, A., THE, S. M. L., BAKX, R., DERIKX, J. P. M., DE VRIES, N., CORPELEIJN, W. E., PALS, S. T., GAGLIANI, N., FRIESE, M. A., MIDDENDORP, S., NIEUWENHUIS, E. E. S., REINSHAGEN, K., GEIJTENBEEK, T. B. H., VAN GOUDOEVER, J. B. & BUNDERS, M. J. 2019. Human Fetal TNF-alpha-Cytokine-Producing CD4(+) Effector Memory T Cells Promote Intestinal Development and Mediate Inflammation Early in Life. *Immunity,* 50**,** 462-476 e8.
SENDER, R., FUCHS, S. & MILO, R. 2016. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol,* 14**,** e1002533.
SHAH, D. K. & ZUNIGA-PFLUCKER, J. C. 2014. An overview of the intrathymic intricacies of T cell development. *J Immunol,* 192**,** 4017-23.

SHEN, S., PRAME KUMAR, K., STANLEY, D., MOORE, R. J., VAN, T. T. H., WEN, S. W., HICKEY, M. J. & WONG, C. H. Y. 2018. Invariant Natural Killer T Cells Shape the Gut Microbiota and Regulate Neutrophil Recruitment and Function During Intestinal Inflammation. *Front Immunol,* 9**,** 999.

SHERMAN, M. P., ADAMKIN, D. H., NIKLAS, V., RADMACHER, P., SHERMAN, J., WERTHEIMER, F. & PETRAK, K. 2016. Randomized Controlled Trial of Talactoferrin Oral Solution in Preterm Infants. *J Pediatr,* 175**,** 68-73 e3.

SIM, K., SHAW, A. G., RANDELL, P., COX, M. J., MCCLURE, Z. E., LI, M. S., HADDAD, M., LANGFORD, P. R., COOKSON, W. O., MOFFATT, M. F. & KROLL, J. S. 2015. Dysbiosis anticipating necrotizing enterocolitis in very premature infants. *Clin Infect Dis,* 60**,** 389-97.

SISK, P. M., LOVELADY, C. A., DILLARD, R. G., GRUBER, K. J. & O'SHEA, T. M. 2007. Early human milk feeding is associated with a lower risk of necrotizing enterocolitis in very low birth weight infants. *Journal Of Perinatology,* 27**,** 428.

SITTIPO, P., LOBIONDA, S., CHOI, K., SARI, I. N., KWON, H. Y. & LEE, Y. K. 2018. Toll-Like Receptor 2-Mediated Suppression of Colorectal Cancer Pathogenesis by Polysaccharide A From Bacteroides fragilis. *Front Microbiol,* 9**,** 1588.

SORDILLO, J. E., ZHOU, Y., MCGEACHIE, M. J., ZINITI, J., LANGE, N., LARANJO, N., SAVAGE, J. R., CAREY, V., O'CONNOR, G., SANDEL, M., STRUNK, R., BACHARIER, L., ZEIGER, R., WEISS, S. T., WEINSTOCK, G., GOLD, D. R. & LITONJUA, A. A. 2017. Factors influencing the infant gut microbiome at age 3-6 months: Findings from the ethnically diverse Vitamin D Antenatal Asthma Reduction Trial (VDAART). *J Allergy Clin Immunol,* 139**,** 482-491 e14.

STEWART, C. J., AJAMI, N. J., O'BRIEN, J. L., HUTCHINSON, D. S., SMITH, D. P., WONG, M. C., ROSS, M. C., LLOYD, R. E., DODDAPANENI, H., METCALF, G. A., MUZNY, D., GIBBS, R. A., VATANEN, T., HUTTENHOWER, C., XAVIER, R. J., REWERS, M., HAGOPIAN, W., TOPPARI, J., ZIEGLER, A. G., SHE, J. X., AKOLKAR, B., LERNMARK, A., HYOTY, H., VEHIK, K., KRISCHER, J. P. & PETROSINO, J. F. 2018. Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature,* 562**,** 583-588.

STEWART, C. J., EMBLETON, N. D., CLEMENTS, E., LUNA, P. N., SMITH, D. P., FOFANOVA, T. Y., NELSON, A., TAYLOR, G., ORR, C. H., PETROSINO, J. F., BERRINGTON, J. E. & CUMMINGS, S. P. 2017a. Cesarean or Vaginal Birth Does Not Impact the Longitudinal Development of the Gut Microbiome in a Cohort of Exclusively Preterm Infants. *Frontiers in Microbiology,* 8**,** 1008.

STEWART, C. J., EMBLETON, N. D., MARRS, E. C. L., SMITH, D. P., FOFANOVA, T., NELSON, A., SKEATH, T., PERRY, J. D., PETROSINO, J. F., BERRINGTON, J. E. & CUMMINGS, S. P. 2017b. Longitudinal development of the gut microbiome and metabolome in preterm neonates with late onset sepsis and healthy controls. *Microbiome,* 5**,** 75.

STEWART, C. J., EMBLETON, N. D., MARRS, E. C. L., SMITH, D. P., NELSON, A., ABDULKADIR, B., SKEATH, T., PETROSINO, J. F., PERRY, J. D., BERRINGTON, J. E. & CUMMINGS, S. P. 2016.

Temporal bacterial and metabolic development of the preterm gut reveals specific signatures in health and disease. *Microbiome,* 4**,** 67.
STOLL, B. J., HANSEN, N. I., BELL, E. F., WALSH, M. C., CARLO, W. A., SHANKARAN, S., LAPTOOK, A. R., SANCHEZ, P. J., VAN MEURS, K. P., WYCKOFF, M., DAS, A., HALE, E. C., BALL, M. B., NEWMAN, N. S., SCHIBLER, K., POINDEXTER, B. B., KENNEDY, K. A., COTTEN, C. M., WATTERBERG, K. L., D'ANGIO, C. T., DEMAURO, S. B., TRUOG, W. E., DEVASKAR, U., HIGGINS, R. D., EUNICE KENNEDY SHRIVER NATIONAL INSTITUTE OF CHILD, H. & HUMAN DEVELOPMENT NEONATAL RESEARCH, N. 2015. Trends in Care Practices, Morbidity, and Mortality of Extremely Preterm Neonates, 1993-2012. *JAMA,* 314**,** 1039-51.
STOLTZ SJOSTROM, E., OHLUND, I., TORNEVI, A. & DOMELLOF, M. 2014. Intake and macronutrient content of human milk given to extremely preterm infants. *J Hum Lact,* 30**,** 442-9.
SULLIVAN, S., SCHANLER, R. J., KIM, J. H., PATEL, A. L., TRAWÖGER, R., KIECHL-KOHLENDORFER, U., CHAN, G. M., BLANCO, C. L., ABRAMS, S., COTTEN, C. M., LAROIA, N., EHRENKRANZ, R. A., DUDELL, G., CRISTOFALO, E. A., MEIER, P., LEE, M. L., RECHTMAN, D. J. & LUCAS, A. 2010. An Exclusively Human Milk-Based Diet Is Associated with a Lower Rate of Necrotizing Enterocolitis than a Diet of Human Milk and Bovine Milk-Based Products. *The Journal of Pediatrics,* 156**,** 562-567.e1.
SUMATOH, H. R., TENG, K. W., CHENG, Y. & NEWELL, E. W. 2017. Optimization of mass cytometry sample cryopreservation after staining. *Cytometry A,* 91**,** 48-61.
SWANN, J. B., ULDRICH, A. P., VAN DOMMELEN, S., SHARKEY, J., MURRAY, W. K., GODFREY, D. I. & SMYTH, M. J. 2009. Type I natural killer T cells suppress tumors caused by p53 loss in mice. *Blood,* 113**,** 6382-5.
SYKES, L., MACINTYRE, D. A., YAP, X. J., PONNAMPALAM, S., TEOH, T. G. & BENNETT, P. R. 2012. Changes in the Th1:Th2 cytokine bias in pregnancy and the effects of the anti-inflammatory cyclopentenone prostaglandin 15-deoxy-Delta(12,14)-prostaglandin J2. *Mediators Inflamm,* 2012**,** 416739.
TALSNESS, C. E., PENDERS, J., JANSEN, E., DAMOISEAUX, J., THIJS, C. & MOMMERS, M. 2017. Influence of vitamin D on key bacterial taxa in infant microbiota in the KOALA Birth Cohort Study. *PLoS One,* 12**,** e0188011.
TASTAN, C., KARHAN, E., ZHOU, W., FLEMING, E., VOIGT, A. Y., YAO, X., WANG, L., HORNE, M., PLACEK, L., KOZHAYA, L., OH, J. & UNUTMAZ, D. 2018. Tuning of human MAIT cell activation by commensal bacteria species and MR1-dependent T-cell presentation. *Mucosal Immunology,* 11**,** 1591-1605.
THOME, J. J., BICKHAM, K. L., OHMURA, Y., KUBOTA, M., MATSUOKA, N., GORDON, C., GRANOT, T., GRIESEMER, A., LERNER, H., KATO, T. & FARBER, D. L. 2016. Early-life compartmentalization of human T cell differentiation and regulatory function in mucosal and lymphoid tissues. *Nat Med,* 22**,** 72-7.
TREND, S., DE JONG, E., LLOYD, M. L., KOK, C. H., RICHMOND, P., DOHERTY, D. A., SIMMER, K., KAKULAS, F., STRUNK, T. & CURRIE,

A. 2015. Leukocyte Populations in Human Preterm and Term Breast Milk Identified by Multicolour Flow Cytometry. *PLoS One,* 10**,** e0135580.
UNDERWOOD, M. A., DAVIS, J. C. C., KALANETRA, K. M., GEHLOT, S., PATOLE, S., TANCREDI, D. J., MILLS, D. A., LEBRILLA, C. B. & SIMMER, K. 2017. Digestion of Human Milk Oligosaccharides by Bifidobacterium breve in the Premature Infant. *Journal of Pediatric Gastroenterology and Nutrition,* 65**,** 449-455.
UNDERWOOD, M. A., GAERLAN, S., DE LEOZ, M. L., DIMAPASOC, L., KALANETRA, K. M., LEMAY, D. G., GERMAN, J. B., MILLS, D. A. & LEBRILLA, C. B. 2015. Human milk oligosaccharides in premature infants: absorption, excretion, and influence on the intestinal microbiota. *Pediatric Research,* 78**,** 670-7.
URBANIAK, C., CUMMINS, J., BRACKSTONE, M., MACKLAIM, J. M., GLOOR, G. B., BABAN, C. K., SCOTT, L., O'HANLON, D. M., BURTON, J. P., FRANCIS, K. P., TANGNEY, M. & REID, G. 2014. Microbiota of human breast tissue. *Appl Environ Microbiol,* 80**,** 3007-14.
URUSHIYAMA, D., SUDA, W., OHNISHI, E., ARAKI, R., KIYOSHIMA, C., KURAKAZU, M., SANUI, A., YOTSUMOTO, F., MURATA, M., NABESHIMA, K., YASUNAGA, S. I., SAITO, S., NOMIYAMA, M., HATTORI, M., MIYAMOTO, S. & HATA, K. 2017. Microbiome profile of the amniotic fluid as a predictive biomarker of perinatal outcome. *Scientific Reports,* 7**,** 12171.
VANTOUROUT, P. & HAYDAY, A. 2013. Six-of-the-best: unique contributions of gammadelta T cells to immunology. *Nat Rev Immunol,* 13**,** 88-100.
VENTURA, V. & BROOKE, O. G. 1987. Plasma amino acids in small pretrm infants fed on human milk or formula. *Arch Dis Child,* 62**,** 1257-1264.
VICTORA, C. 2016. The Lancet Breastfeeding Series Group. Breastfeeding in the 21st century: epidemiology, mechanisms, and lifelong effect. *Lancet,* 387**,** 475-90.
VILLAR, J., GIULIANI, F., BHUTTA, Z. A., BERTINO, E., OHUMA, E. O., ISMAIL, L. C., BARROS, F. C., ALTMAN, D. G., VICTORA, C., NOBLE, J. A., GRAVETT, M. G., PURWAR, M., PANG, R., LAMBERT, A., PAPAGEORGHIOU, A. T., OCHIENG, R., JAFFER, Y. A. & KENNEDY, S. H. 2015. Postnatal growth standards for preterm infants: the Preterm Postnatal Follow-up Study of the INTERGROWTH-21 st Project. *The Lancet Global Health,* 3**,** e681-e691.
VILLAVICENCIO, A., RUEDA, M. S., TURIN, C. G. & OCHOA, T. J. 2017. Factors affecting lactoferrin concentration in human milk: how much do we know? *Biochem Cell Biol,* 95**,** 12-21.
VOGT, N. M., KERBY, R. L., DILL-MCFARLAND, K. A., HARDING, S. J., MERLUZZI, A. P., JOHNSON, S. C., CARLSSON, C. M., ASTHANA, S., ZETTERBERG, H., BLENNOW, K., BENDLIN, B. B. & REY, F. E. 2017. Gut microbiome alterations in Alzheimer's disease. *Sci Rep,* 7**,** 13537.
WALKER, J. C., SMOLDERS, M. A. J. C., GEMEN, E. F. A., ANTONIUS, T. A. J., LEUVENINK, J. & DE VRIES, E. 2010. Development of Lymphocyte Subpopulations in Preterm Infants. *Scandinavian Journal of Immunology,* 73**,** 53-58.
WALKER, L. J., THARMALINGAM, H. & KLENERMAN, P. 2014. The rise and fall of MAIT cells with age. *Scand J Immunol,* 80**,** 462-3.
WANDRO, S., OSBORNE, S., ENRIQUEZ, C., BIXBY, C., ARRIETA, A. & WHITESON, K. 2018. The Microbiome and Metabolome of Preterm

Infant Stool Are Personalized and Not Driven by Health Outcomes, Including Necrotizing Enterocolitis and Late-Onset Sepsis. *mSphere,* 3.
WANG, W.-P., IIGO, M., SATO, J., SEKINE, K., ADACHI, I. & TSUDA, H. 2000. Activation of Intestinal Mucosal Immunity in Tumour-bearing Mice by Lactoferrin. *Japan Journal of Cancer Research,* 91**,** 1022-1027.
WARD, P. & CONNEELY, O. M. 2004. Lactoferrin: Role in iron homeostasis and host defense against microbial infection. *BioMetals,* 17**,** 203-208.
WARNER, B. B., DEYCH, E., ZHOU, Y., HALL-MOORE, C., WEINSTOCK, G. M. & SODERGREN, E. 2016. Gut bacteria dysbiosis and necrotising enterocolitis in very low birthweight infants: a prospective case-control study. *Lancet,* 387.
WHITE, J. R., GONG, H., POPE, B., SCHLIEVERT, P. & MCELROY, S. J. 2017. Paneth-cell-disruption-induced necrotizing enterocolitis in mice requires live bacteria and occurs independently of TLR4 signaling. *Dis Model Mech,* 10**,** 727-736.
WILLIAMS, T., NAIR, H., SIMPSON, J. & EMBLETON, N. 2016. Use of Donor Human Milk and Maternal Breastfeeding Rates: A Systematic Review. *J Hum Lact,* 32**,** 212-20.
WILSON, A. D. 2018. Applications of Electronic-Nose Technologies for Noninvasive Early Detection of Plant, Animal and Human Diseases. *Chemosensors,* 6.
WOJCIK, K. Y., RECHTMAN, D. J., LEE, M. L., MONTOYA, A. & MEDO, E. T. 2009. Macronutrient analysis of a nationwide sample of donor breast milk. *J Am Diet Assoc,* 109**,** 137-40.
WONG, M. T., ONG, D. E., LIM, F. S., TENG, K. W., MCGOVERN, N., NARAYANAN, S., HO, W. Q., CERNY, D., TAN, H. K., ANICETE, R., TAN, B. K., LIM, T. K., CHAN, C. Y., CHEOW, P. C., LEE, S. Y., TAKANO, A., TAN, E. H., TAM, J. K., TAN, E. Y., CHAN, J. K., FINK, K., BERTOLETTI, A., GINHOUX, F., CUROTTO DE LAFAILLE, M. A. & NEWELL, E. W. 2016. A High-Dimensional Atlas of Human T Cell Diversity Reveals Tissue-Specific Trafficking and Cytokine Signatures. *Immunity,* 45**,** 442-56.
WOOD, A. J., RAYNES-GREENOW, C. H., CARBERRY, A. E. & JEFFERY, H. E. 2013. Neonatal length inaccuracies in clinical practice and related percentile discrepancies detected by a simple length-board. *J Paediatr Child Health,* 49**,** 199-203.
XIONG, W., BROWN, C. T., MOROWITZ, M. J., BANFIELD, J. F. & HETTICH, R. L. 2017. Genome-resolved metaproteomic characterization of preterm infant gut microbiota development reveals species-specific metabolic shifts and variabilities during early life. *Microbiome,* 5**,** 72.
YURCHENKO, E., LEVINGS, M. K. & PICCIRILLO, C. A. 2011. CD4+ Foxp3+ regulatory T cells suppress gammadelta T-cell effector functions in a model of T-cell-induced mucosal inflammation. *Eur J Immunol,* 41**,** 3455-66.
ZAHRAN, A. M., SAAD, K., ABDEL-RAHEEM, Y. F., ELSAYH, K. I., EL-HOUFEY, A. A., ABOUL-KHAIR, M. D. & ALBLIHED, M. A. 2019. Characterization of Regulatory T Cells in Preterm and Term Infants. *Archivum Immunologiae et Therapiae Experimentalis,* 67**,** 49-54.
ZALEWSKI, S., STEWART, C. J., EMBLETON, N. D. & BERRINGTON, J. E. 2018. Brief guide to the analysis, interpretation and presentation of microbiota data. *Arch Dis Child Educ Pract Ed,* 103**,** 327-330.

ZHANG, X., MOZELESKI, B., LEMOINE, S., DERIAUD, E., LIM, A., ZHIVAKI, D., AZRIA, E., LE RAY, C., ROGUET, G., LAUNAY, O., VANET, A., LECLERC, C. & LO-MAN, R. 2014. CD4 T cells with effector memory phenotype and function develop in the sterile environment of the fetus. *Sci Transl Med,* 6**,** 238ra72.

ZHANG, Z., ADELMAN, A. S., RAI, D., BOETTCHER, J. & LONNERDAL, B. 2013. Amino acid profiles in term and preterm human milk through lactation: a systematic review. *Nutrients,* 5**,** 4800-21.

ZHAO, Y., LIN, L., XIAO, Z., LI, M., WU, X., LI, W., LI, X., ZHAO, Q., WU, Y., ZHANG, H., YIN, J., ZHANG, L., CHO, C. H. & SHEN, J. 2018. Protective Role of gammadelta T Cells in Different Pathogen Infections and Its Potential Clinical Application. *J Immunol Res,* 2018**,** 5081634.

ZWITTINK, R. D., VAN ZOEREN-GROBBEN, D., MARTIN, R., VAN LINGEN, R. A., GROOT JEBBINK, L. J., BOEREN, S., RENES, I. B., VAN ELBURG, R. M., BELZER, C. & KNOL, J. 2017. Metaproteomics reveals functional differences in intestinal microbiota development of preterm infants. *Mol Cell Proteomics,* 16**,** 1610-1620.