LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Maternal nutrition, seasonality and epigenetics: an exploration of one-carbon metabolism and a novel nutritional supplement design in The Gambia.

Philip T James, MSc

Thesis submitted in accordance with the requirements for the degree of

Doctor of Philosophy

University of London

January 2019

Department of Population Health

Faculty of Epidemiology & Population Health

London School of Hygiene & Tropical Medicine

STATEMENT OF OWN WORK

I, Philip T James, confirm the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Philip T James

11th January 2019

ACKNOWLEDGEMENTS

ABSTRACT

A series of studies in rural Gambia established first-in-human evidence that peri-conceptional environment, including maternal methyl-donor nutritional status, predicts offspring DNA methylation. Methyl donor status is largely determined by the complex interlinking pathways of one-carbon metabolism. Women's methyl donor status ('methylation potential') is higher in the Gambian rainy season than in the dry season. Compared to infants conceived in the rainy season, those conceived in the dry season have decreased methylation at metastable epialleles (MEs). MEs are genomic loci whose methylation states are established in the very early embryo prior to gastrulation. Dry season conceptions are also associated with methylation patterns suggestive of a loss of imprinting at the *VTRNA2-1* gene, a putative tumour suppressor and regulator of innate immunity. The long-term goal is to design an intervention that shifts the maternal metabolome to improve regulation of the infant epigenome by providing micronutrients in the ratio and quantity necessary for optimal one-carbon metabolism all year round. To achieve this we first require a proof-of-concept trial showing that a nutritional supplement can address nutritional imbalances and increase methylation potential in non-pregnant women. This step provides the overall rationale for this thesis.

The research in this thesis contributes three key areas to existing knowledge of nutritional epigenetics in The Gambia. Firstly, it characterises seasonal differences in plasma nutritional biomarkers in a new dataset, validating the findings from previous studies and extending the consideration to new metabolites. Secondly, it explores how maternal nutritional predictors of infant DNA methylation may change between seasons, highlighting the importance of considering seasonality and underlying nutritional status in nutritional epigenetic research. Thirdly, it describes the design of a novel drink powder supplement, tailored to the target population by analysing negative nutritional predictors of homocysteine (designed to increase methylation potential). A clinical trial tested this new supplement alongside an existing multiple micronutrient tablet (UNIMMAP) amongst non-pregnant women in rural Gambia. Both interventions proved to be promising candidates for future epigenetic trials, acting on metabolic pathways to increase methylation potential.

Whilst the explicit targeting of DNA methylation profiles by nutritional interventions remains premature, accumulating evidence supports a focus on nutrition-sensitive epigenetic mechanisms as potential contributors to the developmental origins of health and disease.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AE	Adverse event
ANCOVA	Analysis of covariance
BHMT	Betaine homocysteine methyltransferase
BMI	Body mass index
CBS	Cystathionine-Beta-Synthase
CI	Confidence interval
CpG	Cytosine-phopshate-Guanine
CTH	Cystathionine gamma-lyase
CCV	Coefficient of cyclic variation
CV	Coefficient of variation
DHFR	Dihydrofolate reductase
DMR	Differentially methylated region
DOHaD	Developmental Origins of Health and Disease
DRI	Dietary reference intake
dTMP	Deoxythymidine monophosphate
dTTP	Deoxythymidine triphosphate
ENID	Early Nutrition & Immune Development
EWAS	Epigenome-wide association study
FAD	Flavin adenine dinucleotide
FT	Fourier terms
GNMT	Glycine N-Methyltransferase
GW	Gestational week
Hcy	Homocysteine
HPLC-MS	High performance liquid chromatography mass spectrometry
Lasso	Least absolute shrinkage and selection operator
LCSB	Luxembourg Centre for Systems Biomedicine
LNS	Lipid-based nutrient supplement
MAT	Methionine adenosyltransferase
MDEG-2	Methyl Donors and Epigenetics - 2
ME	Metastable epiallele
MMN	Multiple micronutrients
MRCG	Medical Research Council The Gambia Unit
MS	Methionine synthase
MT	Methyl transferases
MTHFD	Methylenetetrahydrofolate dehydrogenase
MTHFR	Methylenetetrahydrofolate reductase
PA	Pyridoxic acid
PC	Principal component
PL	Pyridoxal
PLP	Pyridoxal phosphate
RDA	Recommended daily allowance
SAH	S-adenosyl homocysteine

SAHH	S-adenosyl homocysteine hydrolase
SAM	S-adenosyl methionine
SCC	Scientific Coordination Committee
SHMT	Serine hydroxymethyltransferase
SoC	Season of conception
TE	Transposable element
TS	Thymidylate synthase
UNIMMAP	United Nations International Multiple Micronutrient Preparation
VA	Village Assistant
VMR	Variably methylated region
1C	One-carbon

GLOSSARY OF KEY TERMS

CpG site: Where a cytosine (5' end) is found next to a guanine (3' end) on a DNA strand along its linear sequence, hence 'cytosine-phosphate-guanine' or CpG site.

DNA methylation: The covalent bonding of a methyl (CH_3) group to a nucleotide, most often at CpG sites where the cytosine base is methylated at the 5' carbon position to form 5-methylcytosine.

Differentially Methylated Region: Regions of the genome exhibiting different methylation states between groups of interest (*i.e.* between intervention arms in a trial, between groups experiencing different exposures in cohort studies etc.).

Developmental Origins of Health and Disease: The hypothesis based on the work of David Barker that environmental exposures experienced early in life may influence future disease risk throughout the life course.

Epigenetics: Modifications to the genome that may affect gene expression without altering the underlying nucleotide sequence. Mechanisms include DNA methylation, histone modifications and RNA-mediated effects. Epigenetic marks are mitotically heritable and can be influenced by environmental factors.

Imprinted genes: A gene where only one allele is expressed and the other is silenced ('monoallelic expression'). They contain differentially methylated regions, where the silenced allele is generally hypermethylated. Imprinted genes are inherited and expressed in a parent-of-origin specific manner. If a gene is 'paternally expressed' it means the expressed allele comes from the father and the maternal allele is imprinted ('silenced').

Metastable Epialleles: These are CpG sites whose non-genetically-determined methylation state varies between individuals, but where variation is correlated across tissues originating from all germ layers in a single individual – indicating that the marks must have been laid down in the first few hours after conception

before cell types start to specialize.

Methylation potential: The extent to which methylation can occur e.g. through the provision of methyl donors and essential co-factors in one-carbon metabolic pathways.

One-carbon metabolism: The interlinking metabolic reactions of folate, homocysteine, methionine, choline, transsulfuration and transmethylation metabolic pathways. One-carbon units are activated and transferred to support numerous reactions, including those required for DNA synthesis, amino acid production and methylation.

Periconceptional period: Often defined as the window 14 weeks before conception to 10 weeks afterwards.

Chapter 1 Background to Thesis

_____

Summary of Chapter

In this chapter I briefly introduce the background to this thesis, summarising the work the research group I am attached to (the MRC International Nutrition Group) did prior to starting my thesis. I describe the rationale to my PhD, its aims and objectives and an overview of the structure of the PhD. I also state the extent to which I was involved in all aspects of the thesis, and provide details of the PhD timeline, publication outputs and funding received.

The technical background is deliberately brief as this is provided in much greater detail in Chapter 3. Further details on the research leading up to this thesis are also provided in Chapter 2.

1.1   Introduction

Why focus on epigenetics?

The field of epigenetics is rapidly generating a great deal of interest, exploring how our very molecular make-up in the form of modifications to the genome can be altered by factors as diverse as ageing, disease, nutrition, stress, alcohol and exposure to pollutants[1]. Epigenetic processes involve the interplay of DNA methylation, histone modifications and RNA-based mechanisms. They can result in changes to the genome that are mitotically heritable and can alter gene expression without changing the underlying DNA sequence[2].

The fusion of sperm and egg to create a new human being involves highly intricate epigenetic processes requiring the erasure of methylation marks to create pluripotent cells, and then the re-establishment of specific methylation patterns as cells differentiate during early embryonic development[3]. These processes use diverse metabolic pathways that are highly dependent upon the mother's nutrition to provide the necessary methyl groups and essential coenzymes for methylation[4]. The one-carbon metabolic pathways are most proximal to the DNA methylation reaction. Deficiencies in the key methyl donors of choline, betaine or folate, alongside deficiencies in vitamins B2, B6 and B12 that play essential co-enzyme roles, may disrupt the metabolic pathways that are responsible for DNA methylation. Errors in this epigenetic reprogramming at periconception cause a range of pathologies with a high burden to human health[5]. The most severe errors are embryonically lethal (causing infertility and later abortions) or can lead to severe developmental defects in the baby[6]. Errors in imprinted genes (genes with parent-of-origin-specific expression) can affect fetal growth and brain development with major implications for long-term health[7]. Finally there is evidence that early embryonic re-programming can capture information from the environment and 'tune' the offspring's metabolic settings to optimise its fitness; suggesting it as a key mediator of the developmental origins of health and disease (DOHaD) paradigm[8,9].

What has previous work from the MRC International Nutrition Group shown?

In the lead-up to this PhD the MRC International Nutrition Group published a series of studies from rural Gambia that established[10], and then replicated[11], first-in-human evidence that periconceptional environment, including maternal methyl-donor nutritional status, predicts offspring DNA methylation. This recapitulates evidence previously available in experimental

animals[12]. The group achieved this by combining a seasonal 'experiment of nature' (through which mothers consume different diets in the dry/harvest season compared to the wet/hungry season) with the study of so-called metastable epialleles (MEs). MEs are genomic regions in which methylation patterns are known to be established in the very early embryo, characterised by having a methylation state that varies between individuals, but where within an individual the variation is correlated across tissues originating from all germ layers[13]. This indicates that the marks have been laid down in the first few days after conception before cell types start to specialize and therefore makes these genomic loci useful for studying periconceptional exposures associated with offspring methylation[10,11]. The key learning from this previous body of work can be summarised as follows:

- Women's methyl donor status ('methylation potential') can be quantified by looking at plasma concentrations of one carbon biomarkers, such as folate, B2, B6, B12, betaine and choline. Higher concentrations of such methyl donors and coenzymes indicate a higher methylation potential.
- The profile of plasma biomarkers in women of rural Gambia suggests a higher methylation potential in the rainy season than in the dry season[14].
- Indeed, methylation levels at certain genomic loci (particularly MEs) were found to be higher in children conceived in the rainy season compared to those conceived in the dry season[11].
- Some of the genetic loci demonstrating differential methylation by season of conception are known to influence health:
    - The first example is the imprinted gene *VTRNA2-1;* a cancer suppressor gene involved also in immune defences against viral infections[15]. Imprinted genes should be hemi-methylated (*i.e.* the allele from one parent is 100% methylated and the other allele is unmethylated). The group instead found significant hypomethylation, suggestive of an excess of loss of imprinting, in dry season conceptions[15].
    - The group also showed that methylation of a critical region of the pro-opioid melanocortin gene (*POMC*) differs by season of conception. This locus is an ME. *POMC* methylation is associated with *POMC* expression and obesity traits in both children and adults[16,17].

In summary, the evidence to date suggests women conceiving in the dry season have a one-carbon metabolic profile that results in a lower maternal methylation potential. There is

preliminary evidence that this in turn affects methylation patterns in the fetal epigenome at loci implicated in a broad range of adverse birth and later life outcomes.

1.2   PhD Rationale

The observational data from my research group to date suggest that it may be possible to 'shift' the periconceptional methyl donor metabolome in such a way as to significantly alter methylation at multiple loci, including at key phenotype-related MEs and imprinted regions. My overall PhD goal was to design and implement a proof-of-concept trial assessing whether a nutritional supplement could target one-carbon pathways to increase methylation potential in non-pregnant women by correcting nutritional imbalances. If successful, the supplement would be a candidate to test in larger pregnancy trials.

Even without an explicit focus on epigenetics an understanding of seasonal one-carbon metabolome patterns and how to ensure women are not deficient in key related micronutrients is crucial for a wide range of health-related reasons[18]. These include improving fertility (both male and female) and enabling normal oocyte maturation and embryogenesis[19,20]; avoiding disruption of DNA synthesis and DNA stability[21,22]; and mitigating risk factors for cardiovascular disease[23,24], cancer[22,25], neurological disorders[26,27], birth defects[28] and potentially pre-eclampsia[29].

1.3   PhD Aim and Objectives

Overall Aim: To characterise Gambian seasonal variation in the maternal one-carbon metabolome and design a proof-of-concept pilot intervention to test whether a nutritional supplement can increase methylation potential in non-pregnant women.

Specific objectives:

a. Conduct a comprehensive literature review of human studies that investigate associations between maternal periconceptional nutritional exposures, the infant epigenome and later phenotypes.

b. Analyse an expanded set of maternal biomarkers (covering core one-carbon metabolites alongside new ones) in plasma obtained from the first antenatal visit of 350 women across the year in West Kiang, The Gambia. Use Fourier series modelling to characterize how the maternal one-carbon metabolome varies across the year. Are there obvious micronutrient deficiencies in metabolites essential for one-carbon metabolism? Are there times of the year when women appear more vulnerable to these deficiencies?

c. Interrogate existing data to explore whether the profile of maternal nutritional predictors of offspring methylation change between the rainy and dry season of The Gambia. This is to help answer the question whether the effect of a micronutrient supplement on DNA methylation would be consistent over the year.

d. Using three datasets, ascertain which nutritional components are consistently associated with increased methylation potential in women of rural Gambia.

e. Use the results from objectives b) to d) above to design a novel nutritional supplement to increase methylation potential and to decide which time of year is most suitable for a trial.

f. Conduct a randomized controlled trial to test the effect of the novel supplement on increasing methylation potential (measured by reducing homocysteine concentration) in 300 non-pregnant Gambian women supplemented daily over 12 weeks.

The overall PhD framework is portrayed in Figure 1.1, split into a planning and implementation phase. Part A comprises objectives a) to d); Phase 2 comprises objectives e) and f).

Part A: Preparation		Part B: Implementation	
• Analysis of maternal plasma biomarkers from a new cohort **(Chapter 4)** • Inclusion of expanded list of metabolites **(Chapters 2, 4)**	• Modelling of biomarker seasonal trends **(Chapter 4)** • Predictors of methylation by season **(Chapter 5)** • Predictors of methylation potential **(Chapter 6)**	• Final selection of supplement ingredients **(Chapter 7)** • Decision on final doses **(Chapter 7)** • Assess aspects of supplement safety **(Chapter 7)**	Randomised controlled trial in 300 non-pregnant women in rural Gambia 12 weeks of daily supplementation to increase methylation potential by reducing homocysteine **(Chapter 8)**

Figure 1.1 PhD Conceptual framework

1.4 PhD Outline

The thesis is presented in a research paper style with additional linking chapters, submitted in accordance with London School of Hygiene and Tropical Medicine regulations. The prepared and published articles are included without adaptation and there is therefore some repetition between chapters on the study context. This has been minimised where possible. Where research papers included supplementary material for publication these have been included either at the end of the relevant chapter, or in the Annex section. The summary of each chapter indicates where all the relevant material can be found.

The thesis contains nine chapters, which include two published papers and one prepared (unpublished) manuscript. A brief overview of each chapter follows.

Chapter 1: An overview of the thesis structure, my involvement, collaborators and study timeline.

Chapter 2: An introduction providing the context to the studies, detailing the locations, populations and datasets used.

Chapter 3: A published research paper containing a literature review of candidate genes linking maternal nutrient exposure to offspring phenotypes via DNA methylation.

Chapter 4: Here I explore how plasma nutritional metabolites in pregnant women co-vary

over the year. I employ longitudinal analysis to describe seasonal trends and to assess whether there are specific times of the year when women appear more vulnerable to micronutrient deficiencies.

Chapter 5: A published research paper describing whether the association of maternal nutrition exposures with offspring DNA methylation is consistent over the year. To do this I stratify existing data on maternal biomarkers and offspring DNA methylation by season of conception and assess whether the profile of maternal nutritional predictors of offspring methylation remains constant or changes between the Gambian seasons.

Chapter 6: I use three datasets to investigate which metabolites are associated with increasing methylation potential, estimated by the extent to which they decrease plasma homocysteine. I run multivariable linear regression models to assess which nutrients independently predict plasma homocysteine in the datasets. I select those that consistently demonstrate an inverse relationship with homocysteine as potential ingredients for a novel nutritional supplement.

Chapter 7: I describe how I use the analyses from previous chapters to design a novel drink powder supplement tailored for West Kiang region of rural Gambia. I assess the chosen ingredients against existing dosage safety information, amounts given in previous trials and relevant plasma nutrient levels to establish final supplement doses well below upper safety limits.

Chapter 8: A research paper (prepared for submission) describing a clinical trial in West Kiang region of Gambia. 298 women were recruited into the three-arm study comprising the novel drink powder arm, a commonly used multiple micronutrient supplement (UNIMMAP) arm, and a control arm. Women in the intervention arms took their supplement daily for 12 weeks. Baseline, midline and endline blood samples were obtained from all participants to assess the effectiveness of the interventions on lowering plasma homocysteine.

Chapter 9: A final discussion chapter to summarise the overall findings, strengths and weaknesses of the thesis by objective, and to reflect on future research priorities in the field of nutritional epigenetics.

1.5 Candidate's Involvement

The original idea for developing and testing a nutritional supplement as a future candidate for pregnancy trials with epigenetic outcomes came from Prof Andrew Prentice, Nutrition Theme Leader for MRC Unit The Gambia. He obtained the project funding and recruited me as a research assistant to develop and implement the aim. I was given free rein to develop and test the supplement as I saw fit, but had constant technical support and advice from Dr Matt Silver, my PhD supervisor, together with Prof Prentice.

For Chapter 3, the literature review, I was responsible for coordinating the collaborative effort on the paper and produced the first draft of the paper. I include a detailed description of the input of co-authors at the beginning of the chapter. For Chapter 4 I organised the sample preparation, shipment and analysis of the plasma samples. I contributed ideas for the expanded biomarker list and visited the laboratory at the University of British Columbia (UBC) to learn more about the analyses. The analyses were overseen by Roger Dyer at UBC, who developed the assays for the mass spectrometry and who worked with Professor Rajavel Elango for the amino acid panel. I performed the longitudinal analyses of the biomarkers. For Chapter 5, Dr Matt Silver had the original idea of exploring seasonal interactions in nutritional predictors of methylation and provided me with the methylation data. I developed the statistical approach and drafted the first version of the paper, which was commented on by all co-authors and edited accordingly. Chapter 6 was my own work after double-checking my strategy with my supervisor. For Chapter 7 I decided on the ingredients and doses for the supplement myself, and wrote up the product safety information. However, I ran the final idea by my advisory committee (Prof Andrew Prentice, Dr Matt Silver, Prof Patrick Stover, Dr Sophie Moore), who checked they were happy with the rationale for the composition. I set up and managed the clinical trial summarised in Chapter 8, with the help of a hard-working field team in Keneba. I conducted all analyses and wrote the first draft of the paper. Chapter 9 is my own work.

Whilst I was given overall ownership of the direction and implementation of the contents of this thesis, this was, necessarily, a collaborative approach. I provide the names and roles of everyone involved as collaborators and field team members in **Annex 1.1.**

1.6   PhD Publications and additional outputs

Published papers:

- **James P,** Sajjadi S, Tomar AS, Saffari A, Fall CHD, Prentice AM, Shrestha S, Issarapu P, Yadav DK, Kaur L, et al. Candidate genes linking maternal nutrient exposure to offspring health via DNA methylation: a review of existing evidence in humans with specific focus on one-carbon metabolism. *Int J Epidemiol.* 2018; dyy153. https://doi.org/10.1093/ije/dyy153 **(Chapter 3).**

- **Philip T James,** Paula Dominguez-Salas, Branwen J Hennig, Sophie E Moore, Andrew M Prentice and Matt J Silver. Maternal one-carbon metabolism and infant DNA methylation between contrasting seasonal environments: A case study from The Gambia. *Current Developments in Nutrition.* 2018; nzy082. https://doi.org/10.1093/cdn/nzy082 **(Chapter 5).**

Manuscript in preparation:

- **Philip T James,** Andrew M Prentice & Matt J Silver. Comparison of a novel drink powder and existing multiple micronutrient tablet in reducing plasma homocysteine in non-pregnant women in rural Gambia: a randomised controlled trial to identify future candidates for epigenetic trials. *Am. J. Clin. Nutr.* (targeted) **(Chapter 8).**

Textbook chapter:

- **James P,** Silver M, Prentice A. Epigenetics, Nutrition, and Infant Health. In: Karakochuk C, Whitfield K, Green T and Kraemer K eds. *The Biology of the First 1,000 Days.* 2017. p. 335–54. CRC Press. **(Annex 3.5)**

Magazine article:

- **James P,** Silver M, Prentice A. Epigenetics, Nutrition, and Infant Health. *Sight and Life.* 2015; 29(1): 35-38. **(Annex 3.6)**

Poster:

- Maternal One-Carbon Biomarkers and the Infant Methylome: The Role of Seasonality Using a Case Study from The Gambia. *FASEB.* 2017; 31 (Supplement 1): Abstract 644.14. Presented at Epigenomics of Common Disease, October 2017, Wellcome Genome Campus, UK. (Developed into **Chapter 5)**.

Invited conference talks:

- 'Can what a mother eats at the time of conception influence the epigenome of her child?'
  - DOHaD 9th World Congress, November 2015, Cape Town, South Africa.
  - Rank Prize Funds Symposium on Preconception Nutrition and Lifelong Health, February 2016, Grasmere, UK.
  - Society for Nutrition, Education and Health Action (SNEHA) Conference, March 2016, Hyderabad, India.

- 'Next generation supplement design', Micronutrient Forum Global Conference, October 2016. Sight and Life Elevator Pitch Finalist, Cancun, Mexico.

- 'Epigenetics, Nutrition and Infant Health', IUNS 21st International Congress of Nutrition, October 2017, Buenos Aires, Argentina.

There were also three papers I was involved with as part of my PhD training as a supervisor to MSc Students and lecturer on the Nutrition in Emergencies module at LSHTM. These are unrelated to the PhD topic.

- Natasha Phillipa O'Sullivan, Natasha Lelijveld, Natasha Lelijveld, Alexandra Rutishauser-Perera and **Philip T James**. Follow-up between 6 and 24 months after discharge from treatment for severe acute malnutrition in children aged 6-59 months: A systematic review. *PLoS ONE.* 2018; 13(8):e0202053.

- Lucas K, **James P,** Choh AC, Lee M, Czerwinski SA, Demerath EW, Johnson W. The positive association of infant weight gain with adulthood body mass index has strengthened over time in the Fels Longitudinal Study. *Pediatr Obes*. 2018; 13:476–84.

- Adnaan Ghanachi, **Philip T James,** Carla Cerami. Guts, Germs and Iron: A Systematic Review of the Effect of Iron Supplementation and Fortification on Diarrhoea in Children Aged 4 to 59 Months. *Current Developments in Nutrition.* 2018 (submitted).

1.7   PhD Timeframe

The PhD took place from October 2015 to October 2018. An overview of the timeframe of key activities is provided in **Annex 1.2.**

1.8   Funding

1.9   References

1.   Perera F & Herbstman J (2011) Prenatal environmental exposures, epigenetics, and disease. *Reprod. Toxicol.* **31**, 363–73.

2.   Jaenisch R & Bird A (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.* **33 Suppl**, 245–254.

3.   Smallwood SA & Kelsey G (2012) De novo DNA methylation: a germ cell perspective. *Trends Genet.* **28**, 33–42.

4.   Fox JT & Stover PJ (2008) Folate-mediated one-carbon metabolism. *Vitam. Horm.* **79**, 1–44.

5.   Kalish JM, Jiang C & Bartolomei MS (2014) Epigenetics and imprinting in human disease. *Int. J. Dev. Biol.* **58**, 291–8.

6.   Schaefer CB, Ooi SKT, Bestor TH, et al. (2007) Epigenetic decisions in mammalian germ cells. *Science* **316**, 398–9.

7.   Moore GE, Ishida M, Demetriou C, et al. (2015) The role and interaction of imprinted genes in human fetal growth. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, 20140074–20140074.

8.   van Dijk SJ, Tellam RL, Morrison JL, et al. (2015) Recent developments on the role of epigenetics in obesity and metabolic disease. *Clin. Epigenetics* **7**, 66.

9.   Burdge GC & Lillycrop KA (2010) Nutrition, Epigenetics, and Developmental Plasticity: Implications for Understanding Human Disease. *Annu. Rev. Nutr.* **30**, 315–339.

10.   Waterland RA, Kellermayer R, Laritsky E, et al. (2010) Season of conception in rural gambia affects DNA methylation at putative human metastable epialleles. *PLoS Genet.* **6**, e1001252.

11.   Dominguez-Salas P, Moore SE, Baker MS, et al. (2014) Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. *Nat. Commun.* **5**, 3746.

12.   Waterland RA & Jirtle RL (2003) Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Mol. Cell. Biol.* **23**, 5293–300.

13.   Rakyan VK, Blewitt ME, Druker R, et al. (2002) Metastable epialleles in mammals. *Trends Genet.* **18**, 348–51.

14.   Dominguez-Salas P, Moore SE, Cole D, et al. (2013) DNA methylation potential: dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women. *Am. J. Clin. Nutr.* **97**, 1217–27.

15.   Silver MJ, Kessler NJ, Hennig BJ, et al. (2015) Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. *Genome Biol.* **16**, 118.

16.   Kühnen P, Handke D, Waterland RA, et al. (2016) Interindividual Variation in DNA Methylation at a Putative POMC Metastable Epiallele Is Associated with Obesity. *Cell Metab.* **24**, 502–9.

17.   Kuehnen P, Mischke M, Wiegand S, et al. (2012) An Alu element-associated hypermethylation variant of the POMC gene is associated with childhood obesity. *PLoS Genet.* **8**, e1002543.

18.   De-Regil LM, Harding KB & Roche ML (2016) Preconceptional Nutrition Interventions for Adolescent Girls and Adult Women: Global Guidelines and Gaps in Evidence and Policy with Emphasis on Micronutrients. *J. Nutr.* **146:**1461S-70S

19.   Steegers-Theunissen RPM, Twigt J, Pestinger V, et al. (2013) The periconceptional period, reproduction and long-term health of offspring: the importance of one-carbon metabolism. *Hum. Reprod. Update* **19**, 640–55.

20.   Zeisel SH (2006) Choline: critical role during fetal development and dietary requirements in adults. *Annu. Rev. Nutr.* **26**, 229–50.

21.   Martiniova L, Field MS, Finkelstein JL, et al. (2015) Maternal dietary uridine causes, and deoxyuridine prevents, neural tube closure defects in a mouse model of folate-

responsive neural tube defects. *Am. J. Clin. Nutr.* **101**, 860–9.

22. Ciappio ED, Mason JB & Crott JW (2011) Maternal one-carbon nutrient intake and cancer risk in offspring. *Nutr. Rev.* **69**, 561–71.

23. Selhub J (1999) Homocysteine metabolism. *Annu. Rev. Nutr.* **19**, 217–46.

24. Homocysteine Studies Collaboration (2002) Homocysteine and risk of ischemic heart disease and stroke: a meta-analysis. *JAMA* **288**, 2015–22.

25. Agnoli C, Grioni S, Krogh V, et al. (2016) Plasma Riboflavin and Vitamin B-6, but Not Homocysteine, Folate, or Vitamin B-12, Are Inversely Associated with Breast Cancer Risk in the European Prospective Investigation into Cancer and Nutrition-Varese Cohort. *J. Nutr.* **146**, 1227–34.

26. Lopes da Silva S, Vellas B, Elemans S, et al. (2014) Plasma nutrient status of patients with Alzheimer's disease: Systematic review and meta-analysis. *Alzheimers. Dement.* **10**, 485–502.

27. Bailey LB, Stover PJ, McNulty H, et al. (2015) Biomarkers of Nutrition for Development-Folate Review. *J. Nutr.* **145**, 1636S–1680S.

28. Molloy AM, Kirke PN, Brody LC, et al. (2008) Effects of folate and vitamin B12 deficiencies during pregnancy on fetal, infant, and child development. *Food Nutr. Bull.* **29**, S101-11.

29. Singh MD, Thomas P, Owens J, et al. (2015) Potential role of folate in pre-eclampsia. *Nutr. Rev.* **73**, 694–722.

Chapter 2 PhD setting and datasets used

_____

Summary of Chapter

This chapter provides a brief overview of the study setting. It provides the geographical context of the West Kiang region of The Gambia, where the datasets I use in this thesis originate, as well as where I conducted my fieldwork for the clinical trial described in Chapter 8.

It describes the detailed methodology for three datasets I refer to throughout the thesis. Two of the datasets were already established when I started the thesis, and one I established for this PhD.

2.1 PhD Setting: West Kiang, The Gambia

The Republic of The Gambia in West Africa is the smallest African nation. It is surrounded by Senegal and has a population of 1.9 million people with a growth rate of 3.3% per annum[1]. Data for this thesis comes from villages of West Kiang district of the Lower River Division, where MRC Keneba is located (Figure 2.1). This is an area of rural subsistence farming, predominantly made up of the sedentary Mandinka ethnic group. The Gambia experiences a rainy season from July to October (Figure 2.2a). This is a planting season with increased energy expenditure associated with agricultural work, and increased incidence of malarial and diarrhoeal diseases. Food stores from the harvest period run low in this season, hence it is also referred to as the 'hungry' season. The cooler dry season (Figure 2.2b) covers the other months, with harvesting occurring particularly in February to April, leading to improved food security[2,3]. Despite overall household food stores being more replete in the dry season, it is the rainy season that is associated with higher plasma levels of certain micronutrients central to one-carbon metabolism[4]. The underlying reasons are yet to be confirmed, but hypotheses include increased availability of green leaves during the rains or the scarcity of staples forcing people to increase food diversity (e.g. through seeking bush foods).
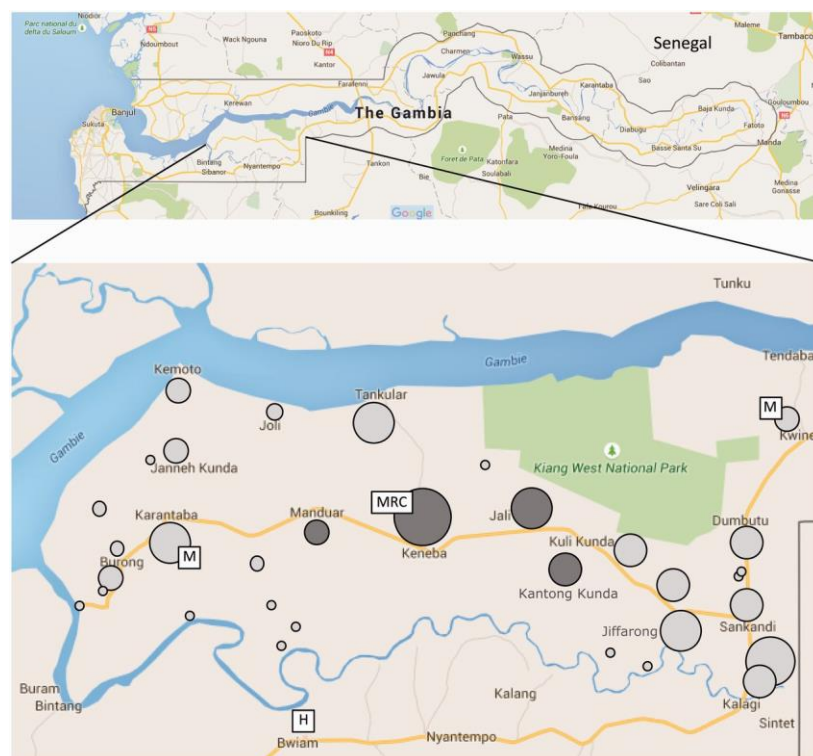
Figure 2.1 Location of MRC Keneba within West Kiang, The Gambia. Village circle sizes proportional to population.

Source: Hennig BJ *et al.* Int J Epidemiol. 2017; 46(2): e13[5]. Reproduced with permission.
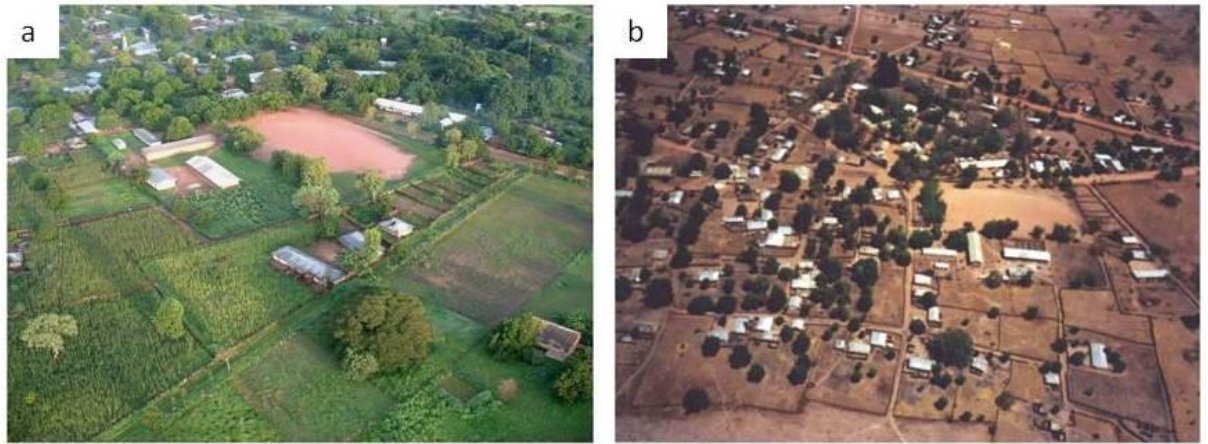
Figure 2.2 The Gambian rainy (a) and dry (b) season in Keneba, West Kiang. Photo credit: Andrew Prentice.

2.2 Datasets used in thesis

Three datasets are referred to in this thesis. Two are from the original Methyl Donors and Epigenetics (MDEG) study, the 'indicator group' and the 'main study group' led by Paula Dominguez-Salas from 2009-2012[4,6,7]. These were established prior to me starting the PhD but I use them for additional analyses. The third, termed 'MDEG-2', forms a new dataset for this thesis and utilises stored samples from the Early Nutrition & Immune Development (ENID) Trial[8]. I was involved in the sample selection and preparation, defined the biomarkers of interest and received training at the University of British Columbia laboratory on the principles of mass spectrometry used in many of the assays.

Table 2.1 summarises the key characteristics of each dataset and where they have been used in this thesis. Fuller details on each dataset are provided below, along with a summary of the key findings generated from each.

2.2.1 MDEG Indicator group

This was an observational study that took place in West Kiang region between July 2009 and June 2010, and written up in Dominguez-Salas *et al.* (2013)[4]. The overall aim was to document the dietary intake and plasma nutrition concentrations of a cohort of non-pregnant women, with a focus on one-carbon related metabolites. This dataset therefore provided a reference for the seasonal variation in these metabolites over a year, hence referred to as the 'indicator group'.

Non-pregnant women between the ages of 18 and 45 years (mean age 31 years) were followed monthly for a year. Each month they provided a fasted 10 mL blood sample and

were observed by field workers for collection of a 48-hour weighed dietary intake. This thesis utilises the blood sample data. Blood samples were collected in the field into EDTA monovettes, transported on ice and fully processed within two hours at the MRC laboratory in Keneba. Samples were spun for 10 minutes at 2,750g and the plasma was removed, aliquoted and immediately stored at -80°C. A sample of the remaining red blood cells were removed, washed and also stored at -80°C. The original list of maternal one-carbon biomarkers analysed were folate, B12, holotranscobalamin (active B12), choline, betaine, dimethylglycine (DMG), methionine, s-adenosyl methionine (SAM), s-adenosyl homocysteine (SAH), homocysteine (Hcy), riboflavin (B2), cysteine, 4-pyridoxic acid (PA), pyridoxal (PL) and pyridoxal 5'-phosphate (PLP). Throughout this thesis this collection of biomarkers is termed the **'core one-carbon biomarker'** set.

Plasma samples were shipped to the Department of Pediatrics, University of British Columbia, Canada for analysis of the following one-carbon metabolites by liquid chromatography-tandem mass spectrometry: SAM, SAH, free choline, betaine, DMG, Hcy, methionine, cysteine and the B6 vitamers (PA, PL, PLP). Plasma B12, active B12 and folate were analysed at UBC using an AxSyM analyser (Abbot laboratories, Chicago, IL). Riboflavin (B2) concentrations were analysed in the washed red blood cells by erythrocyte glutathione reductase activation coefficient (EGRAC) assay at MRC Human Nutrition Research laboratories, Cambridge, UK.

A sample size of 30 women per month was planned for, and to account for drop-out and those being excluded after becoming pregnant 62 women were recruited into the study. Overall between 22 and 30 blood draws were obtained each month throughout the year.

Key findings from the Indicator group[4]:

- Over the year there were notable seasonal fluctuations in plasma concentrations of one-carbon metabolites, particularly, folate, choline, betaine, DMG, homocysteine and B6.
- In the peak rainy season (July-October) compared to the peak dry season (February – April) there were higher concentrations of plasma folate, B2, betaine and SAM:SAH, and lower concentrations of B12, DMG, B6 and Hcy.
- Overall this equated to a higher methylation potential in the rainy season.
- For many metabolites dietary intake measurements demonstrated poor stability from one month to the next after adjusting for seasonality, and did not correlate well

with plasma concentrations. The recommendation was to use plasma concentrations going forwards.

This thesis uses the Indicator group dataset to help identify nutritional predictors of homocysteine (Chapter 6), and to validate the seasonal trend findings described in Chapter 4.

2.2.2 MDEG main study group

This dataset contains 167 mother-child pairs enrolled into the main MDEG cohort. Women of reproductive age (18-45 years) in West Kiang district were invited to participate and followed monthly. Those women conceiving in the peak of the rainy season (July-September 2009) and the peak of the dry season (February – April 2010) were enrolled and continued to be followed monthly until delivery. Their offspring were then followed throughout infancy. Detailed methodology and the main study results have been previously published in Dominguez-Salas *et al.* (2014)[7]. Women provided a 10mL fasting venous blood sample at the point they reported their first missed menses (mean (SD) 8.6 ± 4 weeks gestation). The same set of core one-carbon biomarkers were analysed as described in the Indicator group, using the same laboratory analyses. The seasonal trends obtained from the Indicator group were then used to back-extrapolate the plasma concentrations to the time of conception. Infant DNA was obtained from a 3mL venepuncture taken between 2-8 months after delivery. In this main group study methylation at 6 MEs in the infant DNA were analysed. More recently two other published papers have also utilised the offspring genome-wide DNA methylation data, measured using the Illumina Infinium HumanMethytlation450 array ('450k array')[9,10].

Key findings from the main study group[7]:

- Similar to findings in the Indicator group, the profile of plasma biomarkers back-extrapolated to the time of conception showed a higher methylation potential in the peak rainy season compared to the peak dry season.

- Mean methylation across the 6 MEs in the infant DNA was higher in the rainy season compared to the dry season, corresponding to the higher plasma methylation potential.

- Periconceptional plasma B2 concentration was positively associated with infant DNA methylation whilst Hcy, B6 and cysteine were inversely associated.

In this thesis the MDEG main study dataset is used to look at seasonal interactions in the

association between maternal nutritional predictors and infant DNA methylation in Chapter 5. It is also used to help identify nutritional predictors of homocysteine (Chapter 6) and to validate the seasonal trend findings described in Chapter 4.

2.2.3 MDEG-2

This dataset was established to generate information on an expanded set of biomarkers compared to the previous two. Unlike the MDEG main study, which only had data from six months of the year, the MDEG-2 dataset was designed to capture year-round variation in maternal plasma biomarkers and to be able to associate these with offspring DNA at a later date.

The dataset uses banked samples from the Early Nutrition & Immune Development (ENID) Trial (ISRCTN49285450) testing the effect of different nutritional supplements given to pregnant women on the immune development of their children. The trial provides a detailed bank of samples from multiple time-points across pregnancy and in infancy until 24 months of age. A subsample of 350 mother-child pairs from the ENID trial forms this MDEG-2 dataset. The ENID trial protocol has been previously described in detail[8]. In brief, women were randomised to four intervention groups after a positive pregnancy test at their 'booking' visit (the first clinic visit when pregnancy was confirmed, at approximately 13 weeks gestation): iron-folate (Fe-Fol) tablets, multiple micronutrient (MMN) tablets, a lipid-based nutritional supplement (LNS) fortified with Fe-Fol and LNS fortified with MMN. Supplementation was taken daily from the booking visit until delivery. At the time of booking mothers provided a fasted sample of 10 mL venous blood; 7.5mL into lithium heparin monovettes and 2.6mL into an EDTA monovette. Samples were immediately placed on ice after the venepuncture and taken to the laboratory for processing within one hour. The monovettes were centrifuged at 1800 RCFs for ten minutes at 4°C and the plasma drawn off into 2mL microtubes. The microtubes were immediately frozen at -70°C. The plasma samples experienced two freeze-thaw cycles (one to sub-aliquot and one at the point of analysis). Since maternal booking blood samples were taken before the interventions were started the trial design does not affect interpretation of the maternal plasma nutritional biomarkers. The first booking visit took place in January 2010 and final child was born in February 2014.

In order to understand how the metabolome changes across the year 350 samples were purposively selected to represent an even distribution by month of booking. Within each month the women with the earliest gestational age at booking (assessed by ultrasound) were

selected to capture the metabolome closest to the periconceptional period.

Chapter 4 details the plasma metabolites selected for analysis as part of the MDEG-2 dataset and their laboratory analyses. They include all the core one-carbon biomarkers looked at in the indicator group and MDEG main study, but also consider additional ones such as amino acids and inflammatory markers. As such the MDEG-2 set of metabolites are the termed the **'expanded biomarker'** collection.

The MDEG-2 dataset is used for the seasonal trends analyses of the expanded biomarker collection in Chapter 4, and also to investigate nutritional predictors of homocysteine in Chapter 6.

Table 2.1 Summary of dataset characteristics used in thesis

Dataset	Study design	Total N	Dry Season N (February-April)	Rainy Season N (July –September)	Biomarkers measured	Referred to in thesis
MDEG Indicator group	Non-pregnant women, ~30 followed for one year, with repeated monthly blood samples[2].	48 women with 288 total observations	34 women with 79 observations	28 women with 63 observations	**Core group:** Hcy, B2, PLP (B6), B12, folate, methionine, choline, betaine, DMG, cysteine	**Overview given:** Section 2.2.1 **Also referred to:** Chapters 4, 6
MDEG Main Study	Cross-sectional data comprising a baseline blood sample from pregnant women at first antenatal visit. Samples are taken from the peak of rainy and peak of dry season[4], spanning 6 months of the year.	167	83	84	**Core group**, as above	**Overview given:** Section 2.2.2 **Also referred to:** Chapters 4, 5, 6
MDEG-2	Cross sectional data comprising a baseline blood sample from pregnant women at first antenatal visit. Samples selected from each month of year.	350	87	100	**Expanded group:** Core group plus AGP, Asp, Thr, Ser, Glu, Gly, Ala, Val, Ile, Leu, Tyr, Phe, Lys, His, Arg, Pro	**Overview given:** Section 2.2.3 **Also referred to:** Chapters 4, 6

Abbreviations: AGP, α-1 acid glycoprotein; Ala, alanine; Arg, arginine; Asp, aspartic acid; DMG, dimethylglycine; Glu, glutamic acid; Gly, glycine; Hcy, homocysteine; His, histidine; Ile, isoleucine; Leu, leucine; Lys, lysine; MDEG, Methyl Donors and Epigenetics; Phe, phenylalanine; PLP, pyridoxal 5'-phosphate; Pro, proline; Ser, serine; Thr, threonine; Tyr, tyrosine; Val, valine.

2.3   References

1.   The Gambia Bureau of Statistics (2013) *The Gambia 2013 Population and Housing Census Preliminary Results*. Banjul, The Gambia.

2.   Prentice A, Whitehead R, Roberts S, et al. (1981) Long-term energy balance in child-bearing Gambian women. *Am J Clin Nutr* **34**, 2790–2799.

3.   Moore SE, Cole TJ, Collinson AC, et al. (1999) Prenatal or early postnatal events predict infectious deaths in young adulthood in rural Africa. *Int. J. Epidemiol.* **28**, 1088–95.

4.   Dominguez-Salas P, Moore SE, Cole D, et al. (2013) DNA methylation potential: dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women. *Am. J. Clin. Nutr.* **97**, 1217–27.

5.   Hennig BJ, Unger SA, Dondeh BL, et al. (2017) Cohort Profile: The Kiang West Longitudinal Population Study (KWLPS)--a platform for integrated research and health care provision in rural Gambia. *Int. J. Epidemiol.* **46**(2): e13.

6.   Dominguez-Salas P (2012) Epigenesis in humans: Can maternal methyl-donor-deficient diets induce epigenetic alterations in their offspring? PhD Thesis, London School of Hygiene and Tropical Medicine, University of London.

7.   Dominguez-Salas P, Moore SE, Baker MS, et al. (2014) Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. *Nat. Commun.* **5**, 3746.

8.   Moore SE, Fulford AJ, Darboe MK, et al. (2012) A randomized trial to investigate the effects of pre-natal and infant nutritional supplementation on infant immune development in rural Gambia: the ENID trial: Early Nutrition and Immune Development. *BMC Pregnancy Childbirth* **12**, 107.

9.   Silver MJ, Kessler NJ, Hennig BJ, et al. (2015) Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. *Genome Biol.* **16**, 118.

10.   Hernandez-Vargas H, Castelino J, Silver MJ, et al. (2015) Exposure to aflatoxin B1 in utero is associated with DNA methylation in white blood cells of infants in The Gambia. *Int. J. Epidemiol.* **44**, 1238–48.

Chapter 3 Literature Review

_____

Summary of Chapter

Background: Mounting evidence suggests that nutritional exposures during pregnancy influence the fetal epigenome, and that these epigenetic changes can persist postnatally, with implications for disease risk across the lifecourse.

Methods: We review human intergenerational studies using a three-part search strategy. Search 1 investigates associations between pre-conceptional or pregnancy nutritional exposures, focussing on one-carbon metabolism, and offspring DNA methylation. Search 2 considers associations between offspring DNA methylation at genes found in the first search and growth-related, cardio-metabolic and cognitive outcomes. Search 3 isolates those studies explicitly linking maternal nutritional exposure to offspring phenotype via DNA methylation. Finally, we compile all candidate genes and regions of interest identified in the searches and describe their genomic locations, annotations and coverage on the Illumina Infinium Methylation beadchip arrays.

Results: We summarise findings from the 34 studies found in the first search, the 31 studies found in the second search and the eight studies found in the third search. We provide details of all regions of interest within 45 genes captured by this review.

Conclusions: Many studies have investigated imprinted genes as priority loci, but with the adoption of microarray-based platforms other candidate genes and gene classes are now emerging. Despite a wealth of information, the current literature is characterised by heterogeneous exposures and outcomes, and mostly comprise observational associations that are frequently underpowered. The synthesis of current knowledge provided by this

review identifies research needs on the pathway to developing possible early-life interventions to optimise lifelong health.

Notes

I have inserted the published version. Where the paper refers to supplementary material and tables these can be found in **Annexes 3.1 to 3.4.**

I published two other reviews for different audiences. Since these covered very similar material I have included them in the Annexes.

- **Annex 3.5:** James P, Silver M, Prentice A. Epigenetics, Nutrition, and Infant Health. In: Karakochuk C, Whitfield K, Green T and Kraemer K eds. *The Biology of the First 1,000 Days.* 2017. p. 335–54.

  This is a textbook chapter written for a readership of public health professionals and students in the medical and nutrition field.

- **Annex 3.6:** James P, Silver M, Prentice A. Epigenetics, Nutrition, and Infant Health. 2015; Sight and Life 29(1): 35-38.

  This is a magazine article for a readership of both academics, practitioners and students, and is written as a short primer.

The review in this chapter was also used to identify candidate genes that were not included on the Infinium MethylationEPIC BeadChip array, and hence needed to be pyrosequenced, in the EMPHASIS project (Epigenetic Mechanisms linking Pre-conceptional nutrition and Health Assessed in India and Sub-Saharan Africa, ISRCTN14266771). The protocol paper for this project, for which I am a co-author, can found here:

Chandak GR *et al*. Protocol for the EMPHASIS study; epigenetic mechanisms linking maternal pre-conceptional nutrition and children's health in India and Sub-Saharan Africa. *BMC Nutrition* (2017) **3:**81.

https://bmcnutr.biomedcentral.com/articles/10.1186/s40795-017-0200-0

RESEARCH PAPER COVER SHEET

SECTION A – Student Details

Student	Philip James
Principal Supervisor	Dr Matt Silver
Thesis Title	Maternal nutrition, seasonality and epigenetics: an exploration of one-carbon metabolism and a novel nutritional supplement design in The Gambia.

SECTION B – Paper already published

Where was the work published?	International Journal of Epidemiology		
When was the work published?	17th August 2018		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N/A		
Have you retained the copyright for the work?*	No	Was the work subject to academic peer review?	Yes

If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

Copyright Clearance Center RightsLink®

OXFORD UNIVERSITY PRESS

Title: Candidate genes linking maternal nutrient exposure to offspring health via DNA methylation: a review of existing evidence in humans with specific focus on one-carbon metabolism

Author: James, Philip; Sajjadi, Sara

Publication: International Journal of Epidemiology

Publisher: Oxford University Press

Date: 2018-08-17

Copyright © 2018, Oxford University Press
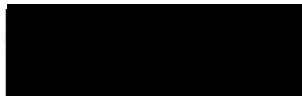
Multi-authored Work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper.

My supervisor and I proposed the idea for this review paper at a steering group meeting for the EMPHASIS project, which is a collaboration between UK, Gambian and Indian institutions to investigate epigenetic effects of preconception nutrition interventions. I coordinated the management of the research paper between all co-authors and produced a complete first draft of the paper for co-author review. I had ownership of the first and third searches in the review. Sara Sajjadi performed the second search for methylation-phenotype associations and provided me with a list of study details for me to write up in the narrative. She also helped me to check that all candidate genes from the literature searches were captured. The genomic coordinates of the regions of interest were looked up by Ashutosh Singh. Ayden Saffari and Matt Silver helped to re-structure the draft of the discussion section and provided valuable suggestions for edits. All co-authors provided comments on the paper, which I incorporated into a final manuscript. All supplementary material (provided in the Annexes) is my own work.
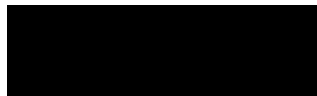
Student Signature:

Date: 15th October 2018

Supervisor Signature:

Date: 15th October 2018

IEA

Original article

Candidate genes linking maternal nutrient exposure to offspring health via DNA methylation: a review of existing evidence in humans with specific focus on one-carbon metabolism

Philip James [ID] ,[1†] Sara Sajjadi,[2†] Ashutosh Singh Tomar,[2†]
Ayden Saffari,[1] Caroline H D Fall,[3] Andrew M Prentice,[1]
Smeeta Shrestha,[2,4] Prachand Issarapu,[2] Dilip Kumar Yadav,[2]
Lovejeet Kaur,[2] Karen Lillycrop,[5‡] Matt Silver,[1‡] Giriraj R Chandak; [2*‡]
the EMPHASIS study group

[1]MRC Unit The Gambia at the London School of Hygiene and Tropical Medicine, London, UK, [2]Genomic Research on Complex Diseases (GRC Group), CSIR-Centre for Cellular and Molecular Biology, Hyderabad, India, [3]MRC Life course Epidemiology Unit, University of Southampton, Southampton General Hospital, Southampton, UK, [4]School of Basic and Applied Sciences, Dayananda Sagar University, Bangalore, India and [5]Research Centre for Biological Sciences, Institute of Developmental Sciences, University of Southampton, Southampton, UK

*Corresponding author. Genomic Research on Complex diseases (GRC Group), CSIR-Centre for Cellular and Molecular Biology (CSIR-CCMB), Hyderabad 500007, India. E-mail: chandakgrc@ccmb.res.in

[†]Joint first authors.

[‡]Joint last authors.

Abstract

Background: Mounting evidence suggests that nutritional exposures during pregnancy influence the fetal epigenome, and that these epigenetic changes can persist postnatally, with implications for disease risk across the life course.

Methods: We review human intergenerational studies using a three-part search strategy. Search 1 investigates associations between preconceptional or pregnancy nutritional exposures, focusing on one-carbon metabolism, and offspring DNA methylation. Search 2 considers associations between offspring DNA methylation at genes found in the first search and growth-related, cardiometabolic and cognitive outcomes. Search 3 isolates those studies explicitly linking maternal nutritional exposure to offspring phenotype via DNA methylation. Finally, we compile all candidate genes and regions of interest identified in the searches and describe their genomic locations, annotations and coverage on the Illumina Infinium Methylation beadchip arrays.

1

Results: We summarize findings from the 34 studies found in the first search, the 31 studies found in the second search and the eight studies found in the third search. We provide details of all regions of interest within 45 genes captured by this review.

Conclusions: Many studies have investigated imprinted genes as priority loci, but with the adoption of microarray-based platforms other candidate genes and gene classes are now emerging. Despite a wealth of information, the current literature is characterized by heterogeneous exposures and outcomes, and mostly comprise observational associations that are frequently underpowered. The synthesis of current knowledge provided by this review identifies research needs on the pathway to developing possible early life interventions to optimize lifelong health.

Key words: Epigenetics, DNA methylation, fetal programming, Developmental Origins of Health and Disease, one-carbon metabolism, candidate genes, metastable epialleles, cognitive development, cardiometabolic outcomes, growth

---

Key Messages

- The body of evidence linking maternal nutritional exposure to offspring phenotype via DNA methylation in humans is rapidly growing yet currently remains complex and inconsistent.
- Candidate genes in the field of intergenerational nutritional epigenetics go beyond imprinted genes to include other gene classes such as metastable epialleles.
- Going forwards, there is a continued need for adequately powered prospective cohort studies with repeated longitudinal measurements and randomized nutritional interventions to track the full continuum from maternal exposure to offspring epigenotype to later phenotype.

Introduction

Epigenetic modifications influence gene expression without altering the nucleotide sequence, through the action of a diverse array of molecular mechanisms including DNA methylation, histone modifications and RNA-mediated effects.[1] Epigenetic processes have been implicated in the aetiology of a variety of diseases,[2] most prominently cancer[3] and fetal growth disorders.[4] Epigenetic marks are mitotically heritable and can be influenced by the environment,[5] suggesting a potential mechanism linking early life exposures to later phenotype,[6,7] a notion supported by animal studies.[8–10] However, the extent to which epigenetics plays a role in fetal programming in humans remains relatively unexplored. In this review we collate evidence from human intergenerational studies, exploring which nutritional exposures during pregnancy may affect DNA methylation in the offspring, and the possible impact of such modifications on health and disease risk across the life course.

DNA methylation and gene expression

Many biological processes rely on DNA methylation, including genomic imprinting, X-chromosome inactivation and tissue-specific gene expression.[11] DNA methylation describes the addition of a methyl group to a cytosine base at the 5' carbon position to form 5-methylcytosine, catalyzed by DNA methyltransferases (DNMTs). This most commonly occurs at cytosine bases adjacent to guanine, termed CpG ('cytosine-phosphate-guanine') sites. Regions of high CpG density are known as 'CpG islands', and approximately two-thirds of human genes contain these in their promoter regions.[12] DNA methylation has been shown to influence transcriptional activity either by blocking transcription factors binding to the DNA, or by the recruitment of histone modifiers which promote a closed chromatin structure and gene silencing.[1] CpG methylation within promoters is typically associated with transcriptional silencing,[13] although not consistently, and the effect of DNA methylation may vary depending on which region within the gene is methylated.[14] There is also increasing evidence that DNA methylation and histone modifications work in concert with non-coding RNAs to regulate gene expression.[15] DNA methylation plays a role in chromatin remodelling, as DNMT enzymes at CpG sites can be physically linked to enzymes which bring about histone methylation and de-acetylation.[13] MicroRNAs (miRNAs) affect gene expression through binding to messenger RNAs (mRNAs) and repressing translation,[16] including mRNAs that control the expression of DNMTs

and histone deacetylases.[15] The transcription of some miRNA classes can be influenced by CpG methylation and histone modifications.[16]

Epigenetics, windows of plasticity and the Developmental Origins of Health and Disease

The Developmental Origins of Health and Disease (DOHaD) hypothesis posits that early life exposure to environmental insults can increase the risk of later adverse health outcomes.[7] David Barker's early cohort studies showed that lower birthweight was associated with an increased risk of hypertension, type 2 diabetes (T2D) and cardiovascular disease in later life,[17] findings that were widely replicated.[18] Risk of disease was further exacerbated by rapid childhood weight gain, adult obesity and other lifestyle factors such as unhealthy diets, smoking and lack of exercise.[19,20] The Dutch Hunger Winter studies showed that exposure to famine during pregnancy was associated with a wide range of phenotypes in the adult offspring, including increased blood pressure,[21] obesity[22] and schizophrenia,[23] effects that depended on the timing of the exposure during pregnancy.[22]

Epigenetic processes are emerging as potential mechanisms to explain these and other associations found in the DOHaD literature. For example the 'thrifty epigenome' hypothesis proposes that *in utero* exposures can shape an epigenetic signature, resulting in a phenotype that is 'adapted' to the early life environment but which may prove to be 'maladapted' if the environment changes in later life.[24] Therefore famine exposure during pregnancy could programme 'thrifty epigenotypes' that are adapted to a nutritionally poor environment, but this may subsequently trigger metabolic disease if the adult environment changes to one that is nutritionally abundant.

The periconceptional period is a time of rapid cell differentiation and epigenetic remodelling, and may therefore represent a critical window during which the developing epigenome is sensitive to environmental influences.[25] We define the periconceptional window from 14 weeks preceding conception until 10 weeks after conception.[26] Within 48 hours of fertilization, there is rapid erasure of methylation marks to render the developing cells pluripotent.[11] After implantation, re-methylation occurs in a tissue-specific manner, and continues throughout pregnancy, enabling differentiation of somatic cells. A second wave of demethylation occurs in the primordial germ cells as they migrate to the genital ridge.[27] At this stage most parental imprints are erased, so that sex-specific imprints can be laid down. In boys the prospermatogonia then undergo re-methylation throughout gestation, whereas in girls the oocytes continue to be re-methylated over the duration of their maturation, with evidence of high activity as each egg ripens before ovulation.[27]

Notable classes of loci that may be especially sensitive to early environmental exposure include imprinted genes, metastable epialleles (MEs) and transposable elements (TEs).[6] Imprinted genes exhibit monoallelic expression, whereby only the maternally or paternally inherited allele is expressed, with expression controlled by regulatory regions whose methylation state is inherited in a parent of origin-specific manner.[28] MEs are genomic loci showing variable methylation between individuals, but showing high correlation in methylation status across tissues within the same individual, indicating establishment of methylation state in the first few days after conception, preceding gastrulation.[29] MEs therefore help to pinpoint the timing of an exposure influencing ME methylation to the periconceptional period.[30,31] TEs are small, mobile sequences of DNA that are thought to comprise 45% of the human genome.[32] They can insert into new genomic locations and become disruptive if transposed into a functional gene or when increasing copy number. Whereas most TEs are silenced epigenetically,[33] some have variable methylation patterns that have been shown to be influenced by nutrition in mice.[9] Their methylation states can alter neighbouring gene expression, exemplified by the Agouti mouse model detailed later.

Influence of nutrition on DNA methylation

A range of maternal exposures have been associated with DNA methylation including nutrition, stress, infection, pollutants, smoking, radiation, level of exercise and parental body composition.[34–36] Animal studies suggest that the epigenome is particularly sensitive to such environmental factors in early life, notably during the prenatal and neonatal periods.[9,25,37] Studies of the effects of early life nutrition on DNA methylation have shown that maternal under- or over-nutrition or differences in protein, fat, sugar or micronutrient intake during gestation can induce epigenetic and phenotypic changes in the offspring.[8,38] Recent studies have also shown that variations in paternal diet or body composition might also induce long-term epigenetic and phenotypic changes in the offspring.[39] One-carbon nutrients and metabolites are thought to be particularly important in the periconceptional period and during embryonic development.[26] One-carbon metabolism (OCM) pathways link the folate, methionine, homocysteine, transsulphuration and transmethylation metabolic pathways together (Figure 1) . These are crucial for many biochemical processes, including DNA methylation.

Nutrition plays a key role in OCM by providing substrates (folate, methionine, choline and betaine) and essential co-factors (vitamins B12, B6 and B2). For example, B12 is required by methionine synthase to methylate
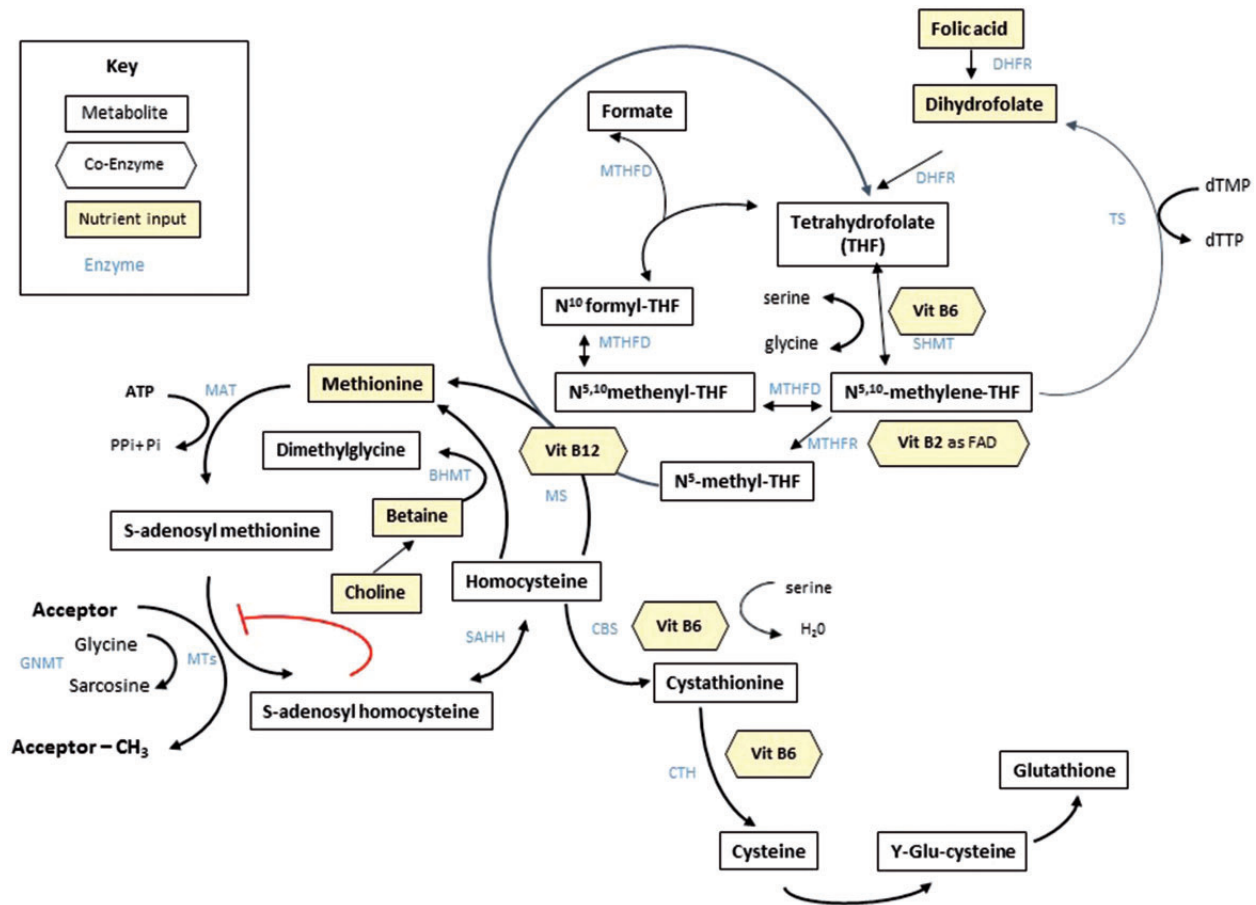
Figure 1. A simplified summary of one-carbon metabolism. BHMT, Betaine Homocysteine MethylTransferase; CBS, Cystathionine-Beta-Synthase; CTH, Cystathionine Gamma-Lyase; DHFR, Dihydrofolate Reductase; dTMP, Deoxythymidine Monophosphate; dTTP, Deoxythymidine Triphosphate; FAD, Flavin Adenine Dinucleotide; GNMT, Glycine N-MethylTransferase; MAT, Methionine AdenosylTransferase; MS, Methionine Synthase; MT, Methyl Transferases; MTHFD, MethyleneTetraHydroFolate Dehydrogenase; MTHF, MethyleneTetraHydroFolate Reductase; SAHH, S-Adenosyl Homocysteine Hydrolase; SHMT, Serine HydroxyMethylTransferase; TS, Thymidylate Synthase. Source: reproduced with permission from James *et al*. Epigenetics, nutrition and infant health. In: Karakochuk C, Whitfield K, Green T, Kraemer K (eds). *The Biology of the First 1000 Days*. Boca Raton, FL: CRC Press, 2017.

homocysteine, B6 is essential in the homocysteine trans-sulphuration pathway, and both B6 and B2 are needed to reduce dietary folate to methyltetrahydrofolate. A more detailed overview of OCM and the role of nutrients in these pathways is provided in Supplementary Material 1, available as Supplementary data at *IJE* online.

The potential for maternal nutrition to both alter offspring DNA methylation and influence phenotype is famously illustrated by the Agouti mouse experiments. Two groups of pregnant dams were fed diets that differed only in nutrients essential to OCM (folic acid, choline, betaine and B12). Increased levels of one-carbon nutrients increased methylation in the isogenic pups at a retrotransposon locus [Intracisternal A Particle (IAP), also an ME] upstream of the Agouti gene. The degree of expression of the Agouti gene depended on the level of IAP methylation, and this in turn altered the pups' fur colour, as well as their appetite, adiposity and glucose tolerance in adulthood.[6,9]

Review methodology

We performed a narrative review of the literature in three stages to form the thematic analysis in this paper. First we searched for studies describing associations between pre-conceptional or pregnancy nutritional exposures and DNA methylation in offspring. We limited this search to human studies that used an intergenerational design. We included nutritional exposures in dietary or supplemental form related to OCM, or broader measures that could influence availability of such nutrients (famine, seasonal diets and macronutrients). We excluded paternal exposures and nutrients not directly involved in OCM, and we only considered epigenetic studies focusing on DNA methylation. Second, we searched for human studies linking infant DNA methylation to a subset of phenotypic outcomes (growth-related, cardiometabolic and cognitive), restricting the included studies to those describing methylation at

genetic loci identified in the first search ('nutrition-sensitive' loci). Third, we isolated those studies explicitly linking maternal nutritional exposure to offspring phenotype via DNA methylation. Three authors (P.J., S.S., A.S.T.) performed the searches in PubMed and Google Scholar, assessing titles and abstracts against the inclusion criteria. Reference sections of included studies and relevant review papers were also used to help confirm that key studies had been included. Searches took place from January to March 2017. Supplementary Material 2, available as Supplementary data at IJE online, details the strategy and gives an example of the search terms used in PubMed.

Review of studies linking maternal nutritional exposure to offspring DNA methylation

We provide a broad overview of the associations found in the literature between maternal nutritional exposure and offspring DNA methylation in Table 1. Below we briefly review the associations by type of exposure, but refer the reader to detailed information on the individual studies ($n = 34$) in Supplementary Table 1, available as Supplementary data at IJE online, which includes information on the nutritional exposures, timing of exposures, study design, DNA tissue, age of offspring and associated genes. All gene names are defined in Table 4 (see candidate gene data summary, below).

Table 1. Summary of associations between maternal one-carbon metabolites and broader nutritional exposures with offspring DNA methylation

Timing of exposure	Maternal exposure[a]	Offspring DNA methylation association (↑/↓: increased/decreased methylation)
Periconception	↑B2	↑PLAGL1 (ZAC1),[40] ↑VTRNA2-1[41]
	↑Betaine	↑ DNMT1,[42] ↑POMC,[43] ↑RXRA[44]
	Famine	↓IGF2,[45] ↑↓[b]IGF2,[46] ↓INSIGF,[46,47] ↑IL10,[47] ↑GNASAS,[47] ↑LEP,[47] ↑ABCA1,[47] ↑MEG3,[47] ↑TACC1,[48] ↑ZNF385A,[48] ↓TMEM105,[48] ↑PAX8,[49] ↓ZFP57,[4], ↓PRDM9[49]
	↑Folates	↓STX11,[50] ↓OTX2,[50] ↓TFAP2A,[50] ↓CYS1,[50] ↓LEP,[44] ↑RXRA[44]
	↑Folic acid	↑LEP,[42] ↓H19,[51] ↑IGF2,[52] ↓IGF2[44]
	↑Multiple micronutrients	↓GNASAS,[53] ↓MEG3,[53] ↓IGF2R,[53] ↓MEST[53]
	Seasonality of one-carbon metabolites[c]	↑POMC,[43] ↑VTRNA2-1,[41] ↑BOLA3,[30] ↑FLJ20433,[30] ↑PAX8,[30] ↑SLITRK1,[30] ↑ZFYVE28,[30] ↑RBM46[31]
1st and 2nd trimester	↑B6	↑MEG3[54]
	↑Betaine	↓ LEP[42]
	↑Carbohydrates	↓RXRA[55]
	↑Choline	↓DNMT1[42]
	Famine	↑FAM150B,[48] ↑SLC38A2,[48] ↑PPAP2C,[48] ↓OSBPL5/MRGPRG,[48] ↑TACC1,[48] ↑ZNF385A,[48] ↑PAX8,[49] ↓ZFP57,[49] ↓PRDM9[49]
	↑Folates	↓PEG3,[56] ↑NR3C1,[57] ↓MEG3,[56] ↓PLAGL1,[56] ↑IGF2,[56] ↓LEP,[42] ↓DNMT1[42]
	↑Folic acid	↓PEG3,[58] ↑IGF2,[58] ↓DNMT1[44]
3rd trimester	↑B2	↑PLAGL1 (ZAC1)[40]
	↑B12	↓IGF2[59]
	↑Choline	↑↓[d] NR3C1,[60] ↑↓[d]CRH,[60] ↑DNMT1,[42,44]
	Famine	↓GNASAS,[47] ↑TACC1,[48] ↑ZNF385A,[48] ↑PAX8,[49] ↓ZFP57,[49] ↓PRDM9[49]
	↑Folates	↑DNMT1,[44] ↓RXRA,[42] ↑LASP1,[61] ↑ACADM,[61] ↑WNT9A,[61] ↑FZD7,[61] ↓ZFP57,[61] ↓LY6E,[61] ↓C21orf56[61]
	↑Folic acid	↑RXRA[42]
	↑ Meat and fish intake	↑HSD2[62]
	↑ High sugar, high fat diet	↑IGF2[63]
	↑Omega-3 PUFA	↓H19,[64] ↑IGF2,[6] mostly ↓associations in EWAS[65]
	↑Omega-6 PUFA	↓MIRLET7BHG[66]

[a]Like nutrients are shaded in the same colour during each time period.
[b]Different associations at different loci within gene.
[c]Rainy season (higher concentration of most one-carbon metabolites) versus dry season.
[d]Different associations between different tissues.
EWAS, epigenome-wide association study; PUFA, polyunsaturated fatty acids.

Folate

Associations between maternal folate exposure and the off-spring methylome are inconsistent, with varying effects according to the form of folate (dietary folates or folic acid supplements)[58] the timing of exposure,[42,58] baseline maternal folate status,[50,61] underlying genotype,[67] the genomic region affected[68] and individual CpG site.[42]

Periconceptional folic acid has been positively associated with offspring methylation at *LEP*,[42] inversely associated with methylation at *H19*,[51] and has demonstrated both positive[52] and inverse[44] associations at *IGF2*. Not all studies have found an effect of periconceptional folic acid exposure.[58] Supplementation started after 12 weeks of gestation has been associated with increased offspring methylation at *IGF2* and decreased methylation at *PEG3*.[58] Folic acid taken up to the end of the second trimester has been inversely associated with *DNMT1* methylation, but positively correlated at the same locus when the folic acid consumption was extended into the third trimester.[44]

Data for dietary folate intakes (assessed using questionnaires or plasma samples) are equally variable. Periconceptional folate intake and offspring DNA methylation were inversely associated with the majority of differentially methylated CpGs in an epigenome-wide screen, although this trend reversed in stratified analysis among women with low intakes (<200 µg/day).[50] Periconceptional intakes have also been inversely associated with methylation at *LEP* and positively associated at *RXRA*.[44] First trimester folate exposure has shown positive associations with DNA methylation at *IGF2*[56] and *NR3C1*,[57] and inverse associations at *MEG3*, *PLAGL1* and *PEG3*.[56] For second trimester folate exposure, studies have reported inverse associations at multiple differentially methylated CpG sites,[68] and at *LEP* and *DNMT1*.[42] Third trimester folate exposure has shown positive associations with methylation at *DNMT1*,[44] and at *LASP1*, *ACADM*, *WNT9A*, *C21orf56* and *FZD7*,[61] but inverse associations at *ZFP57*, *LY6E* and *RXRA*.[42,61]

B vitamins

Maternal serum B12 at first antenatal visit has been inversely associated with cord blood global methylation levels,[67] and inversely associated with offspring *IGF2* methylation when exposure timing was at delivery.[59] Some studies have assessed joint effects of B vitamins. One study assessed pre-pregnancy and third trimester maternal B2, B3, B6, folate and B12 intake, and found a positive correlation between maternal B2 and offspring methylation at *PLAGL1* (*ZAC1*) at both time points.[40] Another study found no associations between first trimester maternal plasma B12 and B6 concentrations with offspring methylation at *H19, PEG10/SGCE* and *PLAGL1*, but there was a positive trend in methylation at *MEG3* across maternal B6 quartiles.[54]

Choline and betaine

To date there is one human intervention study investigating the effect of supplementing mothers' diets with choline (480 mg vs 930 mg) in the third trimester on offspring DNA methylation. The intervention increased methylation at *NR3C1* and *CRH* in fetal placental tissue but reduced methylation in cord blood. No effect was seen at *GNAS-AS, IGF2, IL10* or *LEP*.[60] In observational studies, second trimester choline intake has been inversely associated with *DNMT1* methylation in cord blood.[42] Third trimester choline intake has been positively associated with *DNMT1* methylation in cord blood and in infant buccal cells.[42,44] Maternal periconceptional betaine intake has been positively associated with cord blood methylation at *DNMT1*, *RXRA* and *POMC*,[42–44] and second trimester intake inversely associated with *LEP* methylation.[42]

Polyunsaturated fatty acids

Polyunsaturated fatty acids (PUFAs) are thought to influence OCM by upregulating enzymes responsible for the methylation of homocysteine to methionine and by directly influencing demand for methyl groups via phosphatidylcholine (described in Supplementary Material 1, available as Supplementary data at *IJE* online). There have been several studies of PUFA supplementation in mothers. In one trial, omega-3 PUFA supplementation in the second and third trimesters showed no difference in the cord blood methylation of various gene promoter sites, but the intervention increased global methylation (LINE-1) in offspring of mothers who smoked.[69] It also decreased *H19* methylation, and increased *IGF2* methylation in offspring of overweight mothers.[64] A more recent trial, also implemented in the second and third trimesters, found omega-3 PUFA supplementation was associated with 21 differentially methylated regions (DMRs) at birth.[65] These were predominantly hypomethylated in the intervention group. However, not all omega-3 PUFA supplementations trials have demonstrated an effect on methylation.[70] Maternal plasma omega-6 PUFA concentrations in the third trimester have been inversely associated with offspring *MIRLET7BHG* methylation.[66]

Broader nutrition measures: famine studies, seasonal exposures, macronutrients

Several studies have used broader measures of maternal nutritional exposure, such as famine, season of conception and macronutrient intake. During the Dutch Famine of 1944, there was a large drop in all food intakes, with average energy intake reduced to 500–1000 kcal per day.[71]

In follow-up studies of adults who were exposed to famine *in utero*, exposure in early pregnancy (periconception and up to 10 weeks of gestation) was associated with lower methylation of *INSIF* and *TMEM105*, increased methylation at *IL10*, *GNASAS*, *LEP*, *ABCA1*, *MEG3*, *TACC1* and *ZNF385A*, and both increased and decreased methylation at *IGF2* depending on the loci within the gene.[45–48] Not all these effects were seen in those exposed during late gestation.[45,48] In a candidate gene analysis of putative metastable epialleles, offspring exposed to famine for at least 7 months during gestation in Bangladesh had higher methylation at *PAX8* and lower methylation at *PRDM9* and *ZFP57*, compared with unexposed controls.[49]

One study found an inverse association between maternal second trimester carbohydrate intake and infant *RXRA* methylation.[55] Another study looked at the effect of a prenatal diet high in fat and sugar and found a positive association with offspring *IGF2* methylation.[63] Higher methylation at *GR* has been observed in infants of mothers having higher meat/fish/vegetables and lower bread/potato intake in late pregnancy (>20 weeks of gestation compared with earlier in pregnancy) and increased infant methylation at *HSD2* has been associated with increased maternal meat and fish intake in late pregnancy.[62] In a pilot trial of periconceptional multiple micronutrient supplementation (UNIMMAP) for mothers, there were sex-specific effects on infant methylation at *IGF2R*, *GNASAS*, *MEG3* and *MEST*.[53] The difficulty of such studies, however, is that it is not possible to know which nutrient deficits or imbalances caused the epigenetic effects. In The Gambia, where season has marked effects on maternal diet and body weight,[72] children conceived in the rainy season had higher methylation in peripheral blood lymphocytes at six MEs, at *VTRNA2–1* and at *POMC* compared with those conceived in the dry season.[31,41,43] This may reflect a role of one-carbon-related nutrients; in the rainy season, maternal periconceptional plasma showed higher concentrations of folate, B2, methionine, betaine, S-adenosyl methionine (SAM):S-adenosyl homocysteine (SAH) ratio and betaine:dimethylglycine (DMG) ratio, and lower B12 and homocysteine, indicating higher methylation potential.

Aside from those considered above, the list of maternal exposures associated with changes in infant DNA methylation continues to grow. These include further nutrition-related exposures (e.g. dietary polyphenols,[73] vitamin D[74,75] and vitamin A[76]) non-nutrition-related exposures (e.g. maternal stress[77] and toxin exposure[78]) and factors that span the spectrum of nutrition and health-related considerations (e.g. maternal hyperglycaemia,[79] maternal body mass index (BMI),[80–82] intrauterine growth restriction (IUGR),[83–85] the microbiome[86] and infection[87]). The ongoing challenge is not only to identify relevant exposures, but also to delineate the consequences for human health across the life course. It is to this latter point that we now turn.

Review of studies linking nutrition-associated DNA methylation loci to health outcomes

In animal studies, nutritional exposures in pregnancy bring about distinct phenotypic effects in offspring via epigenetic mechanisms. Differential methylation of genes may induce phenotypic variation by the modulation of gene expression which may alter tissue structure, homeostatic control processes and the activity of metabolic pathways.[88] Often cited examples include the effects of maternal methyl donor supplementation on offspring coat colour and adiposity in the Agouti mouse, and the development of the fertile queen bee from genetically identical larvae by epigenetic silencing of *DNMT3*, caused by preferential feeding of royal jelly.[9,89]

In this section we focus on evidence provided by two types of studies:

i. Those reporting associations between methylation at the nutrition-sensitive epigenetic loci described above and offspring phenotypes; these are summarized in Table 2, with detailed information on all included studies (*n* = 31) in Supplementary Table 2, available as Supplementary data at *IJE* online;

ii. Those linking maternal nutrition exposure, infant DNA methylation and offspring phenotypic effects in a single study (*n* = 8); these are summarized in Table 3.

We consider three broad categories of offspring phenotypic outcomes: growth and body composition, cardiometabolic risk markers and cognitive function.

Growth and body composition

DNA methylation signatures in different tissues such as cord and peripheral blood, placenta, subcutaneous and visceral adipose tissue and buccal cells have been associated with growth outcomes such as size at birth (usually birthweight, with or without adjustment for gestational age), child/adult adiposity and skeletal growth or bone size/quality (see Supplementary Table 2, available as Supplementary data at *IJE* online).

Birth size: most studies investigating growth-related phenotypes have analysed imprinted genes due to their known role in fetal growth regulation.[106] Chromosomal region 11p15.5 contains two imprinting control regions (ICRs): the *H19/IGF2* (ICR1) and *KCNQ1/CDKN1C* (ICR2) domains.[107] Russell–Silver Syndrome (RSS, a

Table 2. Summary of associations between methylation at nutrition-sensitive genetic loci and phenotypes

Direction of DNA methylation/locus	Associated phenotype/direction (↑/↓: increased/decreased)	Tissue analysed	Age at methylation measurement
Birth size			
↑*H19*,[56] ↑*PLAGL1*,[56] ↓*MEG3*,[56] ↓*MIRLET7BHG*,[66] ↑*IGF2*[90]	↑Birthweight	Cord blood	Birth
↑*IGF2 DMR2*[91]	↑Birthweight	Placenta	Birth
↓*IGF2*,[52] ↑*HSD2*[62]	↑Birthweight	Peripheral blood	17 months,[52] 40 years[62]
↑*H19* ICR[62]	↓Birth length	Peripheral blood	40 years
↑*PLAGL1*[40]	↑Estimated fetal weight at 32 weeks of gestation	Cord blood	Birth
↑*HSD2*[62]	↓Neonatal ponderal index	Peripheral blood	40 years
↓*IGF2 DMR0*,[83] ↑*H19*[92]	↑Small for gestational age	Cord blood	Birth
↑*MEST*,[93] ↑*LEP*[94]	↑Small for gestational age	Placenta,[93] cord blood[94]	Birth
↓*IGF2 DMR0*[95]	↑Small for gestational age	Peripheral blood	11 years
Anthropometric measures/adiposity			
↑*PLAGL1*[40]	↑Weight at age 1 year	Cord blood	Birth
↑*PLAGL1*[40]	↑Body mass index (BMI) z-score at age 1 year	Cord blood	Birth
↑*IGF2 DMR2*[91]	↑Height, head and thorax circumference at birth	Placenta	Birth
↑*POMC*[96]	↑Obesity at age 11 years	Peripheral blood	11 years
↑*IGF2/H19* ICR[97]	↓Early childhood head circumference	Peripheral blood	1–10 years
↑*H19* ICR,[62] ↑*HSD2*[62]	↑Weight in adulthood	Peripheral blood	40 years
↑*H19* ICR,[62] ↑*HSD2*,[62] ↑*NR3C1 exon 1C*[62]	↑Waist circumference in adulthood	Peripheral blood	40 years
↑*POMC*,[43] ↑*H19* ICR,[62] ↑*HSD2*,[62] ↑ *NR3C1 exon* 1C,[62] ↓*LEP*[98]	↑BMI in adulthood	Peripheral blood	48,[43] 40,[62] 34.7[98] years
↑*RXRA*[55]	↑Adiposity at age 9 years	Cord blood	Birth
↓*LEP*[99]	↑Obesity at age 10–15 years	Saliva	10–15 years
↓*LEP*[100]	↑Obese subjects with insulin resistance at age 10-16 years	Peripheral blood	10–16 years
↑*IGF2/H19* ICR[97]	↑Skinfold thickness and subcutaneous adiposity at age 17 years	Peripheral blood	17 years
Skeletal growth and bone quality			
↓*RXRA*[75]	↑Bone mineral content at age 4 years	Cord blood	Birth
Cardiometabolic outcomes			
↑*LEP*[98]	↑Fasting low-density lipoproteincholesterol levels in adulthood	Peripheral blood, Subcutaneous adipose tissue	34.7 years
↑*H19* ICR,[62] ↓ *NR3C1 exon* 1F,[62] ↑*HSD2*[62]	↑Blood pressure in adulthood	Peripheral blood	40 years
↓*LEP*[101]	↑High-density lipoprotein (HDL) profile	Peripheral blood	17 months
↑*IGF2*[102]	↑Triglycerides (TG), ↑TG:HDL	Peripheral blood	11.6 years
Cognitive outcomes			
↑*IGF2*[63]	↑Early onset conduct problem, attention-deficit/hyperactivity disorder	Cord blood	Birth
↑*NR3C1*,[103,104] ↓*HSD2*[103,104]	↑Risk of being in a poorly regulated neurobehavioural profile	Placenta, Buccal cells	Birth
↑*LEP*[105]	↑Lethargy and hypotonicity	Placenta	Birth

disorder of impaired growth) is associated with hypomethylation of ICR1 and hypermethylation of ICR2. Beckwith-Wiedemann Syndrome (BWS, an over-growth disorder) is associated with hypermethylation of ICR1 and hypomethylation of ICR2.[108] Some studies indicate that patients with RSS and BWS exhibit abnormal methylation at multiple

Table 3. Studies linking maternal one-carbon metabolites or broader nutritional exposures to offspring DNA methylation and phenotype

Study	Exposure (exposure timing)	Offspring tissue analysed	Genes analysed	Phenotype investigated	Key findings (↑/↓: increased/decreased; ~ associated with)
Azzi S *et al.*[40]	Pre-pregnancy BMI, vitamins B2, B3, B6, folate, B12 (3 months before conception and last trimester)	Cord blood	*PLAGL1 (ZAC1)*	Pre- and post-natal growth	↑Pre-pregnancy and last trimester vitamin B2 ~ ↑*ZAC1* methylation; ↑Pre-pregnancy BMI ~ ↑*ZAC1* methylation; ↑*ZAC1* methylation index ~ ↑estimated fetal weight at 32 weeks of gestation, ↑BMI z-scores at age 1 year
Drake AJ *et al.*[62]	Maternal diet: food group analysis ('Early' <20 weeks and 'late' >20 weeks of gestation)	Peripheral blood	*IGF2, H19* ICR, *HSD2, NR3C1*	Birthweight, current height, weight, waist circumference, blood pressure	↑Meat/fish/vegetables and ↓bread/potato intake in late pregnancy ~ ↑*NR3C1 exon 1F* methylation; ↑Meat/fish intake in late pregnancy ~ ↑*HSD2* methylation; ↑*HSD2* methylation ~ ↓neonatal ponderal index, ↓birthweight, ↑adiposity measures and ↑blood pressure in adulthood (age 40 years); ↑*H19* ICR methylation ~ ↓birth length, ↓weight, ↑waist circumference, ↑BMI and ↑blood pressure in adulthood; ↑*NR3C1 exon 1C* methylation ~ ↑waist circumference, ↑BMI, ↑blood pressure in adulthood
Godfrey KM *et al.*[55]	Maternal carbohydrate intake (2nd trimester)	Cord blood	*RXRA, NOS3, SOD1, IL8, PIK3CD*	Adiposity	↓Maternal carbohydrate intake ~ ↑*RXRA* methylation; ↑*RXRA* methylation ~ ↑childhood fat mass, ↑% fat mass (at age 9 years)
Hoyo C *et al.*[56]	Maternal erythrocyte folate (1st trimester, median 12 weeks of gestation))	Cord blood	*IGF2, H19, PEG1/MEST, PEG3, PLAGL1, MEG3-IG, PEG10/SGCE, NNAT, DLK1/MEG3*	Birthweight	↑Folate levels ~ ↓methylation at *MEG3, PLAGL1, PEG3* and ↑methylation at *IGF2*; ↑Methylation at *H19, PEG10/ SGCE* and *PLAGL1* and

(continued)

Table 3. Continued

Study	Exposure (exposure timing)	Offspring tissue analysed	Genes analysed	Phenotype investigated	Key findings (↑/↓: increased/decreased, ~ associated with)
					↓*MEG3* methylation ~ ↑birthweight; *MEG3* methylation ~ strongest evidence for mediating association between folate and birthweight
Kühnen P et al.[43]	Maternal 1-carbon metabolites/season of conception (periconception)	Peripheral blood/MSH-positive neurons	*POMC*	Obesity/BMI	Gambian rainy season of conception and associated 1-carbon metabolites ~ ↑*POMC* methylation; ↑*POMC* methylation ~ ↑BMI, ↑obesity in children and adults
Lin X et al.[66]	Maternal BMI, glucose, plasma fatty acids, plasma vitamin D, serum B12, B6, folate, iron, zinc, magnesium (3rd trimester; 26–28 weeks of gestation)	Cord blood	Epigenome-wide association study	Birthweight, size and adiposity at 4 years	↑Maternal omega-6 PUFA ~ ↓cg25685359 (*MIRLET7BHG*) methylation; ↓*MIRLET7BHG* methylation ~ ↑birthweight
Rijlaarsdam J et al.[63]	High-fat and -sugar diet (3rd trimester, 32 weeks of gestation)	Cord blood, peripheral blood at age 7 years	*IGF2*	ADHD	Prenatal highfat and high sugar diet ~ ↑*IGF2* methylation; ↑*IGF2* methylation ~ ↑ADHD symptoms in early-onset persistent conduct (EOP) children age 7 years
Steegers-Theunissen RP et al.[52]	Maternal folic acid supplementation (periconception)	Peripheral blood	*IGF2*	Birthweight	↑Folic acid supplementation ~ ↑*IGF2* methylation at 17 months; ↑*IGF2* methylation ~ ↓birthweight

ADHD, attention-deficit/hyperactivity disorder; BMI, body mass index; ICR, imprinting control region; PUFA, polyunsaturated fatty acids.

gene loci.[109] Differences in methylation at these loci have also been associated with less extreme growth-related phenotypes. In a study of 50 French-Canadian mothers and infants, 31% of variance in birthweight was attributed jointly to differential *IGF2/H19* methylation and genotype of a particular *IGF2/H19* polymorphism (rs2107425).[91] The direction of association between methylation and birthweight, however, varies by study and tissue analysed.[90,91] For example, hypomethylation at *IGF2* DMRs have been associated with both increased and decreased birthweight.[52,83,90,95,110] Some studies have found no association with birthweight.[111] Further examples of the complex relationship between DNA methylation at various *IGF2/H19* DMRs and infant growth phenotypes are detailed in Supplementary Table 2, available as Supplementary data at *IJE* online.

The paternally expressed imprinted gene *MEST* acts as an inhibitor of human adipogenesis and is involved in skeletal muscle growth and development.[112] In placenta, increased methylation at the *MEST* transcription start site is correlated with reduced gene expression and IUGR.[93,113] Increased methylation at the paternally expressed *PLAGL1*, which codes for a cell growth suppressor protein, is associated with higher birthweight and weight at 1 year of age.[40]

Some studies have associated other (non-imprinted) genes with birth size. For example, small-for-gestational age newborns had higher methylation at *LEP* in cord blood than appropriate-for-gestational age infants.[94] Methylation at CpGs within *HSD11B2*, which codes for the enzyme responsible for catalyzing the conversion of cortisol to inactive cortisone, has been inversely related to newborn ponderal index in a cohort study.[62]

A small number of studies have investigated links between maternal nutrition, DNA methylation and newborn size. One study found that higher maternal erythrocyte folate levels in the first trimester were associated with decreased methylation in cord blood at *MEG3*, *PLAGL1* and *PEG3*, and increased methylation at *IGF2*.[56] Folate concentration and methylation at five DMRs were positively associated with birthweight. The authors hypothes"sized that the association of folate with birthweight could be mediated by differential methylation at *MEG3*, *H19* and *PLAGL1*, with *MEG3* contributing the strongest effect. Another cohort study found that higher maternal plasma glucose and omega-6 PUFA concentrations in the third trimester were associated with increased infant methylation at *IGDCC4* and *CACNA1G*, and decreased methylation at *MIRLET7BHG*. These methylation patterns were all associated with higher birthweight.[66]

Adiposity: a case-control study in Germany found that obese adults (BMI >35 kg/m^2) demonstrated lower methylation at *MEST* than in controls (BMI <25 kg/m^2), and used a separate dataset to suggest that such outcomes may be partially caused by intrauterine exposure to gestational diabetes mellitus.[114] In obese boys from the USA, an inverse association was reported between *LEP* methylation in buccal DNA and BMI, waist circumference (as z-scores) and percentage body fat.[99] *NR3C1* Exon 1 C methylation has been positively associated with waist circumference and BMI at age 40 years,[62] and increased *IGF2/H19* methylation has been associated with increased skinfold thickness and subcutaneous adiposity at age 17 years.[97]

A number of studies have investigated maternal nutritional exposure, DNA methylation and child adiposity. *POMC* codes for melanocyte-stimulating hormone (MSH) and is involved with leptin in the regulation of body weight. *POMC* is an ME, and children conceived in the dry season in The Gambia had lower DNA methylation at a *POMC* variably methylated region (VMR) compared with those conceived in the rainy season.[43] *POMC* VMR methylation influences *POMC* expression,[96] and methylation at this locus in blood and MSH-positive neurons is associated with BMI and obesity in children and adults.[43] Godfrey *et al.* (2011) found that lower carbohydrate intake during early pregnancy was associated with increased umbilical cord tissue methylation at *RXRA*, which in turn was associated with greater adiposity in the offspring at 9 years of age.[55]

Skeletal growth and bone quality: *RXRA* forms heterodimers with vitamin D (and other nuclear) receptors, facilitating their role in the regulation of bone metabolism.[115,116] Differential methylation of specific CpGs in *RXRA* in cord blood DNA has been inversely associated with percentage bone mineral content and bone mineral content adjusted for body size, measured at age 4 years, and also with maternal free 25(OH)-vitamin D index.[75]

Cardiometabolic outcomes

Maternal nutritional status during pregnancy and factors influencing fetal growth have been implicated in the aetiology of cardiometabolic outcomes such as dyslipidaemia, hypertension, type 2 diabetes (T2D) and cardiovascular disease later in life.[117,118]

Leptin has been studied extensively in the domain of cardiometabolic outcomes, owing to its role in metabolism and regulation of body weight.[119] *LEP* methylation at a specific CpG in blood and subcutaneous adipose tissue has been positively associated with low-density lipoprotein cholesterol levels in very obese (BMI >40 kg/m^2) adults.[98] In the same study, methylation at the *LEP* promoter was inversely correlated with BMI.[98] A different study found an inverse relationship between *LEP* methylation in whole

blood and high-density lipoprotein cholesterol levels in 17-month-old infants.[101] Furthermore, lower methylation in CpGs near the *LEP* transcription start site has been observed in adolescents with obesity and insulin resistance, although not with obesity alone.[100] *IGF2* methylation has also been related to lipid profile in obese children aged 11 years; those with intermediate methylation at the *IGF2* P3 promoter had higher triglycerides (TG) and a higher TG:high-density lipoprotein cholesterol ratio than those with hypomethylation.[102] *HSD2* methylation has been positively associated with systolic blood pressure,[62] and *NR3C1* exon1F and *H19* ICR methylation also show positive associations with both systolic and diastolic blood pressures in adults.[62] Note that adiposity and obesity (reviewed above) are also important risk factors that, alongside other markers, can signal increased risk of adverse cardiometabolic outcomes.[120]

Cognitive outcomes

The glucocorticoid receptors modulate the action of glucocorticoids and are involved in brain development and function.[121] *NR3C1* and *HSD11B2* genes regulate the action of cortisol and have been well studied in relation to neurobehaviour. Increased methylation at the *NR3C1* promoter and decreased methylation in *HSD11B2* in placental and infant buccal cell DNA have been associated with a high-risk neurobehavioural profile characterized by poor attention, high excitability, low quality of movement and signs of stress.[103,104] An increase in *LEP* methylation in placental DNA has been associated with an increased risk of lethargy and hypotonia among male infants.[105] Increased methylation at *IGF2* in cord blood has been associated with early onset persistent attention-deficit/ hyperactivity disorder (ADHD) in children between 7 and 13 years of age.[63]

Candidate gene data summary

In Table 4 we provide further details of the 45 'candidate genes' highlighted so far in this review. This includes information on their genomic location, the studies that considered them, regions of interest (ROIs) analysed and the coverage of ROIs on Illumina Infinium Methylation beadchip arrays.

Discussion

In this review we have described evidence in humans linking maternal nutrition during pregnancy with DNA methylation in the offspring, and linking DNA methylation at nutrition-sensitive loci to phenotypes at birth and outcomes in later life. As with all reviews, publication bias can mean that null findings may have been under-reported, and studies that do report associations may sometimes rely on *post hoc* subgroup analyses for significant findings. There are also numerous challenges specific to both the design and interpretation of intergenerational nutritional epigenetics studies which we discuss in the following sections.

Measuring nutritional exposures

Methods for measuring maternal nutritional exposure have limitations. For example, one of the most commonly used methods for this purpose are food frequency questionnaires, which suffer from recall bias and have differing validity by micronutrient.[123] Weighed records require accurate, context-specific dietary databases and well-trained data collectors, and may not accurately reflect normal eating habits.[124] However, these two approaches have the advantage of capturing food groups and combinations of nutrients that more direct tissue nutritional biomarkers can overlook.[125] Plasma biomarkers are challenging to interpret, given that they represent nutrient levels after absorption and through interaction with genotype, and are not simple reflections of dietary intake. Concentrations do not capture metabolite flux, and can be misleadingly low if tissue uptake is rapid. Of particular relevance to maternal gestational samples is the effect of haemodilution, which can lower several biomarker concentrations.[126] Maternal plasma nutrient concentrations are assumed to reflect dietary intake, and to correlate with cord blood concentrations and nutrient levels in fetal tissue, which may not be the case. Whereas positive correlations between maternal serum and cord blood serum are found for homocysteine, betaine, folate and B12, cord blood levels are multiple times higher, suggesting that these nutrients are homeostatically controlled to ensure fetal supply.[127] In the context of periconceptional studies, more research is needed on which accessible tissues best represent the nutritional milieu surrounding the developing embryo in the initial days after fertilization. In the meantime, serum or plasma levels, though imperfect, are likely to offer a more accurate representation of fetal nutrient exposure than dietary intake methods.

Most of the attention on nutritional exposures has focused on the provision of methyl groups and the necessary co-factors for DNA methylation. However, the periconceptional period is marked by an initial wave of demethylation to erase parental epigenetic marks, before the process of remethylation.[27] It is therefore important to consider the role nutrition could play in influencing demethylation. In demethylation, 5-methylcytosine is sequentially oxidized to 5-hydroxymethylcytosine and 5-formylcytosine (5fC)

Table 4. Candidate genes exhibiting associations between nutritional exposures during periconception and pregnancy and offspring DNA methylation. Links between methylation at nutrition-sensitive genes and offspring phenotype are also included

Blue = ME
Brown = imprinted
Yellow = ME and imprinted

Gene/region of Interest	Genomic features[c]	Exposure (↑/↓: increased/decreased)	Outcome (↑/↓: increased/decreased)	Coordinates of ROI in studies[d,e] (number of CpGs on 450k[a] and EPIC[b] arrays)
ABCA1 (ATP Binding Cassette Subfamily A Member 1)	Promoter marks; CpG island; binding site for multiple TFs	Famine	↑Methylation[47]	chr9: 107, 690, 502-107, 690, 821 (1)[a](5)[b]
ACADM (Acyl-CoA Dehydrogenase, C-4 To C-12 Straight Chain)	Multiple TFs binding sites; Promoter mark; Active Enhancer mark	↑Folate	↑Methylation[61]	chr1: 76, 189, 707-76, 190, 008 (6)[a](7)[b]
BOLA3 (BolA Family Member 3)	Enhancer and Promoter marks; CpG island; binding site for multiple TFs	Rainy season conception	↑Methylation[30]	chr2: chr2: 74, 357, 632-74, 357, 837 (1)[a,b]
CRH (Corticotropin-Releasing Hormone)	Enhancer mark	↑Choline	↓Methylation[60]	chr8: 67, 090, 692-67, 091, 132 (5)[a](8)[b]
CYS1 (Cystin 1)	Multiple TFs binding sites; Promoter mark	↑Folate	↓Methylation[50]	chr2: 10, 220, 719
DNMT1 (DNA Methyltransferase 1)	Multiple TFs binding sites; Promoter mark; Active Enhancer mark	↑Folate	↑Methylation[44], ↓Methylation[42]	chr19: 10, 305, 774-10, 305, 811 (2)[a,b]
	Multiple TFs binding sites; Promoter mark; Active Enhancer mark	↑Folic acid	↓Methylation[44]	chr19: 10, 305, 774-10, 305, 811 (2)[a,b]
	Multiple TFs binding sites; Promoter mark; Active Enhancer mark	↑Choline	↑Methylation[44], ↑↓Methylation[42]	chr19: 10, 305, 774-10, 305, 811 (2)[a,b]
	Multiple TFs binding sites; Promoter mark; Active Enhancer mark	↑Betaine	↑Methylation[42]	chr19: 10, 305, 774-10, 305, 811 (2)[a,b]
EXD3 (FLJ20433) (exonuclease 3'-5' domain containing 3)	Active Enhancer mark; CpG island	Rainy season conception	↑Methylation[30]	chr9: 140, 312, 206-140, 312, 339
FAM150B (Family With Sequence Similarity 150, Member B)	None	Famine	↑Methylation[48]	chr2: 366, 113 (1)[a,b]
FZD7 (Frizzled Class Receptor 7)	Multiple TFs binding sites; Promoter mark	↑Folate	↑methylation[61]	chr2: 202, 901, 045-202, 901, 470 (5)[a](4)[b]
GNASAS (Guanine Nucleotide Binding Protein (G Protein), Alpha Stimulating Activity Antisense RNA 1)	Enhancer marks; Multiple TFs binding sites	Famine (periconceptional)/Famine (late gestation)	↑Methylation/↓Methylation[47]	chr20: 57, 425, 815-57, 426, 108 (3)[a,b]
	CpG island; MYC binding site	UNIMMAP (supplementation)	↓Methylation[53]	chr20: 57, 429, 802-57, 430, 242 (1)[a](2)[b]
H19	Multiple TFs binding sites	↑Methylation	↑Birthweight[56]	chr11: 2, 011, 131-2, 011, 275 (2)[a,b]
	MYC and CTCF binding sites; Active promoter mark; weak enhancer mark	↑Methylation	↑Small for gestational age[92]	chr11: 2, 019, 727-2, 019, 921 (7)[a](6)[b]
	Multiple TFs binding sites	↑ Omega-3 PUFA	↓Methylation[64]	chr11: 2, 024, 197-2, 024, 340

(continued)

Table 4. Continued

Gene/region of Interest Blue = ME Brown = imprinted Yellow = ME and imprinted	Genomic features[c]	Exposure (↑/↓: increased/decreased)	Outcome (↑/↓: increased/decreased)	Coordinates of ROI in studies[d,e] (number of CpGs on 450k[a] and EPIC[b] arrays)
	Multiple TFs binding sites Enhancer Mark; CTCF-binding site	↑Folic acid ↑Methylation	↓Methylation[51] ↓Birth length, ↑weight in adulthood, ↑adult BMI, ↑adult blood pressure[62]	chr11: 2, 024, 254-2, 024, 261 chr11: 2, 021, 072-2, 021, 291 (2)[a,b]
HSD11B2 (Hydroxysteroid 11-Beta Dehydrogenase 2) *(HSD2)*	Multiple TFs binding sites; CpG island	↑Methylation	↓Neonatal ponderal index, ↑birthweight, ↑adult adiposity, ↑adult blood pressure[62]	chr16: 67464346-67464649 (3)[a](4)[b]
	Multiple TFs binding sites; Promoter mark; Active Enhancer mark; CpG island	↑Meat and fish intake	↑Methylation[62]	chr16: 67, 464, 981-67, 465, 111 (1)[a](2)[b]
	Multiple TFs binding sites, Active Enhancer mark	↓Methylation	↑Risk of being in a poorly regulated neurobehavioral profile[103,104]	chr16: 67, 464, 387-67, 464, 417
IGF2 (Insulin-like Growth Factor 2)	POL2A binding site	↑Folic acid	↓Methylation[44]	chr11: 2, 151, 629-2, 151, 721 (3) [a,b]
	POL2A binding site	↑Folate	↓Methylation[56]	chr11: 2, 151, 629-2, 151, 721 (3) [a,b]
	1 reported SNP (rs3741210)	↑Omega-3 PUFA	↑Methylation[64]	chr11: 2, 169, 425-2, 169, 556
	CTCF binding site; Enhancer mark; 2 reported SNPs (rs3741210, rs3741208)	↑Folic acid	↑Methylation[52]	chr11: 2, 169, 459 -2, 169, 796
	CTCF binding site; Enhancer mark; 2 reported SNPs (rs3741210, rs3741208)	↑Methylation	↑Birthweight[52]	chr11: 2, 169, 459 -2, 169, 796
	CTCF binding site; Enhancer mark; 2 reported SNPs (rs3741210, rs3741208)	Famine	↓Methylation[45,46]	chr11: 2, 169, 459-2, 169, 796
	POL2A and USF1 binding sites; 1 CpG island; 1 reported SNP (rs1803647)	↑Folic acid	↑Methylation[58]	chr11: 2, 154, 262-2, 154, 977 (5)[a,b]
	Multiple TFs binding sites; Promoter mark; Active Enhancer mark	↑Methylation	↑ADHD in early-onset persistent youth[63]	(37)[a](35)[b,f]
	Multiple TFs binding sites; Promoter mark; Active Enhancer mark	High-fat and -sugar diet	↓Methylation[63]	(37)[a](35)[b,f]
		↑Omega-3 PUFA	↑Methylation[64]	chr11: 2, 159, 107-2, 159, 965 (3)[a](4)[b]

(continued)

Table 4. Continued

Gene/region of Interest	Genomic features[c]	Exposure (↑/↓: increased/decreased)	Outcome (↑/↓: increased/decreased)	Coordinates of ROI in studies[d,e] (number of CpGs on 450k[a] and EPIC[b] arrays)
Blue = ME Brown = imprinted Yellow = ME and imprinted				
	POL2A binding site; Promoter mark; Active Enhancer mark; CpG island	↑Vitamin B12	↓Methylation[59]	chr11: 2, 161, 115-2, 161, 275 (4)[a,b]
	EZH2 and CTCF binding site; Promoter mark; CpG island	Famine	↓Methylation[46]	chr11: 2, 169, 385-2, 169, 489
	CTCF binding site; Enhancer mark; 2 reported SNPs (rs3741210, rs3741208)	Famine	↓Methylation[46]	chr11: 2, 170, 541-2, 170, 644
	Enhancer mark	↓Methylation	↑Small for gestational age[83]	chr11: 2, 169, 458-2, 169, 796
	CTCF binding site; Enhancer mark; 2 reported SNPs (rs3741210, rs3741208)	Famine	↑Methylation[46]	chr11: 2, 160, 906-2, 161, 372 (14)[a](13)[b]
	EZH2, RAD21 and CTCF binding site; Promoter mark; CpG island	Famine	↑Methylation[46]	chr11: 2, 161, 550-2, 161, 846 (1)[a](2)[b]
	EZH2, ZBTB7A and CTCF binding site; Promoter mark; CpG island	↓Methylation	↑Small for gestational age[95]	chr11: 2, 169, 467-2, 169, 640
	Enhancer mark; 1 reported SNPs (rs3741210)	Famine	↓Methylation[46]	chr11: 2, 155, 447-2, 155, 736 (1)[a,b]
	POLR2A and ZBTB7A binding site	↑Methylation	↑Birthweight, birth height, head and thorax circumference at birth[91]	chr11: 2, 154, 263-2, 154, 457 (2)[a,b]
	CpG island; USF1 and POL2A binding sites	↑Methylation	↑Birthweight[90]	chr11: 2, 169, 518-2, 169, 499
	None; CTCF and REST binding sites; CpG island	↑Methylation	↑TG and TG: HDL[102]	chr11: 2, 160, 374-2, 160, 610 (4)[a,b]
IGF2R (Insulin-like Growth Factor 2 Receptor)	CpG island; associated with SNP rs677882 and rs8191722	↑UNIMMAP (supplementation)	↓Methylation[53]	chr6: 160, 426, 403-160, 426, 850
IGF2/H19 ICR	None	↑Methylation	↑Head circumference between 1–10 years; ↑subcutaneous fat measures at age 17 years[97]	chr11: 2, 064, 402-2, 064, 717
IL10 (Interleukin 10)	Enhancer and Promoter marks; binding site for multiple TFs	Famine	↑Methylation[47]	chr1: 206, 946, 011-206, 946, 339 (2)[a](3)[b]
INSIGF (Insulin- Insulin-like Growth Factor 2)	None	Famine	↓Methylation[46,47]	chr11: 2, 182, 336-2, 182, 640 (5)[a](4)[b]
LASP1 (LIM And SH3 Protein 1)		↑Folate	↑Methylation[61]	chr17: 37, 123, 638-37, 123, 949 (9)[a,b]

(continued)

Table 4. Continued

Gene/region of Interest	Genomic features[c]	Exposure (↑/↓: increased/decreased)	Outcome (↑/↓: increased/decreased)	Coordinates of ROI in studies[d,e] (number of CpGs on 450k[a] and EPIC[b] arrays)
Blue = ME				
Brown = imprinted				
Yellow = ME and imprinted				
LEP (Leptin)	Multiple TFs binding sites; Promoter marks; Enhancer marks; 4 CpG islands; 25 reported SNPs			
	None	↑Folate	↓Methylation[42,44]	chr7: 127, 881, 035-127, 881, 054
	None	↑Betaine	↓Methylation[42]	chr7: 127, 881, 035-127, 881, 054
	None	↑Folic acid	↑Methylation[42]	chr7: 127, 881, 035-127, 881, 054
	CpG island; CEBP binding site; 2 reported SNPs (rs791620, rs2167270)	Famine	↑Methylation[47]	chr7: 127, 881, 054-127, 881, 410 (4)[a] (6)[b]
	CpG island; CEBP binding site; 2 reported SNPs (rs791620, rs2167270)	↑Methylation	↑Small for gestational age[94]	chr7: 127, 881, 127-127, 881, 350 (4)[a] (6)[b]
	CpG island; 1 reported SNP (rs2167270)	↓Methylation	↑BMI[100]	chr7: 127, 881, 280-127, 881, 300 (2)[a] (3)[b]
	CpG island; CEBP binding site; 2 reported SNPs (rs791620, rs2167270)	↓Methylation	↓BMI; ↑hip circumference[98]	chr7: 127, 881, 126-127, 881, 474 (3)[a] (4)[b]
	CpG island; CEBP binding site; 2 reported SNPs (rs791620, rs2167270)	↑Methylation	↑Fasting LDL-C[98]	chr7: 127, 881, 126-127, 881, 474 (3)[a] (4)[b]
	CpG island	↓Methylation	↑BMI[99]	chr7: 127, 881, 036 -127, 881, 057
	CpG island; CEBP binding site; 2 reported SNPs (rs791620, rs2167270)	↑Methylation	↑Lethargy and hypotonicity[105]	chr7: 127, 881, 127-127, 881, 350 (4)[a] (6)[b]
	CpG island; CEBP binding site; 2 reported SNPs (rs791620, rs2167270)	↓Methylation	↓HDL[101]	chr7: 127, 881, 053-127, 881, 410 (4)[a] (6)[b]
LY6E (Lymphocyte Antigen 6 Family Member E)	Multiple TFs binding sites; Promoter mark; Active Enhancer mark	↑Folate	↓Methylation[61]	chr8: 144, 120, 106-144, 120, 706 (8)[a] (9)[b]
MEG3 (Maternally Expressed 3) *(GTL-2)*	CpG island; Promoter mark	↑Vitamin B6	↑Methylation[54]	chr14: 101, 294, 220-101, 294, 391
	CpG island; Promoter mark	↑Folate	↓Methylation[56]	chr14: 101, 294, 220-101, 294, 391
	Enhancer and Promoter marks; CpG island; *POLR2A* binding site	↑UNIMMAP (supplementation)	↓Methylation[53]	chr14: 101, 292, 283– 101, 292, 796 (4)[a] (5)[b]
	CpG island; Promoter mark	↓Methylation	↑Birthweight[56]	chr14: 101, 294, 220-101, 294, 391
	None	Famine	↑Methylation[47]	chr14: 101, 291, 413-101, 291, 642 (5)[a] (6)[b]
MEST (Mesoderm-Specific Transcript) *(PEG1)*	CpG island	↑UNIMMAP (supplementation)	↓Methylation[53]	chr7: 130, 131, 325-130, 131, 792 (11)[a] (9)[b]
	Multiple TFs binding sites; Promoter mark; Enhancer mark; CpG island	↑Methylation	↑Small for gestational age[93]	chr7: 130, 125, 200-130, 126, 400 (16)[a] (17)[b]
MIRLET7BHG (MicroRNA Let-7b Host Gene)	Active Enhancer mark	↑Omega-6 PUFA	↓Methylation[66]	chr22: 46, 473, 721 (1)[a,b]
	Active Enhancer mark	↓Methylation	↑Birthweight[66]	chr22: 46, 473, 721 (1)[a,b]

(continued)

Table 4. Continued

Gene/region of Interest Blue = ME Brown = imprinted Yellow = ME and imprinted	Genomic features[c]	Exposure (↑/↓: increased/decreased)	Outcome (↑/↓: increased/decreased)	Coordinates of ROI in studies[d,e] (number of CpGs on 450k[a] and EPIC[b] arrays)
NR3C1 (Nuclear Receptor Subfamily 3 Group C Member 1) (*GR*)	Multiple TFs binding sites; Promoter mark; Enhancer mark; CpG island; 2 reported SNPs (rs10482604, rs10482605)	↑Methylation	↑Risk of being in a poorly regulated neurobehavioural profile[103,104]	chr5: 142, 783, 501-142, 783, 640 (4)[a,b]
	Multiple TFs binding sites; Promoter mark; Enhancer mark; CpG island; 2 reported SNPs (rs10482604, rs10482605)	↑Choline	↑Methylation[60]	chr5: 142, 783, 501-142, 783, 908 (5)[a] (7)[b]
	Multiple TFs binding sites; Promoter mark; Enhancer mark; CpG island	↑Methylation	↑Adult waist circumference, ↑adult BMI[62]	chr5: 142, 782, 759-142, 783, 164 (2)[ab]
	Multiple TFs binding sites; Promoter mark; Enhancer mark; CpG island; 1 reported SNP (rs10482604)	↑Meat/fish and vegetable intake, ↓bread/potato intake in late pregnancy	↑Methylation[62]	chr5: 142, 783, 579-142, 783, 714 (3)[a,b]
	Multiple TFs binding sites; Promoter mark; Enhancer mark; CpG island; 1 reported SNP (rs10482604)	↑Methylation	↓Adult blood pressure[62]	chr5: 142, 783, 578 -142, 783, 714 (3)[a,b]
OSBPL5/MRGPRG (Oxysterol-Binding Protein Like 5/MAS Related GPR Family Member G)	Enhancer mark; CpG island	Famine	↓Methylation[48]	chr11: 3, 225, 076 (1)[a,b]
OTX2 (Orthodenticle Homeobox 2)	CpG island; EZH2 binding site	↑Folate	↓Methylation[50]	chr14: 57, 278, 729 (1)[a,b]
PAX8 (Paired Box[8])	Multiple TFs binding sites; Promoter mark; Active Enhancer mark	Rainy season conception	↑Methylation[30]	chr2: 113, 993, 262-113, 993, 391(2)[a,b]
	Multiple TFs binding sites; Promoter mark; Active enhancer mark	Famine	↑Methylation[49]	chr2: 113, 992, 866-113, 993, 036(2)[a,b] chr2: 113, 992, 762-113, 993, 313 (8)[a] (7)[b]
PEG3 (Paternally Expressed 3)	Multiple TFs binding sites; 2 CpG islands; 1 reported SNP (rs2302376)	↑Folate	↓Methylation[56]	chr19: 57, 351, 945-57, 352, 096 (4)[a] (3)[b]
	Multiple TFs binding sites; 2 CpG islands; 1 reported SNP (rs2302376)	↑Folic acid	↓Methylation[56]	chr19: 57, 351, 945-57, 352, 096 (4)[a] (3)[b]
	Multiple TFs binding sites; 2 CpG islands; 1 reported SNP (rs2302376)	↑Folic acid	↓Methylation[58]	chr19: 57, 351, 944-57, 352, 096 (4)[a] (3)[b]
	Multiple TFs binding sites; 2 CpG islands; 1 reported SNP (rs2302376)	↑Folate	↓Methylation[56]	chr6: 144, 329, 109-144, 329, 231 (1)[a,b]

(continued)

Table 4. Continued

Gene/region of Interest Blue = ME Brown = imprinted Yellow = ME and imprinted	Genomic features[c]	Exposure (↑/↓: increased/decreased)	Outcome (↑/↓: increased/decreased)	Coordinates of ROI in studies[d,e] (number of CpGs on 450k[a] and EPIC[b] arrays)
PLAGL1 (PLAG1-Like Zinc Finger 1) (ZAC1)	Multiple TFs binding sites; Promoter mark; Active Enhancer mark; CpG island			chr6: 144, 329, 109-144, 329, 231 (1)[a,b]
	Multiple TFs binding sites; Promoter mark; Active Enhancer mark; CpG island	↑Methylation	↑Birthweight[56]	chr6: 144, 329, 390-144, 329, 740 (4)[a,b]
	Multiple TFs binding sites; Promoter mark; CpG island	↑Methylation index	↑Fetal weight at 32 weeks of gestation, weight and BMI at 1 year[40]	chr6: 144, 329, 390-144, 329, 740 (4)[a,b]
	Multiple TFs binding sites; Promoter mark; CpG island	↑ Vitamin B2	↑Methylation index[40]	chr6: 144, 329, 390-144, 329, 740 (4)[a,b]
POMC (Proopiomelanocortin)	Multiple TFs binding sites; Promoter mark; Active Enhancer mark; CpG island	↑Methylation	↑BMI[43,96]	chr2: 25, 384, 508-25, 384, 832 (3)[a,b]
	Multiple TFs binding sites; Promoter mark; Active Enhancer mark; CpG island	↑SAM:SAH ratio; ↑betaine	↑Methylation[43]	chr2: 25, 384, 508-25, 384, 832 (3)[a,b]
PPAP2C (PLPP2) (Phosphatidic Acid Phosphatase 2c)	CpG island	Famine	↑Methylation[48]	chr19: 292, 167 (1)[a,b]
PRDM9 (PR-Domain Containing Protein 9)	Multiple transcription factor binding sites; Promoter mark, Active enhancer mark; 2 reported SNPs (rs10077095, rs1994929)	Famine	↓Methylation[49]	chr5: 23, 507, 030-23, 507, 752 (12)[a](11)[b]
RBM46 (RNA-Binding Motif Protein 46)	CpG island	Rainy season conception	↑Methylation[31]	chr4: 155, 702, 818-155, 703, 110 (1)[a,b]
RXRA (Retinoid X Receptor Alpha)	Multiple TFs binding sites; Enhancer mark	↑Methylation	↑Fat mass; % fat mass[55]	chr9: 137, 215, 697 -137, 216, 117 (1)[a,b]
	Multiple TFs binding sites; Enhancer mark	↑Methylation	↑BMI[55]	chr9: 137, 215, 697 -137, 216, 117 (1)[a,b]
	Multiple TFs binding sites; Enhancer mark	↑Carbohydrate intake	↓Methylation[55]	chr9: 137, 215, 697 -137, 216, 117 (1)[a,b]
	Multiple TFs binding sites; Enhancer mark	↑Methylation	↓Bone mineral content; % BMC[75]	chr9: 137, 215, 697 -137, 216, 117 (1)[a,b]
		↑Folate	↓Methylation[42]	chr9: 137, 217, 097-137, 217, 132

(continued)

Table 4. Continued

Gene/region of Interest Blue = ME Brown = imprinted Yellow = ME and imprinted	Genomic features[c]	Exposure (↑/↓: increased/decreased)	Outcome (↑/↓: increased/decreased)	Coordinates of ROI in studies[d,e] (number of CpGs on 450k[a] and EPIC[b] arrays)
	Multiple TFs binding sites; Promoter mark; Active Enhancer mark; CpG island	↑Folate	↑Methylation[44]	chr9: 137, 217, 097-137, 217, 132
SLC38A2 (Solute Carrier Family 38 Member 2)	Multiple TFs binding sites; Promoter mark; Active Enhancer mark; CpG island Enhancer mark	Famine	↑Methylation[48]	chr12: 46, 737, 123 (1)[a,b]
SLITRK1 (SLIT And NTRK-like Family Member 1)	Promoter mark; Enhancer mark; CpG island	Rainy season conception	↑Methylation[30]	chr13: 84, 453, 741-84, 453, 828 chr13: 84, 454, 210-84, 454, 281
SPATC1L (C21orf56) (Spermatogenesis And Centriole Associated 1 Like)	Multiple TFs binding sites; Promoter mark; Active Enhancer mark	↑Folate	↓Methylation[61]	chr21: 47, 604, 052-47, 604, 654 (5)[a,b]
STX11 (Syntaxin 11)	Multiple TFs binding sites; Promoter mark; CpG island	↑Folate	↓Methylation[50]	chr6: 144, 471, 564 (1)[a,b]
TACC1 (Transforming Acidic Coiled-Coil Containing Protein 1)	Promoter mark; Enhancer mark	Famine	↑Methylation[48]	chr8: 38, 586, 183 (1)[a,b]
TFAP2A (Transcription Factor AP-2 Alpha)	E2F1 and EZH2 binding site; Promoter mark; Active Enhancer mark; CpG island	↑Folate	↓Methylation[50]	chr6: 10, 411, 911 (1)[a,b]
TMEM105 (Transmembrane Protein 105)	Enhancer mark; Active Enhancer mark; CpG island	Famine	↓Methylation[48]	chr17: 79, 283, 915 (1)[a,b]
VTRNA2-1 (Vault RNA 2-1)	Multiple TFs binding sites; Promoter mark; Active Enhancer mark; CpG island	Rainy Season; ↑vitamin B2; ↑methionine; ↓dimethylglycine	↑Methylation[41]	chr5: 135, 415, 762-135, 416, 613 (15)[a] (13)[b]
WNT9A (Wnt Family Member 9A)	NRF1 binding site; Promoter mark; Active Enhancer mark; CpG island	↑Folate	↑Methylation[61]	chr1: 228, 075, 423-228, 075, 749 (5)[a] (3)[b]
ZFP57 (Zinc Finger Protein 57)	YY1 binding site; Promoter mark; Active Enhancer mark; multiple reported SNPs	↑Folate	↓Methylation[61]	chr6: 29, 648, 161-29, 649, 084 (24)[a] (25)[b]
	Promoter mark; Active Enhancer mark; multiple reported SNPs	Famine	↓Methylation[49]	chr6: 29, 648, 345-29, 649, 024 (19)[a] (18)[b]
ZFYVE28 (Zinc Finger FYVE-Type Containing 28)	Multiple TFs binding sites; Promoter mark; CpG island	Rainy season conception	↑Methylation[30]	chr4: 2, 366, 658-2, 366, 739 (1)[a,b] chr4: 2, 366, 909-2, 367, 003

(continued)

by 10-11 translocation (TET) dioxygenases that use vitamin C (ascorbate) as a co-factor.[128] 5fC can then either be further oxidized to 5-carboxylcytosine or converted to an unmethylated cytosine by base excision repair. Adding vitamin C to mouse or human embryonic stem cells *in vitro* increases the activity of TET enzymes, resulting in active demethylation in the germline.[129] However, to our knowledge there have been no human *in vivo* studies exploring effects of periconceptional vitamin C deficiency on offspring DNA methylation.

Nutritional compounds do not act in isolation, and ideally analyses should recognize this by considering their interactions in metabolic pathways. For example, one-carbon metabolism is governed by intricately controlled feedback loops which help protect the flux of metabolites, through key reactions over a range of nutrient and co-factor concentrations.[130,131] This means that associations between individual micronutrients and methylation (e.g. the commonly analysed methyl donors folate and betaine) can disappear after adjustment for other metabolites (e.g. SAM and DMG, which can inhibit transmethylation reaction rates). Advances in measurement technology that allow the measurement of a greater range of nutritional biomarkers (e.g. metabolomics), combined with more sophisticated analytical techniques,[132,133] should enable a more nuanced understanding of the ways in which nutritional biomarkers combine to jointly influence methylation.

Measuring DNA methylation

A single CpG site in a single cell is either methylated or unmethylated, but measurements are typically made at the tissue level where methylation is a quantitative measure corresponding to the proportion of methylated cells.[134] Accurate assessment of tissue-level DNA methylation patterns presents a challenge, given the sensitivity of the measurements to both technical and biological variation. The advent of high-throughput, genome-wide microarray platforms, such as the Illumina HumanMethylation 450 K and EPIC arrays,[135–137] has helped in this regard, first by helping to standardize aspects of epigenome-wide association study (EWAS) design, and second by reducing the cost of genome-wide methylation assays required for adequately powered large studies.

Microarray-based EWAS have a number of limitations. First, by design, only a small proportion of the methylome is interrogated. These platforms attempt to include CpGs sites from all annotated genes, but the number of CpG sites per gene is low and equal coverage is typically not given to all genomic features and/or CpG contexts, with the focus having traditionally been on sites in promoters and CpG islands. Second, arrays provide no information on

Table 4. Continued

Gene/region of Interest Blue = ME Brown = imprinted Yellow = ME and imprinted	Genomic features[c]	Exposure (↑/↓: increased/decreased)	Outcome (↑/↓: increased/decreased)	Coordinates of ROI in studies[d,e] (number of CpGs on 450k[a] and EPIC[b] arrays)
ZNF385A (Zinc Finger Protein 385A)	Multiple TFs binding sites; Promoter mark; CpG island	Famine	↑Methylation[48]	chr12: 54, 764, 265 (1)[a,b]

LBW, low birthweight; LDL-C, low-density lipoprotein cholesterol; ME, metastable epiallele; ROI, region of interest; SAH, s-adenosyl homocysteine; SAM, s-adenosyl methionine; UNIMMAP, United Nations International Multiple Micronutrient Preparation.

[a]Number of CpGs covered on Infinium HumanMethylation450K BeadChip array.
[b]Number of CpGs covered on Infinium MethylationEPIC array.
[c]The following regulatory features were checked: enhancer/promoter marks (histone), overlapping binding sites for various transcription factors (e.g. CTCF, POL2A etc.) within region of interest (ROI) and presence of nearby reported GWAS single nucleotide polymorphisms (SNPs).
[d]Coordinates based on genome build hg19. The BiSearch Web server[122] was used to find genomic coordinates for ROIs where only primers were available.
[e]HumanMethylation450 v1.2 and Infinium MethylationEPIC v1.0 B4 Manifest Files were referred to report ROI coverage on Illumina Infinium Methylation BeadChip arrays.
[f]A total of 37 probes from 450k array were found within the gene and considered for analysis.

sequence-level variation, which is known to influence methylation status.[138,139] Finally, bioinformatics and analytical expertise are required (as well as the necessary computational resources) to process and model the data, and to correct for batch and other technical effects, in order to obtain reliable, high-quality methylation profiles.[140] As an alternative, true genome-wide approaches such as whole-genome bisulphite sequencing (WGBS) are available which interrogate all ∼28 million CpG sites in the methylome, although this is currently prohibitively expensive for larger samples. Targeted high-resolution platforms[141,142] offer a potential compromise between coverage and cost, but their utility, convenience and cost-effectiveness for performing EWAS remain to be established. Given the importance of demethylation during periconceptional epigenetic remodelling, it may also be important to consider the oxidized forms of 5-methyl cytosine (e.g. 5-hydroxymethylcytosine) which occur as intermediate products in the demethylation pathway.[143]

Tissue specificity, confounding and stability of methylation across the life course

The tissue-specific nature of DNA methylation presents a major challenge for epigenetic association studies.[134,144] The majority of studies reported in this review are constrained to accessible tissues such as cord blood that may be unrelated to the phenotype of interest, and different tissues may be sensitive to different environmental exposures. In this case reference epigenomes from different tissues and cell types in both healthy and diseased individuals[145] may inform the choice of tissue as well as providing data for investigating the tissue specificity of identified signals. Where exposure-related effects occur during early embryonic development, before gastrulation, methylation changes may be concordant across multiple tissues,[146] so that methylation states in accessible tissues such as blood and buccal cells may serve as a proxy for methylation in the target tissue.

Furthermore, numerous biological factors may act as potential confounders, for example age, sex, smoking status and BMI. Tissue-specific methylation differences arising from cell type heterogeneity, notably in blood, can also act as confounders,[147] although there are well-established methods that can be used to correct for this.[147,148]

DNA sequence polymorphisms are also known to influence DNA methylation status and may confound observed associations.[149] Heritability of DNA methylation is estimated to be in the range of 18% to 37%.[150,151] Consistent with this, many studies have shown that methylation quantitative trait loci (mQTL)—genetic variants associated with methylation differences at the population level—are

widespread. To account for this, ideally high-throughput genotype data on the sample being studied should be used[149] but, if such data are unavailable, population-level reference mQTL data can be informative.[139]

Finally, methylation changes associated with an early-life exposure may change throughout the life course, with implications for their utility as biomarkers of exposure or predictors of later phenotype.[152–154] Depending on the research question, this may suggest the need to assess long-term stability of methylation at specific loci, through the collection of longitudinal samples.

Linking methylation changes to gene function

Many of the DNA methylation changes reported in studies covered in this review are small, often within the margins of error of the measuring technology, making it difficult to draw conclusions on their functional relevance.[155] Indeed, relatively few methylation studies measure gene expression. The link between DNA methylation and expression is complex, depending on genomic context (e.g. location with gene bodies, promoters and enhancers).[156] This could in part explain seemingly contradictory findings from different studies measuring associations at the same gene. Moreover, a change in methylation may influence transcription factor binding and the induction of a specific signalling pathway in order to observe a change in gene expression. To aid further understanding, future studies should therefore consider measuring transcription factor binding, markers of gene transcription (mRNA levels), and/or translation (protein levels), to better map the potential effects of DNA methylation differences on gene function.[157]

Capturing phenotypes

In this review we have focused on phenotypic outcomes most commonly considered in the DOHaD context. However, we do not wish to exclude the possibility that there may be a broader range of phenotypes that are implicated. For example, exposure to the Dutch Hunger Winter famine during pregnancy has been associated with a wide variety of offspring phenotypes, varying according to the timing of famine exposure during gestation.[45,47] Consideration of the 'thrifty epigenotype' hypothesis[24] would suggest that famine-imposed epigenetic modifications in early life are adaptive where similar environment conditions persist, but maladaptive otherwise. There could therefore be a spectrum of phenotypes according to how great the mismatch is between *in utero* and later life environments. In the case of complex traits such as obesity, the resultant phenotype may also be influenced by factors such

as diet and lifestyle in conjunction with methylation differences and genotype of the individual.[158]

Causal inference

A major goal of nutritional epigenetic studies, also covered in this review, is to assess the potential for epigenetic marks to mediate links between nutritional exposures and health outcomes. In this context, the use of prospective study designs with randomization including negative controls, and techniques such as mediation analysis based on regression systems,[159] structural equation modelling[160] or network-based techniques,[161] parametric/semi-parametric methods,[162] or instrumental variable approaches such as Mendelian randomization,[80,163,164] can help to strengthen causal inference. More broadly, triangulating findings from diverse studies, each with their own strengths, limitations, assumptions and opposing biases, will maximize the potential for robust findings.[165,166]

Study design considerations

The literature in this area is dominated by observational studies. This increases the risk of spurious associations due to confounding or reverse causation,[149] the latter being a particular problem with methylation association studies where the direction of causality can be hard to establish. Added to this, effect sizes are generally modest, with group-level differences in mean methylation typically less than 10% and often in the region of 1–5% for many of the exposures and phenotypes studied.[155,167,168] This has implications for the design of studies characterizing genome-wide, population-level methylation differences, as they need to be adequately powered to detect potentially small effects after adjusting for multiple testing.[169]

Current interest in periconceptional nutrition has stimulated a number of preconceptional nutrition trials.[170–174] In these studies, supplementation before conception is necessary to ensure that the conception period is covered and that a maximal effect on maternal nutritional status at conception is achieved. Nonetheless, accurately pinpointing the timing of nutritional exposures to conception is challenging.

Conclusions

The body of evidence linking maternal nutritional exposure to offspring phenotype via DNA methylation in humans is rapidly growing yet currently remains complex and inconsistent. It is characterized by heterogeneous exposures and outcomes, and mainly observational associations that are frequently under-powered. Existing evidence suggests that the effect of nutritional exposures on DNA methylation depends on the form of the nutritional component, the timing of exposure during periconception and pregnancy, the underlying nutritional status of the mother, maternal and offspring genotype and the specific loci under investigation. The picture is more complex than methylation being determined simply by availability of methyl donors. Many studies have investigated imprinted genes as priority loci for their vulnerability to nutritional exposures, but with the adoption of microarray-based platforms, other candidate genes and gene classes are emerging, for example metastable epialleles.

The utility of this emerging evidence in terms of its translation into effective interventions and therapies remains an open question. Epigenetic marks like DNA methylation may act as integrators of multiple exposures and genetic risk factors, as well as molecular mediators of the effect of exposures on phenotype. Where robust associations are established, DNA methylation can serve as a proxy measure or biomarker of earlier nutritional exposures.[175] As mediators of the effect on later phenotype, nutritionally sensitive DNA methylation changes can provide a means to identify genes and pathways for targeted interventions. Whereas there is still much work to do in this area, there are grounds for optimism that epigenomic approaches will provide insights into the molecular basis of the developmental origins of health and disease, which could in turn lead to the development of next-generation interventions.

Supplementary data

Supplementary data are available at *IJE* online.

Funding

The EMPHASIS study group includes

Lena Acolatse, MRC Unit The Gambia at the London School of Hygiene and Tropical Medicine, The Gambia, Meraj Ahmed, Genomic Research on Complex diseases (GRC Group), CSIR-Centre for Cellular and Molecular Biology, Hyderabad, India, Modupeh Betts, MRC Unit The Gambia at the London School of Hygiene and Tropical Medicine, The Gambia, Giriraj R Chandak,

CSIR-Centre for Cellular and Molecular Biology, Hyderabad, India, Harsha Chopra, Centre for the Study of Social Change, Mumbai, India, Cyrus Cooper, MRC Life Course Epidemiology Unit, University of Southampton, UK, Momodou K Darboe, MRC Unit The Gambia at the London School of Hygiene and Tropical Medicine, The Gambia, Chiara Di Gravio, MRC Life Course Epidemiology Unit, University of Southampton, UK, Caroline HD Fall, MRC Life Course Epidemiology Unit, University of Southampton, UK, Meera Gandhi, Centre for the Study of Social Change, Mumbai, India, Gail R Goldberg, MRC Elsie Widdowson Laboratory, Cambridge, UK, Prachand Issarapu, Genomic Research on Complex diseases (GRC Group), CSIR-Centre for Cellular and Molecular Biology, Hyderabad, India, Philip James, MRC Unit The Gambia at the London School of Hygiene and Tropical Medicine, UK, Ramatoulie Janha, MRC Unit The Gambia at the London School of Hygiene and Tropical Medicine, The Gambia, Landing M A Jarjou, MRC Unit The Gambia at the London School of Hygiene and Tropical Medicine, The Gambia, Lovejeet Kaur, Genomic Research on Complex diseases (GRC Group), CSIR-Centre for Cellular and Molecular Biology, Hyderabad, India, Sarah H Kehoe, MRC Life Course Epidemiology Unit, University of Southampton, UK, Kalyanaraman Kumaran, MRC Life Course Epidemiology Unit, University of Southampton, UK and CSI Holdsworth Memorial Hospital, Mysore, India, Karen A Lillycrop, University of Southampton, UK, Mohammed Ngum, MRC Unit The Gambia at the London School of Hygiene and Tropical Medicine, The Gambia, Suraj S Nongmaithem, Genomic Research on Complex diseases (GRC Group), CSIR-Centre for Cellular and Molecular Biology, Hyderabad, India, Stephen Owens, Institute of Health and Society, Newcastle University, UK, Ramesh D Potdar, Centre for the Study of Social Change, Mumbai, India, Andrew M Prentice, MRC Unit The Gambia at the London School of Hygiene and Tropical Medicine, The Gambia, Ann Prentice, MRC Unit The Gambia, Elsie Widdowson Laboratory, Cambridge, UK and MRC Life Course Epidemiology Unit, University of Southampton, UK, Tallapragada Divya Sri Priyanka, Genomic Research on Complex diseases (GRC Group), CSIR-Centre for Cellular and Molecular Biology, Hyderabad, India, Ayden Saffari, MRC Unit The Gambia at the London School of Hygiene and Tropical Medicine, UK, Sirazul Ameen Sahariah, Centre for the Study of Social Change, Mumbai, India, Sara Sajjadi, Genomic Research on Complex diseases (GRC Group), CSIR-Centre for Cellular and Molecular Biology, Hyderabad, India, Harshad Sane, Centre for the Study of Social Change, Mumbai, India, Smeeta Shrestha, Genomic Research on Complex diseases (GRC Group), CSIR-Centre for Cellular and Molecular Biology, Hyderabad, India, Matt J Silver, MRC Unit The Gambia at the London School of Hygiene and Tropical Medicine, UK, Ashutosh Singh Tomar, Genomic Research on Complex diseases (GRC Group), CSIR-Centre for Cellular and Molecular Biology, Hyderabad, India, Kate A Ward, MRC Elsie Widdowson Laboratory, Cambridge and MRC Life course Epidemiology Unit, University of Southampton, UK, Dilip Kumar Yadav, Genomic Research on Complex diseases (GRC Group), CSIR-Centre for Cellular and Molecular Biology, Hyderabad, India, Chittaranjan S Yajnik, Diabetes Unit, KEM Hospital and Research Centre, Pune, India.

Conflict of interest: None declared.

References

1. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* 2003;**33(Suppl)**:245–54.
2. Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. *Nature* 2004;**429**: 457–63.
3. Esteller M. Epigenetics in cancer. *N Engl J Med* 2008;**358**: 1148–59.
4. Moore GE, Ishida M, Demetriou C *et al*. The role and interaction of imprinted genes in human fetal growth. *Philos Trans R Soc B Biol Sci* 2015;**370**:20140074.
5. Bernstein BE, Meissner A, Lander ES. The mammalian epigenome. *Cell* 2007;**128**:669–81.
6. Waterland RA, Michels KB. Epigenetic epidemiology of the developmental origins hypothesis. *Annu Rev Nutr* 2007;**27**:363–88.
7. Burdge GC, Lillycrop KA. Nutrition, epigenetics, and developmental plasticity: implications for understanding human disease. *Annu Rev Nutr* 2010;**30**:315–39.
8. Lillycrop KA, Burdge GC. Epigenetic mechanisms linking early nutrition to long term health. *Best Pract Res Clin Endocrinol Metab* 2012;**26**:667–76.
9. Waterland RA, Jirtle RL. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Mol Cell Biol* 2003;**23**:5293–300.
10. Vickers MH. Early life nutrition, epigenetics and programming of later life disease. *Nutrients* 2014;**6**:2165–78.
11. Messerschmidt DM, Knowles BB, Solter D. DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes Dev* 2014;**28**:812–28.
12. Wang Y, Leung FCC. An evaluation of new criteria for CpG islands in the human genome as gene markers. *Bioinformatics* 2004;**20**:1170–77.
13. Klose RJ, Bird AP. Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci* 2006;**31**:89–97.
14. Illingworth RS, Bird AP. CpG islands—'a rough guide'. *FEBS Lett* 2009;**583**:1713–20.
15. Guil S, Esteller M. DNA methylomes, histone codes and miRNAs: tying it all together. *Int J Biochem Cell Biol* 2009;**41**: 87–95.

16. Sato F, Tsuchiya S, Meltzer SJ, Shimizu K. MicroRNAs and epigenetics. *FEBS J* 2011;**278**:1598–609.

17. Barker DJ, Osmond C, Golding J, Kuh D, Wadsworth ME. Growth in utero, blood pressure in childhood and adult life, and mortality from cardiovascular disease. *BMJ* 1989;**298**:564–67.

18. Eriksson JG, Forsen T, Tuomilehto J, Winter PD, Osmond C, Barker DJP. Catch-up growth in childhood and death from coronary heart disease: longitudinal study. *BMJ* 1999;**318**:427–31.

19. Barker DJP, Osmond C, Kajantie E, Eriksson JG. Growth and chronic disease: findings in the Helsinki Birth Cohort. *Ann Hum Biol* 2009;**36**:445–58.

20. Li Y, Ley SH, Tobias DK et al. Birth weight and later life adherence to unhealthy lifestyles in predicting type 2 diabetes: prospective cohort study. *BMJ* 2015;**351**:h3672.

21. Roseboom TJ, van der Meulen JH, van Montfrans GA et al. Maternal nutrition during gestation and blood pressure in later life. *J Hypertens* 2001;**19**:29–34.

22. Ravelli GP, Stein ZA, Susser MW. Obesity in young men after famine exposure in utero and early infancy. *N Engl J Med* 1976;**295**:349–53.

23. Susser E, Neugebauer R, Hoek HW et al. Schizophrenia after prenatal famine. *Arch Gen Psychiatry* 1996;**53**:25–31.

24. Stöger R. The thrifty epigenotype: an acquired and heritable predisposition for obesity and diabetes? *BioEssays* 2008;**30**:156–66.

25. Langley-Evans SC. Nutrition in early life and the programming of adult disease: a review. *J Hum Nutr Diet* 2015;**28**:1–14.

26. Steegers-Theunissen RPM, Twigt J, Pestinger V, Sinclair KD. The periconceptional period, reproduction and long-term health of offspring: the importance of one-carbon metabolism. *Hum Reprod Update* 2013;**19**:640–55.

27. Smallwood SA, Kelsey G. De novo DNA methylation: a germ cell perspective. *Trends Genet* 2012;**28**:33–42.

28. Ishida M, Moore GE. The role of imprinted genes in humans. *Mol Aspects Med* 2013;**34**:826–40.

29. Rakyan VK, Blewitt ME, Druker R, Preis JI, Whitelaw E. Metastable epialleles in mammals. *Trends Genet* 2002;**18**:348–51.

30. Waterland RA, Kellermayer R, Laritsky E et al. Season of conception in rural Gambia affects DNA methylation at putative human metastable epialleles. *PLoS Genet* 2010;**6**:e1001252.

31. Dominguez-Salas P, Moore SE, Baker MS et al. Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. *Nat Commun* 2014;**5**:3746.

32. Hancks DC, Kazazian HH. Active human retrotransposons: variation and disease. *Curr Opin Genet Dev* 2012;**22**:191–203.

33. Slotkin RK, Martienssen R. Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* 2007;**8**:272–85.

34. Perera F, Herbstman J. Prenatal environmental exposures, epigenetics, and disease. *Reprod Toxicol* 2011;**31**:363–73.

35. Reul JMHM, Collins A, Saliba RS et al. Glucocorticoids, epigenetic control and stress resilience. *Neurobiol Stress* 2015;**1**:44–59.

36. Fowden AL, Forhead AJ. Hormones as epigenetic signals in developmental programming. *Exp Physiol* 2009;**94**:607–25.

37. Wolff GL, Kodell RL, Moore SR, Cooney CA. Maternal epigenetics and methyl supplements affect agouti gene expression in Avy/a mice. *FASEB J* 1998;**12**:949–57.

38. Jiménez-Chillarón JC, Díaz R, Martínez D et al. The role of nutrition on epigenetic modifications and their implications on health. *Biochimie* 2012;**94**:2242–63.

39. Soubry A. Epigenetic inheritance and evolution: a paternal perspective on dietary influences. *Prog Biophys Mol Biol* 2015;**118**:79–85.

40. Azzi S, Sas TCJ, Koudou Y et al. Degree of methylation of ZAC1 (PLAGL1) is associated with prenatal and post-natal growth in healthy infants of the EDEN mother child cohort. *Epigenetics* 2014;**9**:338–45.

41. Silver MJ, Kessler NJ, Hennig BJ et al. Independent genome-wide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. *Genome Biol* 2015;**16**:118.

42. Pauwels S, Ghosh M, Duca RC et al. Dietary and supplemental maternal methyl-group donor intake and cord blood DNA methylation. *Epigenetics* 2017;**12**:1–10.

43. Kühnen P, Handke D, Waterland RA et al. Interindividual variation in DNA methylation at a putative POMC metastable epiallele is associated with obesity. *Cell Metab* 2016;**24**:502–09.

44. Pauwels S, Ghosh M, Duca RC et al. Maternal intake of methyl-group donors affects DNA methylation of metabolic genes in infants. *Clin Epigenetics* 2017;**9**:16.

45. Heijmans BT, Tobi EW, Stein AD et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc Natl Acad Sci USA* 2008;**105**:17046–49.

46. Tobi EW, Slagboom PE, van Dongen J et al. Prenatal famine and genetic variation are independently and additively associated with DNA methylation at regulatory loci within IGF2/H19. *PLoS One* 2012;**7**:e37933.

47. Tobi EW, Lumey LH, Talens RP et al. DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Hum Mol Genet* 2009;**18**:4046–53.

48. Tobi EW, Slieker RC, Stein AD et al. Early gestation as the critical time-window for changes in the prenatal environment to affect the adult human blood methylome. *Int J Epidemiol* 2015;**44**:1211–23.

49. Finer S, Iqbal MS, Lowe R et al. Is famine exposure during developmental life in rural Bangladesh associated with a metabolic and epigenetic signature in young adulthood? A historical cohort study. *BMJ Open* 2016;**6**:e011768.

50. Gonseth S, Roy R, Houseman EA et al. Periconceptional folate consumption is associated with neonatal DNA methylation modifications in neural crest regulatory and cancer development genes. *Epigenetics* 2015;**10**:1166–76.

51. Hoyo C, Murtha AP, Schildkraut JM et al. Methylation variation at IGF2 differentially methylated regions and maternal folic acid use before and during pregnancy. *Epigenetics* 2011;**6**:928–36.

52. Steegers-Theunissen RP, Obermann-Borst SA, Kremer D et al. Periconceptional maternal folic acid use of 400 microg per day is related to increased methylation of the IGF2 gene in the very young child. *PLoS One* 2009;**4**:e7845.

53. Cooper WN, Khulan B, Owens S et al. DNA methylation profiling at imprinted loci after periconceptional micronutrient

supplementation in humans: results of a pilot randomized controlled trial. *FASEB J* 2012;**26**:1782–90.

54. McCullough LE, Miller EE, Mendez MA, Murtha AP, Murphy SK, Hoyo C. Maternal B vitamins: effects on offspring weight and DNA methylation at genomically imprinted domains. *Clin Epigenetics* 2016;**8**:8.

55. Godfrey KM, Sheppard A, Gluckman PD *et al*. Epigenetic gene promoter methylation at birth is associated with child's later adiposity. *Diabetes* 2011;**60**:1528–34.

56. Hoyo C, Daltveit AK, Iversen E *et al*. Erythrocyte folate concentrations, CpG methylation at genomically imprinted domains, and birth weight in a multiethnic newborn cohort. *Epigenetics* 2014;**9**:1120–30.

57. van Mil NH, Bouwland-Both MI, Stolk L *et al*. Determinants of maternal pregnancy one-carbon metabolism and newborn human DNA methylation profiles. *Reproduction* 2014;**148**: 581–92.

58. Haggarty P, Hoad G, Campbell DM, Horgan GW, Piyathilake C, McNeill G. Folate in pregnancy and imprinted gene and repeat element methylation in the offspring. *Am J Clin Nutr* 2013;**97**:94–99.

59. Ba Y, Yu H, Liu F *et al*. Relationship of folate, vitamin B12 and methylation of insulin-like growth factor-II in maternal and cord blood. *Eur J Clin Nutr* 2011;**65**:480–85.

60. Jiang X, Yan J, West AA *et al*. Maternal choline intake alters the epigenetic state of fetal cortisol-regulating genes in humans. *FASEB J* 2012;**26**:3563–74.

61. Amarasekera M, Martino D, Ashley S *et al*. Genome-wide DNA methylation profiling identifies a folate-sensitive region of differential methylation upstream of ZFP57-imprinting regulator in humans. *FASEB J* 2014;**28**:4068–76.

62. Drake AJ, McPherson RC, Godfrey KM *et al*. An unbalanced maternal diet in pregnancy associates with offspring epigenetic changes in genes controlling glucocorticoid action and foetal growth. *Clin Endocrinol* 2012;**77**:808–15.

63. Rijlaarsdam J, Cecil CAM, Walton E *et al*. Prenatal unhealthy diet, insulin-like growth factor 2 gene (IGF2) methylation, and attention deficit hyperactivity disorder symptoms in youth with early-onset conduct problems. *J Child Psychol Psychiatr* 2017;**58**:19–27.

64. Lee H-S, Barraza-Villarreal A, Biessy C *et al*. Dietary supplementation with polyunsaturated fatty acid during pregnancy modulates DNA methylation at IGF2/H19 imprinted genes and growth of infants. *Physiol Genomics* 2014;**46**:851–57.

65. Dijk SJ. V, Zhou J, Peters TJ *et al*. Effect of prenatal DHA supplementation on the infant epigenome: results from a randomized controlled trial. *Clin Epigenetics* 2016;**8**:114.

66. Lin X, Lim IY, Wu Y *et al*. Developmental pathways to adiposity begin before birth and are influenced by genotype, prenatal environment and epigenome. *BMC Med* 2017;**15**:50.

67. McKay JA, Groom A, Potter C *et al*. Genetic and non-genetic influences during pregnancy on infant global and site specific DNA methylation: role for folate gene variants and vitamin B12. *PLoS One* 2012;**7**:e33290.

68. Joubert BR, Dekker HT. D, Felix JF *et al*. Maternal plasma folate impacts differential DNA methylation in an epigenome-wide meta-analysis of newborns. *Nat Commun* 2016;**7**:10577.

69. Lee H-S, Barraza-Villarreal A, Hernandez-Vargas H *et al*. Modulation of DNA methylation states and infant immune

system by dietary supplementation with ω-3 PUFA during pregnancy in an intervention study. *Am J Clin Nutr* 2013;**98**: 480–87.

70. Amarasekera M, Noakes P, Strickland D, Saffery R, Martino DJ, Prescott SL. Epigenome-wide analysis of neonatal CD4 + T-cell DNA methylation sites potentially affected by maternal fish oil supplementation. *Epigenetics* 2014;**9**:1570–76.

71. Lumey LH, Stein AD, Kahn HS *et al*. Cohort Profile: The Dutch Hunger Winter families study. *Int J Epidemiol* 2007;**36**: 1196–204.

72. Prentice A, Whitehead R, Roberts S, Paul A. Long-term energy balance in child-bearing Gambian women. *Am J Clin Nutr* 1981;**34**:2790–99.

73. Fang M, Chen D, Yang CS. Dietary polyphenols may affect DNA methylation. *J Nutr* 2007;**137**(**Suppl 1**):223–28S.

74. Pereira F, Barbáchano A, Singh PK, Campbell MJ, Muñoz A, Larriba MJ. Vitamin D has wide regulatory effects on histone demethylase genes. *Cell Cycle* 2012;**11**:1081–89.

75. Harvey NC, Sheppard A, Godfrey KM *et al*. Childhood bone mineral content is associated with methylation status of the RXRA promoter at birth. *J Bone Miner Res* 2014;**29**:600–07.

76. Feng Y, Zhao L-Z, Hong L, Shan C, Shi W, Cai W. Alteration in methylation pattern of GATA-4 promoter region in vitamin A-deficient offspring's heart. *J Nutr Biochem* 2013;**24**: 1373–80.

77. Babenko O, Kovalchuk I, Metz GAS. Stress-induced perinatal and transgenerational epigenetic programming of brain development and mental health. *Neurosci Biobehav Rev* 2015;**48**: 70–91.

78. Anway MD, Skinner MK. Epigenetic transgenerational actions of endocrine disruptors. *Endocrinology* 2006;**147**(**Suppl 6**): S43–49.

79. Hajj NE, Schneider E, Lehnen H, Haaf T. Epigenetics and life-long consequences of an adverse nutritional and diabetic intra-uterine environment. *Reproduction* 2014;**148**:R111–20.

80. Sharp GC, Lawlor DA, Richmond RC *et al*. Maternal pre-pregnancy BMI and gestational weight gain, offspring DNA methylation and later offspring adiposity: findings from the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* 2015;**44**:1288–304.

81. Azzi S, Brioude F, Bouc YL, Netchine I. Human imprinting anomalies in fetal and childhood growth disorders: clinical implications and molecular mechanisms. *Curr Pharm Des* 2014;**20**:1751–63.

82. Burris HH, Baccarelli AA, Byun H-M *et al*. Offspring DNA methylation of the aryl-hydrocarbon receptor repressor gene is associated with maternal BMI, gestational age, and birth weight. *Epigenetics* 2015;**10**:913–21.

83. Bouwland-Both MI, van Mil NH, Stolk L *et al*. DNA methylation of IGF2DMR and H19 is associated with fetal and infant growth: the generation R study. *PLoS One* 2013;**8**:e81731.

84. Einstein F, Thompson RF, Bhagat TD *et al*. Cytosine methylation dysregulation in neonates following intrauterine growth restriction. *PLoS One* 2010;**5**:e8887.

85. Toure DM, Baccaglini L, Opoku ST *et al*. Epigenetic dysregulation of insulin-like growth factor (IGF)-related genes and adverse pregnancy outcomes: a systematic review. *J Matern Fetal Neonatal Med* 2016;**18**:1–11.

86. Davie JR. Inhibition of histone deacetylase activity by butyrate. *J Nutr* 2003;**133**(Suppl 7):2485–93S.

87. Claycombe KJ, Brissette CA, Ghribi O. Epigenetics of inflammation, maternal infection, and nutrition. *J Nutr* 2015;**145**: 1109–15S.

88. Burdge GC, Hanson MA, Slater-Jefferies JL, Lillycrop KA. Epigenetic regulation of transcription: a mechanism for inducing variations in phenotype (fetal programming) by differences in nutrition during early life? *Br J Nutr* 2007;**97**:1036.

89. Kucharski R, Maleszka J, Foret S, Maleszka R. Nutritional control of reproductive status in honeybees via DNA methylation. *Science* 2008;**319**:1827.

90. Hoyo C, Fortner K, Murtha AP *et al*. Association of cord blood methylation fractions at imprinted insulin-like growth factor 2 (IGF2), plasma IGF2, and birth weight. *Cancer Causes Control* 2012;**23**:635–45.

91. St-Pierre J, Hivert M-F, Perron P *et al*. IGF2 DNA methylation is a modulator of newborn's fetal growth and development. *Epigenetics* 2012;**7**:1125–32.

92. Qian Y-Y, Huang X-L, Liang H *et al*. Effects of maternal folic acid supplementation on gene methylation and being small for gestational age. *J Hum Nutr Diet* 2016;**29**:643–51.

93. Kappil MA, Green BB, Armstrong DA *et al*. Placental expression profile of imprinted genes impacts birth weight. *Epigenetics* 2015;**10**:842–49.

94. Lesseur C, Armstrong DA, Paquette AG, Koestler DC, Padbury JF, Marsit CJ. Tissue-specific leptin promoter DNA methylation is associated with maternal and infant perinatal factors. *Mol Cell Endocrinol* 2013;**381**:160–67.

95. Murphy R, Thompson JM, Tost J, Mitchell EA; Auckland Birthweight Collaborative Study Group. No evidence for copy number and methylation variation in H19 and KCNQ10T1 imprinting control regions in children born small for gestational age. *BMC Med Genet* 2014;**15**:67.

96. Kuehnen P, Mischke M, Wiegand S *et al*. An Alu element-associated hypermethylation variant of the POMC gene is associated with childhood obesity. *PLoS Genet* 2012;**8**:e1002543.

97. Huang R-C, Galati JC, Burrows S *et al*. DNA methylation of the IGF2/H19 imprinting control region and adiposity distribution in young adults. *Clin Epigenetics* 2012;**4**:21.

98. Houde A-A, Légaré C, Biron S *et al*. Leptin and adiponectin DNA methylation levels in adipose tissues and blood cells are associated with BMI, waist girth and LDL-cholesterol levels in severely obese men and women. *BMC Med Genet* 2015;**16**:29.

99. Dunstan J, Bressler JP, Moran TH *et al*. Associations of LEP, CRH, ICAM-1, and LINE-1 methylation, measured in saliva, with waist circumference, body mass index, and percent body fat in mid-childhood. *Clin Epigenet* 2017;**9**:29.

100. García-Cardona MC, Huang F, García-Vivas JM *et al*. DNA methylation of leptin and adiponectin promoters in children is reduced by the combined presence of obesity and insulin resistance. *Int J Obes* 2014;**38**:1457–65.

101. Wijnands KPJ, Obermann-Borst SA, Steegers-Theunissen RPM. Early life lipid profile and metabolic programming in very young children. *Nutr Metab Cardiovasc Dis* 2015;**25**: 608–14.

102. Deodati A, Inzaghi E, Liguori A *et al*. IGF2 methylation is associated with lipid profile in obese children. *Horm Res Paediatr* 2013;**79**:361–67.

103. Paquette AG, Lester BM, Lesseur C *et al*. Placental epigenetic patterning of glucocorticoid response genes is associated with infant neurodevelopment. *Epigenomics* 2015;**7**:767–79.

104. Lester BM, Marsit CJ, Giarraputo J, Hawes K, LaGasse LL, Padbury JF. Neurobehavior related to epigenetic differences in preterm infants. *Epigenomics* 2015;**7**:1123–36.

105. Lesseur C, Armstrong DA, Murphy MA *et al*. Sex-specific associations between placental leptin promoter DNA methylation and infant neurobehavior. *Psychoneuroendocrinology* 2014; **40**:1–9.

106. Reik W, Walter J. Genomic imprinting: parental influence on the genome. *Nat Rev Genet* 2001;**2**:21–32.

107. Nordin M, Bergman D, Halje M, Engström W, Ward A. Epigenetic regulation of the Igf2/H19 gene cluster. *Cell Prolif* 2014;**47**:189–99.

108. Piedrahita JA. The role of imprinted genes in fetal growth abnormalities. *Birth Defects Res Part A Clin Mol Teratol* 2011; **91**:682–92.

109. Azzi S, Rossignol S, Steunou V *et al*. Multilocus methylation analysis in a large cohort of 11p15-related foetal growth disorders (Russell Silver and Beckwith Wiedemann syndromes) reveals simultaneous loss of methylation at paternal and maternal imprinted loci. *Hum Mol Genet* 2009;**18**:4724–33.

110. Liu Y, Murphy SK, Murtha AP *et al*. Depression in pregnancy, infant birth weight and DNA methylation of imprint regulatory elements. *Epigenetics* 2012;**7**:735–46.

111. Tobi EW, Heijmans BT, Kremer D *et al*. DNA methylation of IGF2, GNASAS, INSIGF and LEP and being born small for gestational age. *Epigenetics* 2011;**6**:171–76.

112. Karbiener M, Glantschnig C, Pisani DF *et al*. Mesoderm-specific transcript (MEST) is a negative regulator of human adipocyte differentiation. *Int J Obes* 2015;**39**:1733–41.

113. McMinn J, Wei M, Schupf N *et al*. Unbalanced placental expression of imprinted genes in human intrauterine growth restriction. *Placenta* 2006;**27**:540–49.

114. Hajj NE, Pliushch G, Schneider E *et al*. Metabolic programming of MEST DNA methylation by intrauterine exposure to gestational diabetes mellitus. *Diabetes* 2013;**62**:1320–28.

115. Ahuja HS, Szanto A, Nagy L, Davies PJA, The retinoid X receptor and its ligands: versatile regulators of metabolic function, cell differentiation and cell death. *J Biol Regul Homeost Agents* 2003;**17**:29–45.

116. Yee YK, Chintalacharuvu SR, Lu J, Nagpal S. Vitamin D receptor modulators for inflammation and cancer. *Mini Rev Med Chem* 2005;**5**:761–78.

117. Barker DJ, Gluckman PD, Godfrey KM, Harding JE, Owens JA, Robinson JS. Fetal nutrition and cardiovascular disease in adult life. *Lancet* 1993;**341**:938–41.

118. Barker DJP. Fetal growth and adult disease. *Br J Obstet Gynaecol* 1992;**99**:275–76.

119. Vickers MH. Developmental programming and adult obesity: the role of leptin. *Curr Opin Endocrinol Diabetes Obes* 2007; **14**:17–22.

120. Després J-P, Lemieux I, Bergeron J *et al*. Abdominal obesity and the metabolic syndrome: contribution to global cardiometabolic risk. *Arterioscler Thromb Vasc Biol* 2008;**28**:1039.

121. Harris A, Seckl J. Glucocorticoids, prenatal stress and the programming of disease. *Horm Behav* 2011;**59**:279–89.

122. Arányi T, Tusnády GE. BiSearch: ePCR tool for native or bisulfite-treated genomic template. *Methods Mol Biol* 2007; **402**:385–402.

123. Kroke A, Klipstein-Grobusch K, Voss S *et al*. Validation of a self-administered food-frequency questionnaire administered in the European Prospective Investigation into Cancer and Nutrition (EPIC) Study: comparison of energy, protein, and macronutrient intakes estimated with the doubly labeled water, urinary nitrogen, and repeated 24-h dietary recall methods. *Am J Clin Nutr* 1999;**70**:439–47.

124. Black AE, Prentice AM, Goldberg GR *et al*. Measurements of total energy expenditure provide insights into the validity of dietary measurements of energy intake. *J Am Diet Assoc* 1993; **93**:572–79.

125. Biró G, Hulshof KFAM, Ovesen L, Amorim Cruz JA; EFCOSUM Group. Selection of methodology to assess food intake. *Eur J Clin Nutr* 2002;**56**(**Suppl 2**):S25–32.

126. Faupel-Badger JM, Hsieh C-C, Troisi R, Lagiou P, Potischman N. Plasma volume expansion in pregnancy: implications for biomarkers in population studies. *Cancer Epidemiol Biomarkers Prev* 2007;**16**:1720–23.

127. Wallace JM, Bonham MP, Strain J *et al*. Homocysteine concentration, related B vitamins, and betaine in pregnant women recruited to the Seychelles Child Development Study. *Am J Clin Nutr* 2008;**87**:391–97.

128. Young JI, Züchner S, Wang G. Regulation of the epigenome by vitamin C. *Annu Rev Nutr* 2015;**35**:545–64.

129. Blaschke K, Ebata KT, Karimi MM *et al*. Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature* 2013;**500**:222–26.

130. Reed MC, Gamble MV, Hall MN, Nijhout HF. Mathematical analysis of the regulation of competing methyltransferases. *BMC Syst Biol* 2015;**9**:69.

131. Nijhout HF, Best J, Reed MC. Escape from homeostasis. *Math Biosci* 2014;**257**:104–10.

132. Aurich MK, Thiele I. Computational modeling of human metabolism and its application to systems biomedicine. *Methods Mol Biol* 2016;**1386**:253–81.

133. Thiele I, Swainston N, Fleming RMT *et al*. A community-driven global reconstruction of human metabolism. *Nat Biotechnol* 2013;**31**:419–25.

134. Michels KB, Binder AM, Dedeurwaerder S *et al*. Recommendations for the design and analysis of epigenome-wide association studies. *Nat Methods* 2013;**10**:949–55.

135. Bibikova M, Le J, Barnes B *et al*. Genome-wide DNA methylation profiling using Infinium® assay. *Epigenomics* 2009;**1**: 177–200.

136. Bibikova M, Barnes B, Tsan C *et al*. High density DNA methylation array with single CpG site resolution. *Genomics* 2011;**98**: 288–95.

137. Pidsley R, Zotenko E, Peters TJ *et al*. Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biol* 2016;**17**:208.

138. Wagner JR, Busche S, Ge B, Kwan T, Pastinen T, Blanchette M. The relationship between DNA methylation, genetic and expression inter-individual variation in untransformed human fibroblasts. *Genome Biol* 2014;**15**:R37.

139. Gaunt TR, Shihab HA, Hemani G *et al*. Systematic identification of genetic influences on methylation across the human life course. *Genome Biol* 2016;**17**:61.

140. Morris TJ, Beck S. Analysis pipelines and packages for Infinium HumanMethylation450 BeadChip (450k) data. *Methods* 2015; **72**:3–8.

141. Li Q, Suzuki M, Wendt J *et al*. Post-conversion targeted capture of modified cytosines in mammalian and plant genomes. *Nucleic Acids Res* 2015;**43**:e81.

142. Teh AL, Pan H, Lin X *et al*. Comparison of methyl-capture sequencing vs. infinium 450K methylation array for methylome analysis in clinical samples. *Epigenetics* 2016;**11**:36–48.

143. Ulahannan N, Greally JM. Genome-wide assays that identify and quantify modified cytosines in human disease studies. *Epigenetics Chromatin* 2015;**8**:5.

144. Roadmap Epigenomics Consortium: Kundaje A, Meuleman W, Ernst J *et al*. Integrative analysis of 111 reference human epigenomes. *Nature* 2015;**518**:317–30.

145. Bujold D, Morais DAL, Gauthier C *et al*. The international human epigenome consortium data portal. *Cell Syst* 2016;**3**: 496–99.e2.

146. van Baak TE, Coarfa C, Dugué PA *et al*. Epigenetic supersimilarity of monozygotic twin pairs. *Genome Biol* 2018;**19**:2.

147. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biol* 2014;**15**:R31.

148. Houseman EA, Accomando WP, Koestler DC *et al*. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* 2012;**13**:86.

149. Birney E, Davey Smith G, Greally JM. Epigenome-wide association studies and the interpretation of disease -omics. *PLoS Genet* 2016;**12**:e1006105.

150. Bell JT, Tsai PC, Yang TP *et al*. Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. *PLoS Genet* 2012;**8**: e1002629.

151. Grundberg E, Meduri E, Sandling JK *et al*. Global analysis of DNA methylation variation in adipose tissue from twins reveals links to disease-associated variants in distal regulatory elements. *Am J Hum Genet* 2013;**93**:876–90.

152. Richmond RC, Simpkin AJ, Woodward G *et al*. Prenatal exposure to maternal smoking and offspring DNA methylation across the life course: findings from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Hum Mol Genet* 2015;**24**:2201–17.

153. Tsaprouni LG, Yang T-P, Bell J *et al*. Cigarette smoking reduces DNA methylation levels at multiple genomic loci but the effect is partially reversible upon cessation. *Epigenetics* 2014;**9**:1382–96.

154. Richmond RC, Sharp GC, Ward ME *et al*. DNA methylation and BMI: investigating identified methylation sites at *HIF3A* in a causal framework. *Diabetes* 2016;**65**:1231–44.

155. Leenen FAD, Muller CP, Turner JD. DNA methylation: conducting the orchestra from exposure to phenotype? *Clin Epigenetics* 2016;**8**:92.

156. Schultz MD, He Y, Whitaker JW *et al*. Human body epigenome maps reveal noncanonical DNA methylation variation. *Nature* 2015;**523**:212–16.

157. Cheung WA, Shao X, Morin A *et al*. Functional variation in allelic methylomes underscores a strong genetic contribution and reveals novel epigenetic alterations in the human epigenome. *Genome Biol* 2017;**18**:50.

158. Campión J, Milagro FI, Martínez JA. Individuality and epigenetics in obesity. *Obes Rev* 2009;**10**:383–92.

159. Preacher KJ. Advances in mediation analysis: a survey and synthesis of new developments. *Annu Rev Psychol* 2015;**66**:825–52.

160. Li R, Tsaih S-W, Shockley K *et al*. Structural model analysis of multiple quantitative traits. *PLoS Genet* 2006;**2**:e114.

161. Schadt EE, Lamb J, Yang X *et al*. An integrative genomics approach to infer causal associations between gene expression and disease. *Nat Genet* 2005;**37**:710–17.

162. Millstein J, Zhang B, Zhu J, Schadt EE. Disentangling molecular relationships with a causal inference test. *BMC Genet* 2009;**10**:23.

163. Relton CL, Davey Smith G. Two-step epigenetic Mendelian randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease. *Int J Epidemiol* 2012;**41**:161–76.

164. Yamada L, Chong S. Epigenetic studies in Developmental Origins of Health and Disease: pitfalls and key considerations for study design and interpretation. *J Dev Orig Health Dis* 2017;**8**:30–43.

165. Lin X, Barton S, Holbrook JD. How to make DNA methylome wide association studies more powerful. *Epigenomics* 2016;**8**:1117–29.

166. Munafò MR, Davey Smith G. Repeating experiments is not enough. *Nature* 2018;**553**:399–401.

167. Heijmans BT, Mill J. Commentary: the seven plagues of epigenetic epidemiology. *Int J Epidemiol* 2012;**41**:74–78.

168. Mill J, Heijmans BT. From promises to practical strategies in epigenetic epidemiology. *Nat Rev Genet* 2013;**14**:585–94.

169. Saffari A, Silver MJ, Zavattari P *et al*. Estimation of a significance threshold for epigenome-wide association studies. *Genet Epidemiol* 2018;**42**:20–33.

170. Kumaran K, Yajnik P, Lubree H *et al*. The Pune Rural Intervention in Young Adolescents (PRIYA) study: design and methods of a randomised controlled trial. *BMC Nutr* 2017;**3**:41.

171. Potdar RD, Sahariah SA, Gandhi M *et al*. Improving women's diet quality preconceptionally and during gestation: effects on birth weight and prevalence of low birth weight—a randomized controlled efficacy trial in India (Mumbai Maternal Nutrition Project). *Am J Clin Nutr* 2014;**100**:1257–68.

172. Owens S, Gulati R, Fulford AJ *et al*. Periconceptional multiple-micronutrient supplementation and placental function in rural Gambian women: a double-blind, randomized, placebo-controlled trial. *Am J Clin Nutr* 2015;**102**:1450–59.

173. Nguyen PH, Young M, Gonzalez-Casanova I *et al*. Impact of preconception micronutrient supplementation on anemia and iron status during pregnancy and postpartum: a randomized controlled trial in rural Vietnam.. *PLoS One* 2016;**11**:e0167416.

174. Hambidge KM, Krebs NF, Westcott JE *et al*. Preconception maternal nutrition: a multi-site randomized controlled trial. *BMC Pregnancy Childbirth* 2014;**14**:111.

175. Ladd-Acosta C, Fallin MD. The role of epigenetics in genetic and environmental epidemiology. *Epigenomics* 2016;**8**:271–83.

Chapter 4 Variation in plasma biomarker concentrations over the year in the Methyl Donors and Epigenetics-2 dataset

_____

Summary of Chapter

Background: In The Gambia previous studies have shown that plasma concentrations of one-carbon biomarkers vary over the year, and that maternal concentrations of some of these at the time of conception predict infant DNA methylation at various genomic loci.

Objective: The four objectives of this chapter were to describe the variation in concentrations of an expanded set of plasma biomarkers over the year, including amino acids; to assess the degree of replication of previous yearly trends; to check for new seasonal windows of nutritional vulnerability, and to provide a dataset ready for downstream nutrition-methylation analysis.

Methods: Banked plasma samples from 350 pregnant women in their first trimester were analysed for 26 metabolites. Samples were purposively selected to span the whole year. All plasma biomarkers were pre-adjusted for inflammation, age, body mass index and gestational age, and were assessed for evidence of periodic variation over the year. Those demonstrating a periodic pattern were fitted using Fourier term linear regression models to visualise the periodic variation over the year. The amplitude of the periodic variation was calculated using the coefficient of cyclic variation (CCV). The timing of peaks and troughs of the trends were compared with data from previous studies.

Results: All core one-carbon biomarkers apart from methionine and dimethylglycine (DMG) demonstrated a periodic pattern, along with six of the other amino acids. For the biomarkers showing evidence of periodic variation, the Fourier term regressions explained between 2.3% to 10.1% of the variability. The magnitude of the periodic variation (CCV) varied from 3.5% to 16.2%. Comparison of the peaks and troughs of biomarker trends with previous studies demonstrated good replication, with the exception of DMG.

Conclusions: Fitted trends using Fourier term regression models explained relatively little of

the overall variation of biomarker concentrations. However, small fluctuations in individual biomarkers may combine to exert larger effects on methylation potential and DNA methylation, to be considered in future analyses. There is no clear-cut season of deficiency and adequacy, although a possible new window of nutritional vulnerability to consider is September to October.

Notes

Four documents in the Annexes provide additional details for this chapter:

- An overview of how a systems biology approach can be used to consider the entirety of human metabolic pathways **(Annex 4.1).**
- A detailed primer on how to use Fourier term regression modelling to represent periodic trends, and how such models are used to back-extrapolate data to the time of conception **(Annex 4.2).**
- Details of all the plasma biomarkers assessed, with information on their role in one-carbon metabolism, possible changes in concentration during pregnancy and storage, and their correlation with dietary intake data **(Annex 4.3).**
- Annotated one-carbon metabolic pathway diagrams further justifying the rationale for biomarker choice and providing background on biochemistry **(Annex 4.4).**

4.1   Introduction

In this chapter I explore the extent to which a broad range of nutritional plasma biomarkers in pregnant women from rural Gambia demonstrate a periodic variation over the year and whether they co-vary in the expected ways. This builds on previous work performed by Dominguez-Salas et al. (2013)[1] using the 'Indicator group', described in Chapter 2. In brief, the Indicator group study team followed approximately 30 non-pregnant women in West Kiang region for one year. Each month they took a fasted venepuncture sample and analysed plasma concentrations of one-carbon biomarkers. Regression models using Fourier terms were used to represent the periodic variation that the biomarkers displayed over the year. The Indicator group data were also used to assess whether there were differences in plasma concentrations of one-carbon biomarkers between the peak rainy (July-September) and peak dry (February-April) seasons, under the hypothesis that marked variation in dietary intake between the two seasons would influence nutrient status. The Indicator group findings suggested that one-carbon biomarkers did show evidence of periodic patterns, and that folate, betaine, B2 and the S-adenosyl methionine (SAM): S-adenosyl homocysteine (SAH) ratio were higher in the rainy season, suggesting a higher methylation potential.

I investigate a subsample of 350 pregnant mothers enrolled into the ENID trial to form the 'MDEG-2' dataset, introduced in Chapter 2. The plasma samples form a baseline that spans the whole year, enabling an analysis of periodic patterns across the year. A key difference, however, is that the MDEG-2 samples are cross-sectional (one time-point, 350 women), whereas the indicator group is composed of repeated measurements.

I had four main aims forming the rationale of the analyses in this chapter:

1. **Expanding the set of plasma biomarkers to include amino acids**

   Whilst the one-carbon metabolic pathways investigated in the Indicator group are the ones most proximal to DNA methylation, my collaborators and I wanted to broaden the investigation to other related pathways that may have been previously overlooked. We therefore included a panel of amino acids since several amino acids are thought to contribute one-carbon units[2], and in turn deficiencies in one-carbon-related analytes may affect amino acid metabolism[3–5]. Originally as part of this aim my collaborators and I had considered integrating as much plasma concentration information as resources enabled into a systems biology model that sought to

represent the whole human metabolome. I provide details of the original systems biology plans in **Annex 4.1.** The intention of this approach was to use a platform that could enable us to interrogate other metabolic pathways outside the core one-carbon pathways. Although this collaboration did not work out I chose to continue the exploration of the additional metabolites in the analyses described in this chapter, in the expectation they could still potentially identify pathways distal to the 'core' one-carbon ones to consider.

2. **Replication of previous yearly trends**

Analysis of the MDEG-2 dataset provided an opportunity to assess the extent of replication of yearly trends seen in both the MDEG Indicator group and main study group. These studies have been introduced in Chapter 2 (summary in Table 2.1). As part of this aim I sought to determine the whether the MDEG-2 dataset, a cross-sectional sample of women taken from throughout the year, would enable any periodic patterns to emerge, or whether there would be too much inter-individual variation for this.

3. **Defining seasonal windows of nutritional vulnerability**

I wanted to investigate in greater detail whether the 'peak seasons' used in previous studies (the Indicator group and MDEG main study[1,6]) were the best time points to consider, or whether there were other windows where women were particularly vulnerable to nutritional deficiencies. This might inform a suitable timeframe for the supplementation trial detailed in Chapter 8, as well as delineate specific windows for investigation in future studies using this dataset.

4. **Preparation for downstream nutrition-methylation analysis**

Finally, I aimed to provide a detailed description of the nutritional status of this subsample in preparation for future planned studies that will explore nutritional predictors of infant DNA methylation throughout the year (and not just in peak seasons as in the MDEG main study[6]). I intend to using the Fourier term regression models outlined in this chapter to be able to back-extrapolate the biomarker concentrations to the time of conception. I provide an example of how this will be done in **Annex 4.2**. This will enable continued exploration of periconceptional nutritional exposures and their association with offspring methylation. This chapter

will help establish where the equations come from for the back-extrapolation, and will also provide transparency on the strengths and weaknesses of the modelling approach.

The role of nutritional components in one-carbon metabolic pathways has been reviewed in Chapter 3. However, **Annex 4.3** provides a summary of the core one-carbon biomarkers with a reminder of their relevance, as well as detailing factors that may affect their plasma concentration (such as stage of pregnancy and storage). The one-carbon metabolic pathways are further explained in **Annex 4.4,** which provides annotated metabolic pathway diagrams highlighting the role of the chosen metabolites. These illustrate in more detail how the transmethylation, folate, methionine and transsulfuration pathways intersect and justify the choice of the additional metabolites included in the expanded biomarker set.

4.2   Methods

4.2.1   Sample population

I used a subsample of banked plasma aliquots obtained from pregnant women in their first trimester enrolling onto the ENID Trial, described in Chapter 2. Women provided their blood samples on enrolment, referred to as their 'booking' visit. This was before they received any supplements from the ENID trial and formed the baseline measurements. I selected the subsample of 350 women aiming to obtain as equal a proportion of women per booking month as possible, and choosing the earliest gestational ages within each month of booking. Figure 4.1 shows the flow diagram for the purposive selection of the 350 women in the subsample from 688 eligible women enrolled into the ENID Trial. Gestational age was estimated by ultrasound by study midwives. The sample size was chosen on the basis that this provided approximately 30 women per month, which was the sample size of the Indicator group.

4.2.2   Laboratory analyses

The mothers' data is comprised of nutritional biomarkers analysed from EDTA-treated plasma. The following biomarkers were measured: choline, betaine, dimethylglycine (DMG), homocysteine, folate, vitamin B12, B6 vitamers (4-pyridoxic acid (PA), pyridoxal (PL) and pyridoxal-5'-phosphate (PLP)), vitamin B2, α-1-acid glycoprotein (AGP), cysteine, methionine, serine, glycine, alanine, arginine, aspartic acid, glutamic acid, histidine, isoleucine, leucine, lysine, phenylalanine, proline, threonine, tryptophan, tyrosine and valine.

```
┌─────────────────────┐
│   Pregnant women    │
│ enrolled into ENID  │
│       trial         │
│      N = 875        │
└─────────────────────┘
```

Withdrew during pregnancy   n = 42

Still birth   n = 21

Miscarriage   n = 12

Withdrew before child age 2   n = 111

Child died during study   n= 1

```
┌─────────────────────┐
│ Women eligible for  │
│  sub-sample         │
│    selection        │
│     N = 688         │
└─────────────────────┘
```

```
┌─────────────────────┐
│ Subsample selection │
│   for maternal      │
│   booking plasma    │
│    biochemistry     │
│     N = 350         │
│                     │
│  Based on equal     │
│  representation by  │
│  month of booking   │
│  and earliest       │
│  gestational age    │
└─────────────────────┘
```
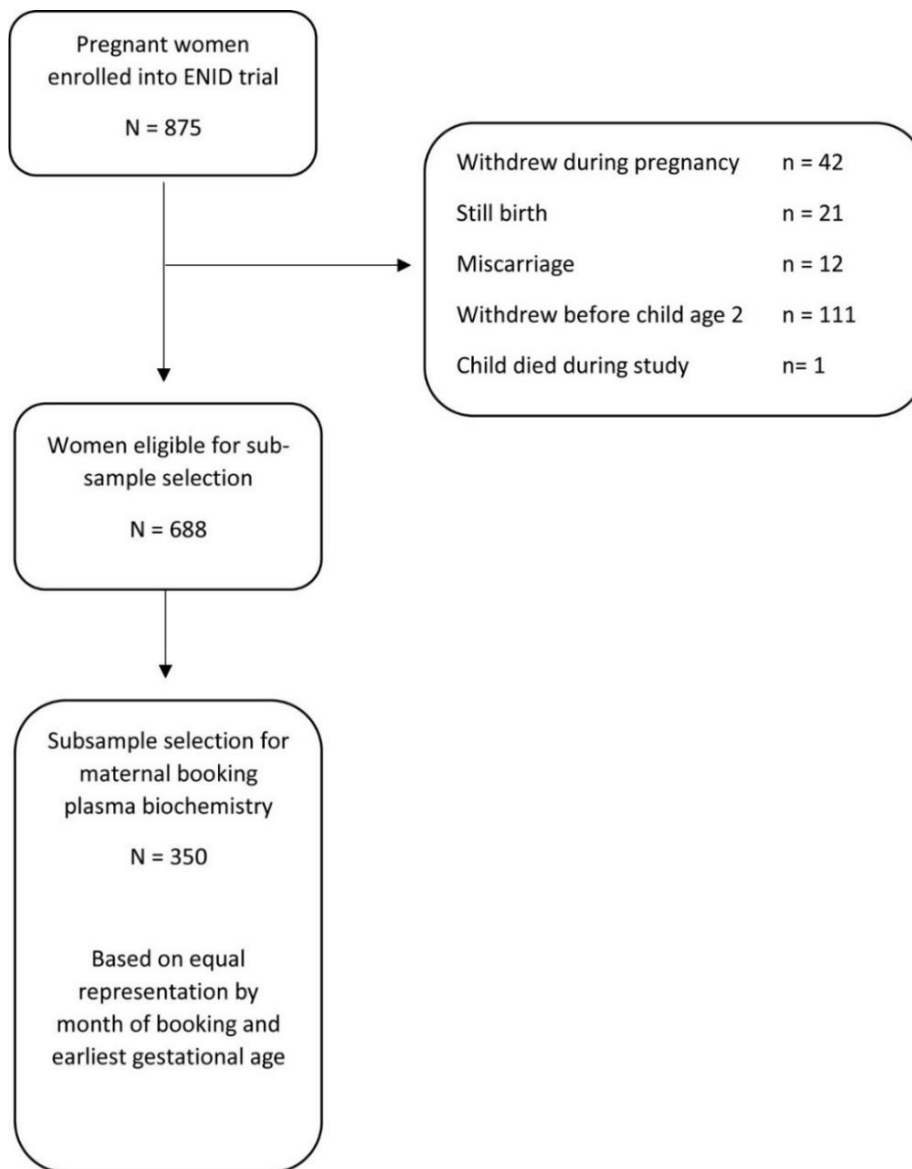
Figure 4.1 Flow diagram showing the number of women enrolled onto the ENID trial and subsequent subsampling for this MDEG-2 study.

Vitamin B12 and folate were measured using an Abbott AxSYM autoanalyzer. Choline, betaine, DMG, homocysteine, folate, B6 vitamers (PA, PL, PLP) and vitamin B2 were measured using liquid chromatography-tandem mass spectrometry. The amino acids were analysed using a Hitachi L-8900 amino acid analyser. All of these biomarkers were measured at the Child and Family Research Institute at the University of British Columbia. The inflammatory marker AGP was measured using the Cobas Integra 400 plus autoanalyser at MRC The Gambia, Keneba field station. All intra- and inter-assay coefficients of variation are

presented in **Supplementary Table 4.1**. Of the B6 vitamers I chose to solely present PLP, since this is the active component used as a coenzyme in reactions relevant to one-carbon metabolism.

The original plan was also to measure S-adenosyl methionine (SAM), S-adenosyl homocysteine (SAH), decarboxylated SAM (dc-SAM), formate and active B12 (holotranscobalamin). This was not possible, either due to degradation of the samples (SAM, SAH and dc-SAM), insufficient plasma volume (active B12) or assay failure (formate).

In the analyses the biomarkers that were measured both in the Indicator group and in this dataset are termed 'core one-carbon biomarkers' (choline, betaine, DMG, homocysteine, folate, vitamin B12, PLP, vitamin B2, cysteine and methionine). The new metabolites included in this dataset are referred to as the 'other amino acids' (serine, glycine, alanine, arginine, aspartic acid, glutamic acid, histidine, isoleucine, leucine, lysine, phenylalanine, proline, threonine, tryptophan, tyrosine and valine).

4.2.3   Data preparation

Outliers in the plasma biomarkers were defined as those having a z-score of <-6 or >6. The number of exclusions made per biomarker, along with the justification, is summarised in **Supplementary Table 4.1**, including the number of samples that did not have sufficient plasma to assay.

4.2.4   Statistical analyses

I used Stata 15.1 (StataCorp LLC, USA) for all analyses. I assessed the normality of the biomarker data by performing a skewness-kurtosis test based on D'Agostino *et al.* (1990)[7] (using the command 'sktest' in Stata). Biomarkers were skewed to the right and were therefore log-transformed to better approximate a normal distribution. For the baseline table I summarised overall (12-month) geometric means and 95% confidence intervals. For comparison with the Indicator group I also summarised the portions of the subsample falling into the peak rainy (July-October) and peak dry (February-April) seasons. I compared the rainy and dry season subsamples using the Wilcoxon rank-sum test.

I pre-adjusted the plasma biomarkers before use in regression models for four *a priori* confounders: inflammation, age, body mass index (BMI) and gestational age. This was done to remove known associations with biomarkers that might mask any underlying periodic variation. Inflammatory markers are often overlooked in one-carbon biomarker analyses, but

have been shown in previous studies to be inversely associated with concentrations of PLP[8]. Age has the potential to influence nutrient status through influence on dietary intake, pathological and physiological changes[9–11]. BMI has been associated with several nutrient concentrations, including one-carbon biomarkers, in previous studies[12,13]. Gestational age may have an influence on biomarker concentrations through physiological changes in pregnancy, for example, through haemodilution (see examples provided in **Annex 4.3**). As in Dominguez-Salas (2014)[6], I generated the first three orthogonal polynomials for gestational age. I then regressed the log-transformed biomarker against the gestational age polynomials, maternal age, maternal BMI and AGP (a measure of chronic inflammation). I added the regression residuals to the constant to obtain the adjusted variables to take forwards into Fourier term regressions.

Periodic patterns are variations around a mean that can be modelled using Fourier terms (FTs) in linear regression models. These models introduce sine and cosine waves to represent the periodic / cyclical fluctuation of biomarker concentrations over the year. In **Annex 4.2** I provide a primer of how to generate the FTs, how to assess whether it is appropriate to model a periodic pattern or not, and if so, how to choose the number of pairs of FTs to include in the linear regression model. For those biomarkers demonstrating a periodic pattern I measured the magnitude by calculating the 'coefficient of cyclic variation (CCV)'. This is determined by the square root of half the sum of the squared Fourier term coefficients[14]. The CCV is a standardised statistic that enables us to directly compare the extent of periodic variation between biomarkers and between datasets, regardless of the original units of the biomarker concentrations.

The biomarker data were visualised by plotting the exponentials of the individual pre-adjusted data points along with their fitted periodic variation curves (the mean and 95% confidence intervals (CI) of the fitted predictions from the appropriate FT regression). If there was no evidence for a trend I simply plotted the mean (95% CI) of the adjusted biomarkers. Given the biomarkers have a variety of measurement units I also represented the periodic variations as change in percentage of the geometric mean in order to plot groups of metabolites on the same axis according to their metabolic pathways.

I compared the trends with the Indicator group and the MDEG main study by checking the timing of the peaks and troughs of the trends over the year. The Indicator group trends are published in Figures 4E-G of Dominguez-Salas *et al*. (2013)[1]. The MDEG main study trends for the two three-month peak season windows (before they were back-extrapolated to the

time of conception) are shown in Figure 1b and Supplementary Figure 1 of Dominguez-Salas *et al.* (2014)[6]. I also compared the CCVs with the Indicator group (these were not available for the MDEG main study since the data only covered 'peak' seasons comprising six months of the year).

4.3   Results

4.3.1   Sample population

The proportion of women represented in the subsample per month of booking and their mean gestational age is shown in Figures 4.2 and 4.3 respectively. I also include a visual reminder of the definitions of the peak seasons, with peak dry season February to April and peak rainy season July to September. I was able to obtain data from each month of booking, although proportions were imbalanced, ranging from 6.0% in May to 10.9% in July. Mean (95% CI) gestational age per month of booking were within approximately two weeks of each other across the year, ranging from 11.2 (10.7, 11.7) weeks in July to 13.4 (12.0, 14.8) weeks in January.
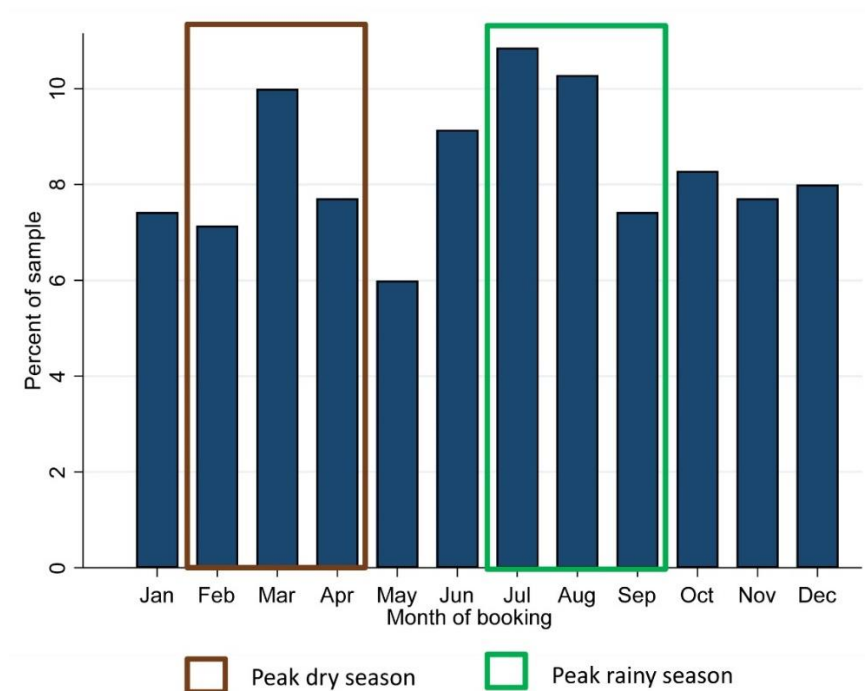
Figure 4.2 Proportion of women selected for the subsample per month of booking (n=350).
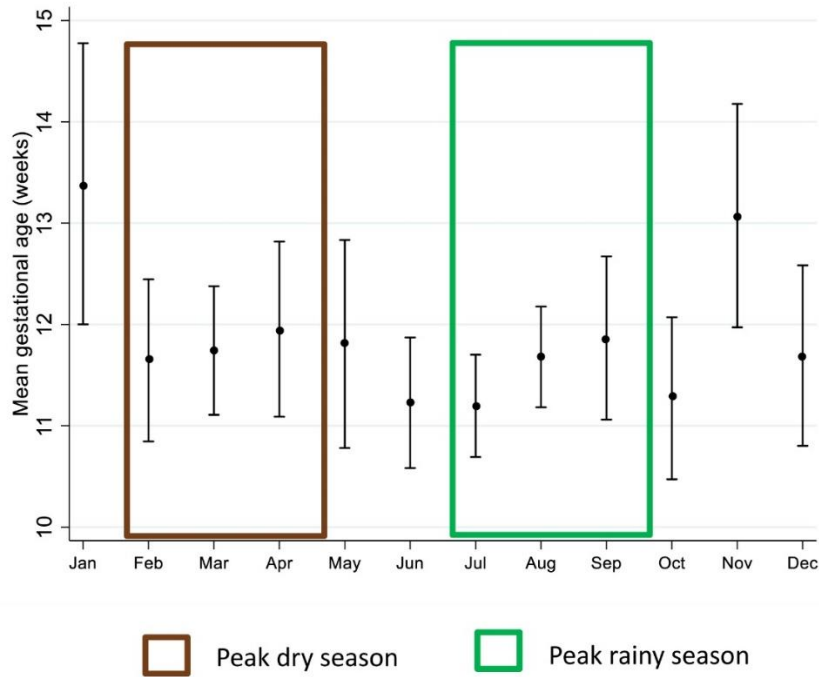
Figure 4.3 Mean (95% confidence interval) gestational age in weeks per month of booking (n=350).

4.3.2 Baseline characteristics

In Table 4.1 I summarise the overall geometric mean (95% CI) plasma biomarker concentrations, age, gestational age and BMI of the subsample. I also stratify these by peak season for comparison with Dominguez Sala et al. (2013)[1]. Looking at the core one-carbon biomarkers, over 30% of the subsample had low B2, PLP and methionine status and approximately quarter of the subsample had low folate and betaine status. The subsample were relatively replete in B12, choline and cysteine and only 4% of the subsample were categorised as having hyperhomocysteinemia. Regarding the other amino acids the majority (>80%) of the population had geometric mean concentrations below the 10th percentile of a healthy population in Canada for arginine, histidine, leucine, lysine, tyrosine and valine. This proportion was between 60-80% for isoleucine, phenylalanine, proline, serine and threonine, whilst it approximately half the sample for alanine and glycine. The majority of the subsample (87.3%) showed no sign of chronic inflammation as measured by AGP. The women had a geometric mean (95% CI) age of 28.8 (28.1, 29.5) years and a geometric mean BMI of 20.8 (20.5-21.1) kg/m^2.

For the core one-carbon biomarkers where differences were detected between the peak seasons, homocysteine, PLP, B12, cysteine and DMG had higher concentrations in the peak dry season. Folate had a higher concentration in the peak rainy season. There was no

difference between peak season concentration of B2 and betaine. Amongst the other amino acids glycine was higher in the peak rainy season and histidine, leucine and lysine were higher in the peak dry season. There was no evidence of a difference amongst the remaining amino acids. There was also no evidence to suggest a difference in inflammation between the peak seasons. Women enrolled into the trial during the peak rainy season were slightly older (2.5 years) than those enrolling in the peak dry season, but there was no evidence for a difference in gestational age nor BMI.

4.3.3   Periodic patterns in biomarkers over the year

Figures 4.4A-Y show the plasma biomarker values adjusted for age, gestational age, inflammation and BMI. Where there is evidence of a periodic pattern they are overlaid with the FT regression fitted curves, showing the mean and 95% CIs of the fitted values. Table 4.2 provides further details of these models, presenting the number of pairs of FTs introduced into the regressions where there is evidence of a trend, the proportion of the variance explained by the model (R^2 and adjusted R^2), and the size of the periodic variation as determined by the CCV. I also compare whether the peaks and troughs of the core one-carbon biomarkers coincide with those seen in previous studies. Figures 4.5A-D express the periodic co-variation of biomarkers in percentage of the geometric mean. Although the metabolic pathways overlap, I portray the folate, betaine and transsulfuration pathways separately and summarise the co-variation of the other amino acids displaying a trend.

There was evidence of periodic variation for all of the core one-carbon biomarkers apart from methionine and DMG. The FTs explained between 2.2% (for B12) and 12.9% (for folate) of the variation in biomarker concentrations (adjusted R^2). Six of the other 15 amino acids showed evidence of periodic patterns (glycine, histidine, leucine, lysine, proline and serine). The biomarker concentration variation explained by the models ranged from 2.3% (for serine) to 10.1% (for lysine). Comparing the magnitude of the periodic variation for biomarkers with a trend, measured by the CCV, shows that B2 had the greatest periodic variation (16.2%), followed by folate (15.1%). The smallest periodic variation was found amongst the amino acids histidine, leucine and serine, with CCVs of approximately 3.5%.

In the folate pathway (Fig. 4.5A) homocysteine showed inverse periodic variation with folate, and to a slightly lesser extent, with B2. Peak folate /B2 and trough homocysteine were in June-July. The peak of B12 did not coincide with the other B vitamins, coming earlier in the year during April. In the betaine pathway (Fig. 4.5B) betaine and choline had peaks at

different times of the year, and these did not mirror homocysteine troughs, nor did DMG show evidence of any trend. In the transsulfuration pathway (Fig. 4.5C) PLP and cysteine periodic variations largely coincided with peaks in April-May and troughs in September-October, although the amplitude of PLP was greater than that of cysteine. Homocysteine showed an inverse pattern to these, although with a lag of two months with its trough in July. Within the six other amino acids showing periodic variation (Fig. 4.5D), histidine and lysine had a similar pattern with peak concentration in May and lowest concentration in August-September. Proline and serine had similar trajectories to each other, although with a one-month lag and a more complex pattern due to each being fitted with three pairs of FTs. Leucine and glycine showed inverse associations, with peak leucine / trough glycine in March and trough leucine / peak glycine in September.

4.3.4   Comparison with Indicator group and MDEG main study

In the core one-carbon biomarkers there was a similar periodic pattern between the MDEG-2 results and the previous Indicator group and MDEG main study cohort for homocysteine, folate, PLP and B12. Similar peaks but different troughs were found for B2 and betaine, and similar troughs but different peaks for choline and cysteine. Methionine and DMG showed different yearly patterns to past studies. The results suggested that neither methionine nor DMG showed evidence of a periodic pattern. In the previous studies although methionine had been fitted with Fourier terms the amplitude was very low, generating near-flat trends fairly similar to the current results. DMG differed the most out of the biomarkers, in past studies showing a periodic variation peaking in March and with a trough in July. Compared to the MDEG-2 findings the CCVs of the Indicator group were similar for B12, betaine and cysteine; smaller for B2, and larger for homocysteine, folate, PLP, choline, methionine and DMG.

Table 4.1 Baseline concentrations of nutritional biomarkers and characteristics of the sample population

Variables	12-month		Cut-off for low / abnormal status; n (%)	Dry season		Rainy season		P value*
	N	Geometric mean (95% CI)		N	Geometric mean (95% CI)	N	Geometric mean (95% CI)	
Core one-carbon biomarkers								
Homocysteine (µmol/L)	349	7.8 (7.5-8.0)	>15µmol/L[15]; 15 (4.3%)[a]	87	8.1 (7.7-8.5)	100	7.3 (6.8-7.8)	**0.023**
Folate (nmol/L)	350	13.1 (12.5-13.7)	<10nmol/L[16]; 90 (25.7%)	87	11.5 (10.8-12.4)	100	13.8 (12.6-15.2)	**0.006**
B2 (nmol/L)	347	13.5 (12.4-14.5)	<10nmol/L[17]b; 136 (39.1%)	86	14.5 (12.3-17.1)	100	12.5 (10.8-14.4)	0.202
PLP (nmol/L)	349	24.6 (23.5-25.8)	<20nmol/L[18]c; 110 (31.5%)	87	28.4 (25.7-31.4)	100	23.1 (21.5-24.9)	**0.001**
B12 (pmol/L)	350	374.5 (358.4-391.5)	<221pmol/L[19]; 33 (9.4%)	87	402.5 (368.7-439.3)	100	341.2 (313.1- 372.0)	**0.005**
Choline (µmol/L)	348	7.3 (7.1-7.4)	<5µmol/L[15]c; 9 (2.6%)	87	7.4 (7.1-7.7)	100	7.1 (6.8-7.3)	0.076
Betaine (µmol/L)	350	20.3 (19.6-21.0)	<16µmol/L[15]c; 85 (24.3%)	87	21.1 (19.7-22.5)	100	21.6 (20.2 -23.0)	0.676
Methionine (µmol/L)	350	21.2 (20.8-21.6)	<20µmol/L[20]d; 128 (36.6%)	87	21.5 (20.7-22.3)	100	21.1 (20.4-21.8)	0.490
Cysteine (µmol/L)	350	212.2 (208.1-216.3)	<36µmol/L[20]d; 0 (0%)	87	224.5 (216.8-232.4)	100	197.5 (190.6-204.7)	**<0.001**
Dimethylglycine (µmol/L)	346	2.0 (1.9-2.2)	Not established[e]	85	2.3 (2.0-2.5)	99	1.9 (1.7-2.1)	**0.024**
Other amino acids								
Alanine (µmol/L)	348	236.7 (231.1 -242.6)	<240µmol/L[20]d; 172 (49.4%)	86	233.6 (223.1-244.7)	98	243.5 (232.0-255.6)	0.198
Arginine (µmol/L)	346	51.0 (49.6-52.4)	<68µmol/L[20]d; 303 (87.6%)	86	52.0 (49.3-54.8)	98	50.4 (47.6-53.3)	0.508
Aspartate (µmol/L)	345	5.2 (5.0-5.4)	<2µmol/L[20]d; 2 (0.6%)	86	5.3 (4.9-5.7)	98	5.0 (4.6-5.5)	0.706
Glutamate (µmol/L)	344	53.0 (50.7-55.4)	<11µmol/L[20]d; 0 (0%)	85	53.2 (48.9-58.0)	98	50.6 (46.9-54.6)	0.516
Glycine (µmol/L)	346	177.1 (172.1-182.2)	<183µmol/L[20]d; 189 (54.6%)	86	159.6 (151.9-167.8)	98	191.4 (181.2-202.1)	**<0.001**
Histidine (µmol/L)	346	65.5 (64.6-66.4)	<77µmol/L[20]d; 308 (89.0%)	86	67.8 (66.0-69.6)	98	62.5 (60.9-64.2)	**<0.001**
Isoleucine (µmol/L)	346	43.8 (43.0-44.6)	<47µmol/L[20]d; 222 (64.1%)	86	44.7 (42.7-46.7)	98	43.1 (41.7-44.6)	0.443
Leucine (µmol/L)	346	82.9 (81.5-84.4)	<101µmol/L[20]d; 302 (86.8%)	86	85.2; (81.9-88.6)	98	79.0 (76.6-81.5)	**0.008**
Lysine (µmol/L)	346	108.3 (106.2-110.4)	<157µmol/L[20]d; 342 (98.3%)	86	111.9 (107.7-116.2)	98	101.5 (97.7-105.5)	**<0.001**

Variables	12-month		Cut-off for low / abnormal status; n (%)	Dry season		Rainy season		P value*
	N	Geometric mean (95% CI)		N	Geometric mean (95% CI)	N	Geometric mean (95% CI)	
Phenylalanine (µmol/L)	346	45.1 (44.5-45.7)	<47µmol/L[20]d; 231 (66.8%)	86	45.7 (44.2-47.1)	98	45.8 (44.7-47.0)	0.601
Proline (µmol/L)	345	103.4 (100.7-106.1)	<113µmol/L[20]d; 231 (67.0%)	86	101.8 (97.2-106.6)	98	109.4 (103.6-115.4)	0.062
Serine (µmol/L)	346	92.5 (90.7-94.3)	<101µmol/L[20]d; 245 (70.8%)	86	94.5 (91.1-98.1)	98	94.1 (90.9-97.3)	0.978
Threonine (µmol/L)	346	98.3 (96.1-100.7)	<104µmol/L[20]d; 209 (60.4%)	86	97.1 (92.2-102.3)	98	101.7 (97.2-106.4)	0.334
Tyrosine (µmol/L)	346	38.4 (37.6-39.1)	<46µmol/L[20]d; 288 (83.2%)	86	38.4 (36.9-40.0)	98	39.1 (37.6-40.6)	0.564
Valine (µmol/L)	346	150.9 (148.6-153.2)	<178µmol/L[20]d; 300 (86.7%)	86	151.8 (146.7-157.1)	98	147.1 (142.9-151.5)	0.192
Inflammatory markers								
AGP (g/L)	346	0.74 (0.70-0.78)	>1g/L[21]f; 44 (12.7%)	83	0.71 (0.64-0.80)	100	0.72 (0.68-0.76)	0.872
Other variables								
Age (years)	350	28.8 (28.1-29.5)	-	87	27.1 (25.8-28.5)	100	29.7 (28.4-31.0)	**0.016**
BMI (kg/m^2)	350	20.8 (20.5-21.1)	-	87	20.3 (19.8-20.8)	100	20.7 (20.1-21.4)	0.431
Gestational age (weeks)	350	11.6 (11.4-11.9)	-	87	11.6 (11.2-12.0)	100	11.4 (11.0- 11.8)	0.320

Abbreviations: CI, confidence interval; PLP, pyridoxal 5'-phophsate; AGP, Alpha-1-acid glycoprotein; BMI, body mass index.

*P value from Wilcoxon rank-sum test comparing the peak dry (February-April) and rainy (July-October) season samples, for comparison with Dominguez-Salas et al. (2013)[1]. Results <0.05 in bold.

aDenotes hyperhomocysteinemia.

bThere are no clearly defined plasma cut-offs for riboflavin deficiency (normally erythrocyte riboflavin is used). This cut-off represents the 5th percentile of healthy controls in the European Prospective Investigation into Cancer and Nutrition study[17].

cThere are no clearly defined plasma cut-offs for deficiency. The suggested cut-offs indicate below the normal range and can be considered 'low status'.

dThe amino acids cut-offs represent the 10th percentile of a healthy population age > 16years in Canada[20]. Note these cut-offs do not necessarily represent low status nor deficiency.

eBy way of comparison, in US, Nordic, Asian and Australian populations plasma dimethylglycine geometric mean ranges from 3.7-5.0µmol/L[22].

fIndicates chronic inflammation.

Figure 4.4A-Y: Plasma metabolite concentrations over the year, adjusted for age, gestational age, inflammation and BMI, showing individual data points and mean (95% CI) concentrations. Those with evidence of periodic variation show the fit obtained from the Fourier term linear regression models.

G. Betaine

H. Methionine

I. Cysteine

J. Dimethylglycine

K. Alanine

L. Arginine

M. Aspartate

N. Glutamate

O. Glycine

P. Histidine

Q. Isoleucine

R. Leucine

S. Lysine

T. Phenylalanine

U. Proline

V. Serine

W. Threonine

X. Tyrosine

Y. Valine

Adjusted plasma concentration (μmol/L)

Date of blood measurement
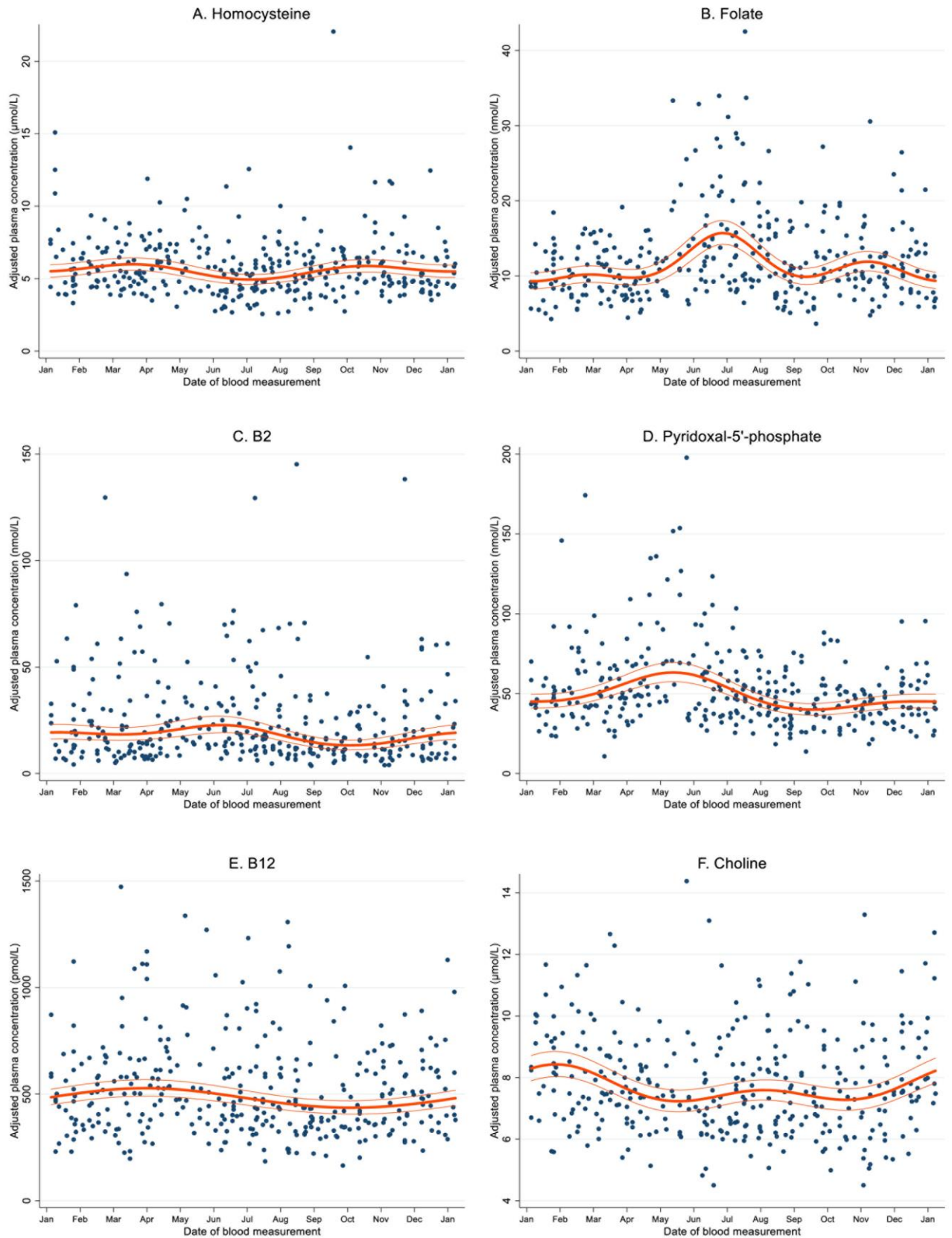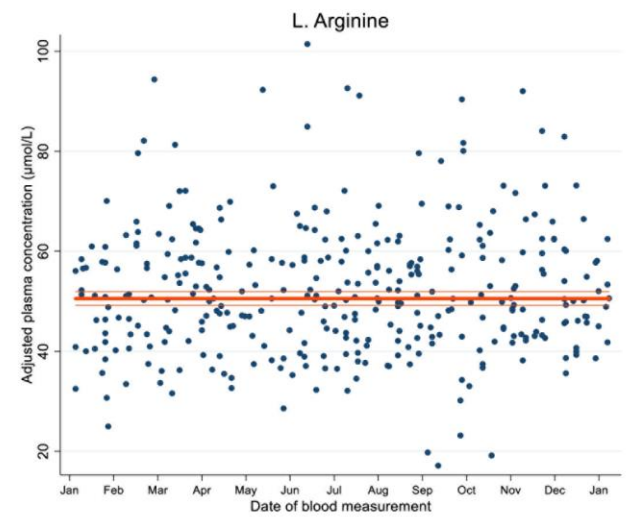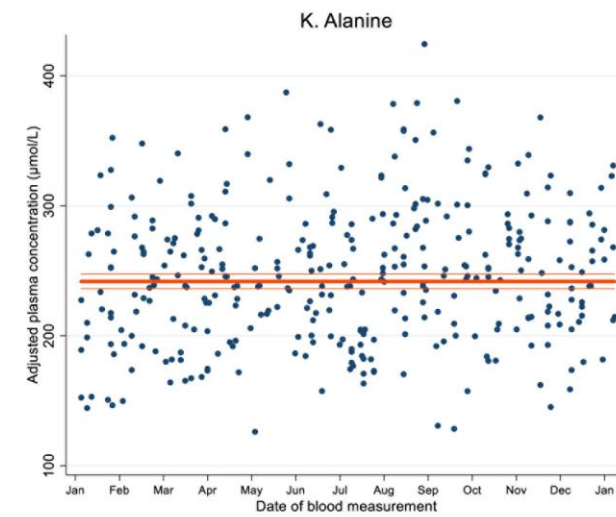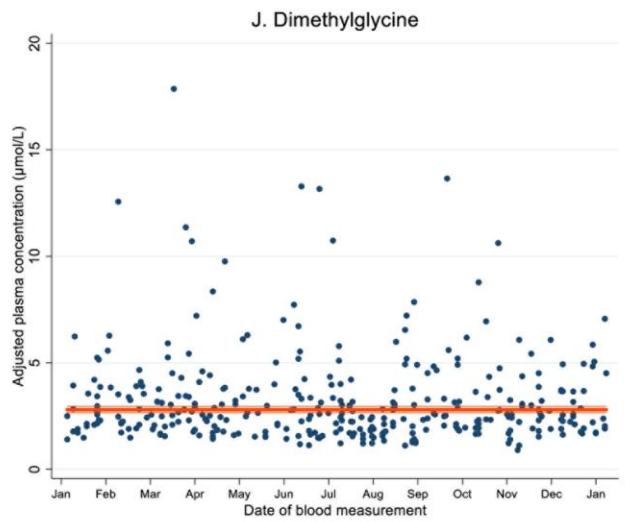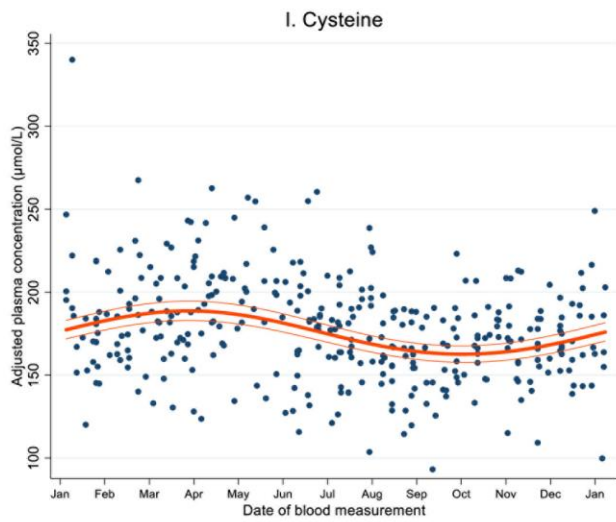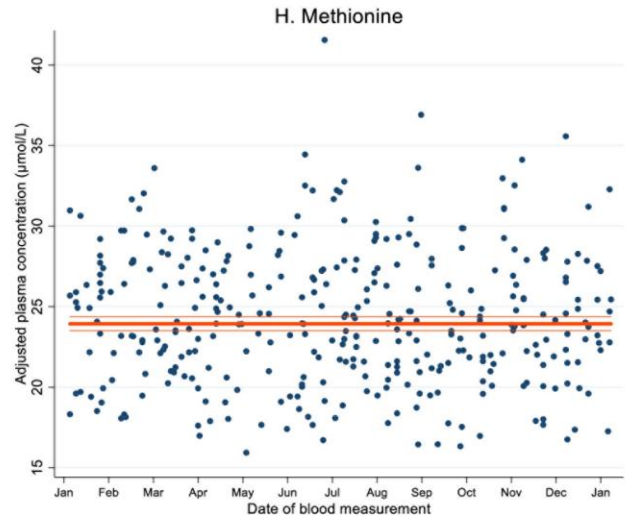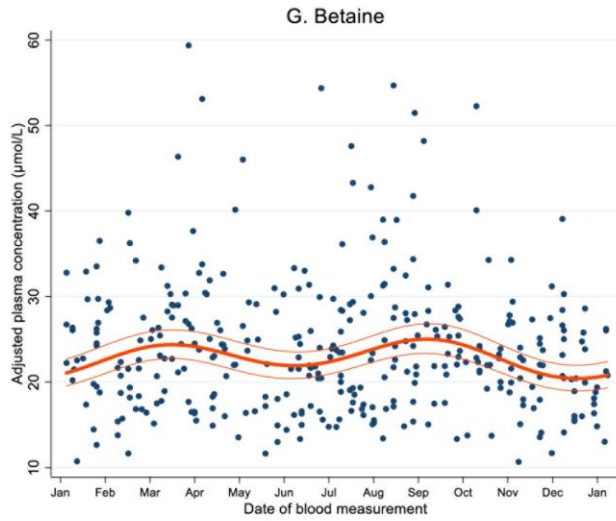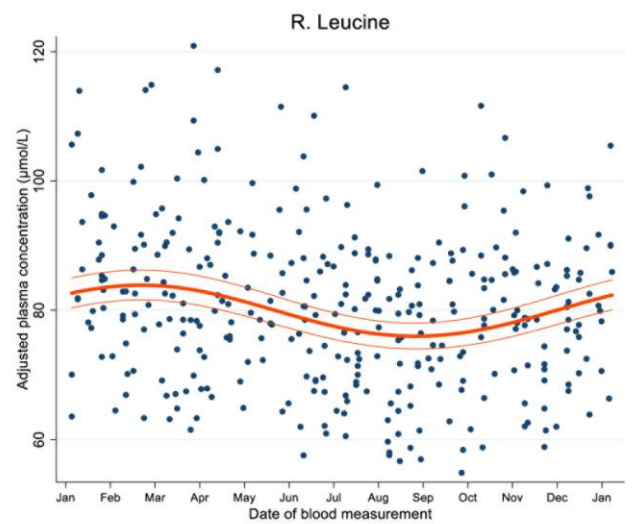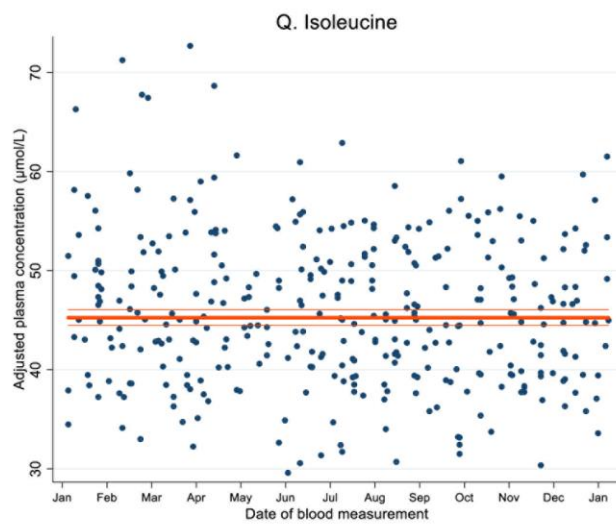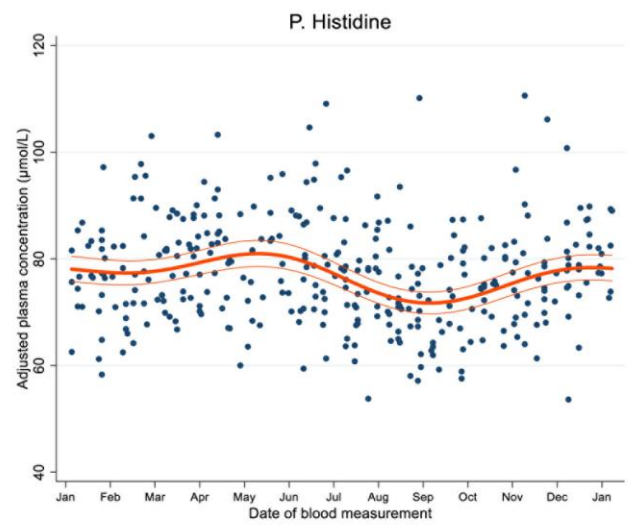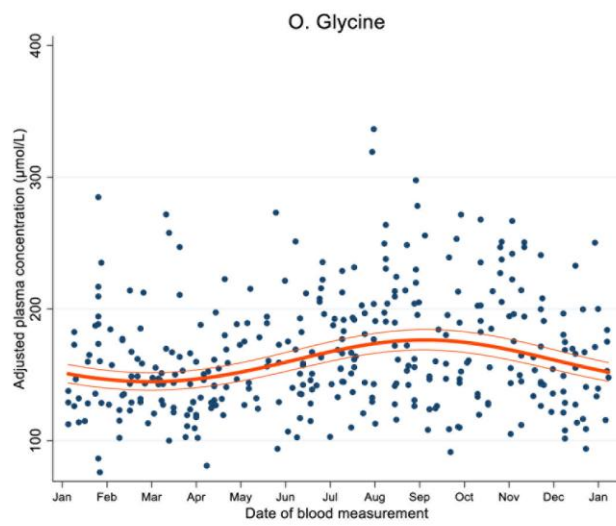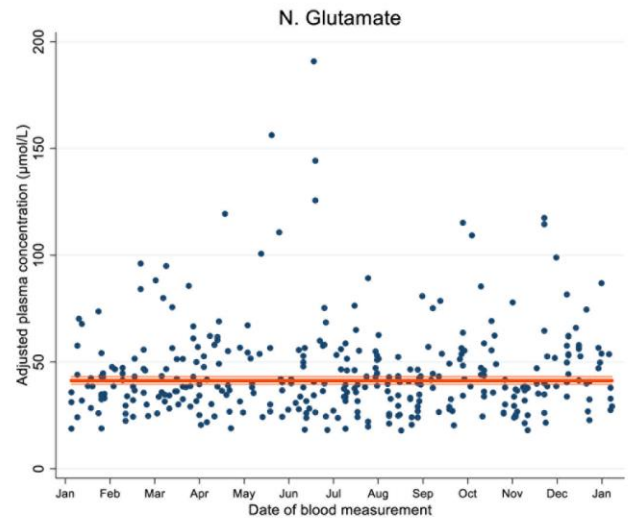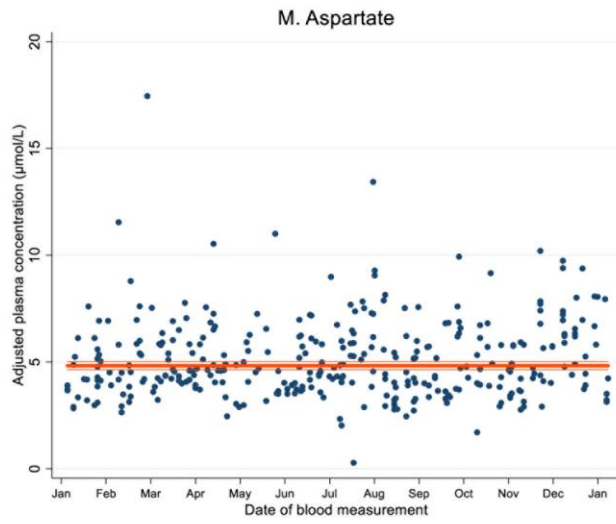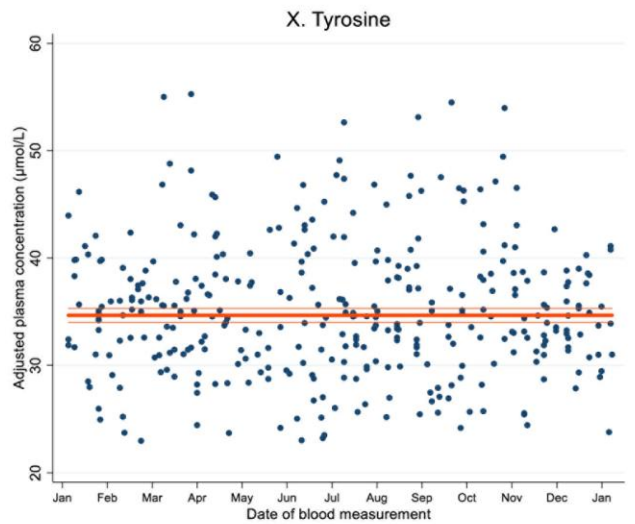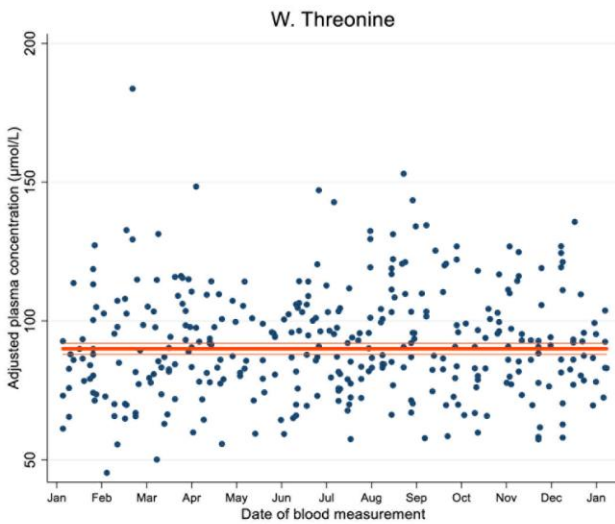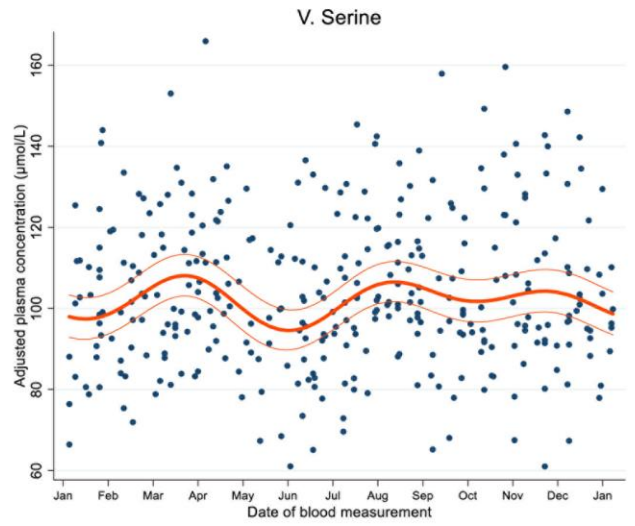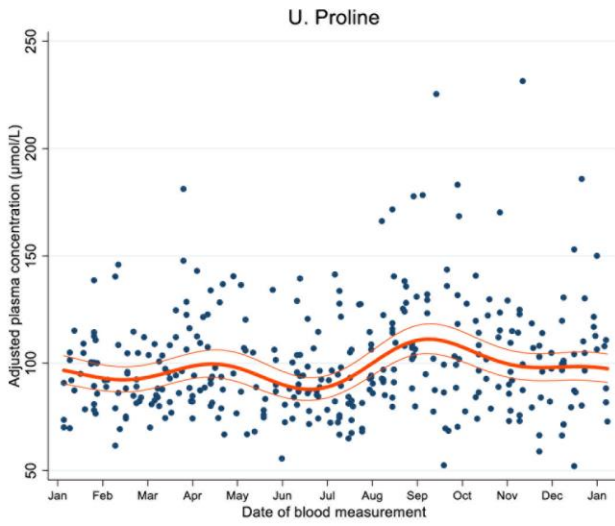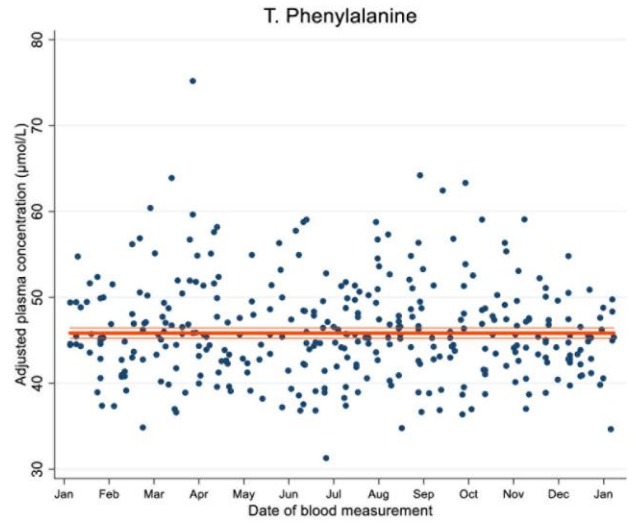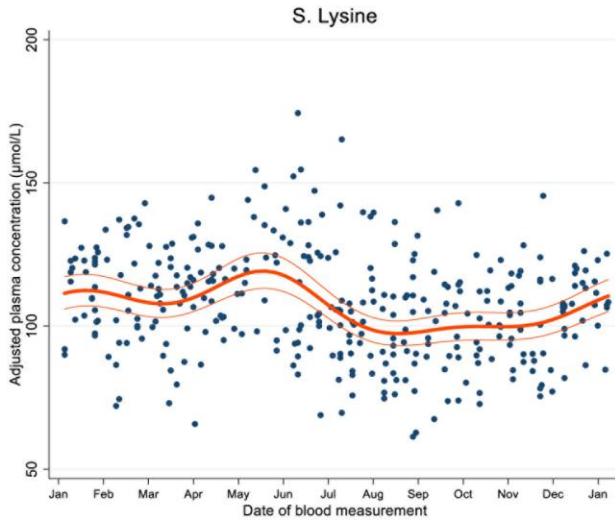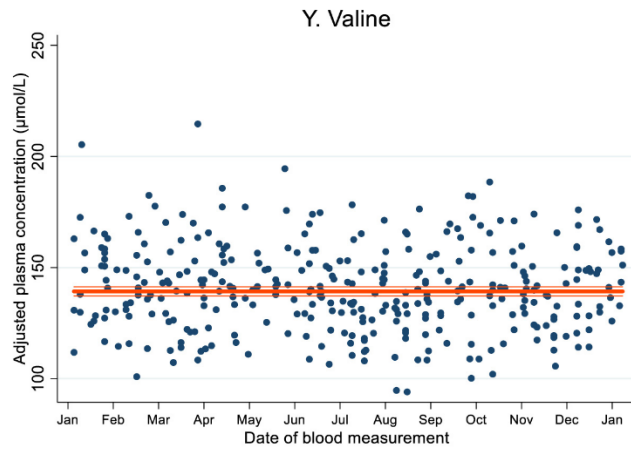
Table 4.2 Information on periodic variation fit for biomarkers in MDEG-2 and comparison with Indicator group and MDEG main study cohort.

Biomarker	Evidence for periodic variation?[a]	No. of pairs of Fourier terms	N	R^2	Adj. R^2	CCV (%) (95% CI)[b]	Timing of peaks and troughs	Similar peak and trough months in Indicator group?[c]	Similar peak and trough months in MDEG1 main study?[d]
Core one-carbon biomarkers									
Homocysteine	Yes	2	349	0.0390	0.0281	5.8 (2.8, 8.9)	Peaks: Apr & Oct Troughs: Jan & July	Yes – same timing Larger CCV: 10.2 (7.9-12.5)%	Similar Peak: Apr &Sept Trough: June
Folate	Yes	3	350	0.1443	0.1293	15.1 (11.2, 19.0)	Peak: June / July & Nov Trough: Sep & Jan-Apr	Yes – same timing Larger CCV: 23.8 (20.5-26.5)%	Yes – same timing
B2[e]	Yes	2	347	0.0481	0.0369	16.2 (8.5, 23.8)	Peak: June Trough: Oct	Similar peak but trough in Feb. Lower CCV: 9.0 (7.6-10.4)%	Similar peak. Trough not covered in window.
PLP[f]	Yes	2	349	0.1190	0.1087	14.6 (10.4, 18.8)	Peak: May Trough: Sep	Yes – same timing Larger CCV: 26.2 (22.4-30.0)%	Yes – same timing
B12	Yes	1	350	0.0278	0.0222	6.7 (2.5, 10.9)	Peak: Mar/Apr Trough: Sep/Oct	Yes – same timing Similar CCV: 3.5 (1.9-5.1)%	Similar peak, indiscernible trough
Choline	Yes	2	348	0.0533	0.0423	4.9 (2.7, 7.1)	Peaks: Feb & Aug Troughs: May & Oct	Similar trough in May but peak is Dec. Higher CCV: 11.1 (8.9-13.4)%	Similar troughs Peak not covered in window
Betaine	Yes	2	350	0.0385	0.0274	6.0 (2.8, 9.2)	Peaks: Mar & Sep Troughs: June & Dec	Similar peak in Sep but not in Mar. One trough in Feb. Similar CCV: 9.8 (7.6-12.0)%	Yes – same timing

Biomarker	Evidence for periodic variation?[a]	No. of pairs of Fourier terms	N	R^2	Adj. R^2	CCV (%) (95% CI)[b]	Timing of peaks and troughs	Similar peak and trough months in Indicator group?[c]	Similar peak and trough months in MDEG1 main study?[d]
Methionine	No	-	350	-	-	-	N/A	Similar very flat trend, possible peak Aug & Mar, possible trough Dec. CCV of 3.9 (2.0-5.8)%	Similar flat trend
Cysteine	Yes	1	350	0.0895	0.0842	5.3 (3.5, 7.0)	Peak: Mar/Apr Trough: Oct	Flat trend. Similar trough but possible peak Jun/Jul. Similar CCV: 4.1 (3.0-5.2)%	Flat trend, no discernible peaks and troughs
DMG	No	-	346	-	-	-	N/A	No – peak in Mar and trough in Jul. CCV 16.3 (13.4-19.3)%.	Trough also in Jul, peak not covered in window
Other amino acids									
Alanine	No	-	346	-	-	-	N/A	N/A	N/A
Arginine	No	-	346	-	-	-	N/A	N/A	N/A
Aspartate	No	-	345	-	-	-	N/A	N/A	N/A
Glutamate	No	-	344	-	-	-	N/A	N/A	N/A
Glycine	Yes	1	346	0.0767	0.0713	7.0 (4.4, 9.6)	Peak: Sep Trough: Mar	N/A	N/A
Histidine	Yes	2	346	0.0818	0.0703	3.6 (2.3, 5.0)	Peak: May Trough: Sep	N/A	N/A
Isoleucine	No	-	346	-	-	-	N/A	N/A	N/A
Leucine	Yes	1	346	0.0547	0.0492	3.5 (2.0, 5.0)	Peak: Feb/Mar Trough: Aug/Sep	N/A	N/A
Lysine	Yes	3	346	0.1163	0.1007	6.2 (4.4, 8.1)	Peak: May/Jun Trough: Aug-Nov	N/A	N/A
Phenylalanine	No	-	346	-	-	-	N/A	N/A	N/A

Biomarker	Evidence for periodic variation?[a]	No. of pairs of Fourier terms	N	R^2	Adj. R^2	CCV (%) (95% CI)[b]	Timing of peaks and troughs	Similar peak and trough months in Indicator group?[c]	Similar peak and trough months in MDEG1 main study?[d]
Proline	Yes	3	346	0.0729	0.0564	6.1 (7.3, 9.5)	Peaks: Apr & Sep Trough: Feb & Jun	N/A	N/A
Serine	Yes	3	346	0.0401	0.0232	3.6 (1.7, 5.4)	Peaks: Mar & Aug Troughs: Jan & Jun	N/A	N/A
Threonine	No	-	346	-	-	-	N/A	N/A	N/A
Tyrosine	No	-	346	-	-		N/A	N/A	N/A
Valine	No	-	346	-	-		N/A	N/A	N/A

Abbreviations: Adj., adjusted; CCV, coefficient of cyclic variation; CI, confidence interval; DMG, dimethylglycine; PLP, pyridoxal-5'-phosphate.

[a]Likelihood ratio test comparing model with Fourier terms against null model. P<0.05 interpreted as evidence for periodic variation, with successive pairs of Fourier terms being added based on further likelihood ratio tests. See Annex 4.2 for further details.

[b]Calculated as the square root of half the sum of the squared Fourier term coefficients[14].

[c]Visual comparison of timing of peaks and troughs of periodic patterns from 12-month trend Dominguez-Salas et al. (2013)[1].

[d]Visual comparison of timing of peaks and troughs of periodic patterns from 6-month window in Dominguez-Salas et al. (2014)[6].

[e]Note in MDEG-2 this is plasma B2 whilst in the Indicator group and MDEG main study this is 1/EGRAC coefficient from a red blood cell assay.

[f]MDEG-2 plots PLP, the active biomarker. The Indicator group and MDEG main study plot total B6. However, in additional material Dominguez-Salas (2012) shows how the B6 vitamers are highly correlated in their seasonal trends[23].

Figure 4.5A-D: Summary diagrams showing covariation of metabolites by metabolic pathway. All metabolites adjusted for age, gestational age, BMI and inflammation. Periodic patterns are fitted from Fourier term linear regression models (mean and 95% CI bars).
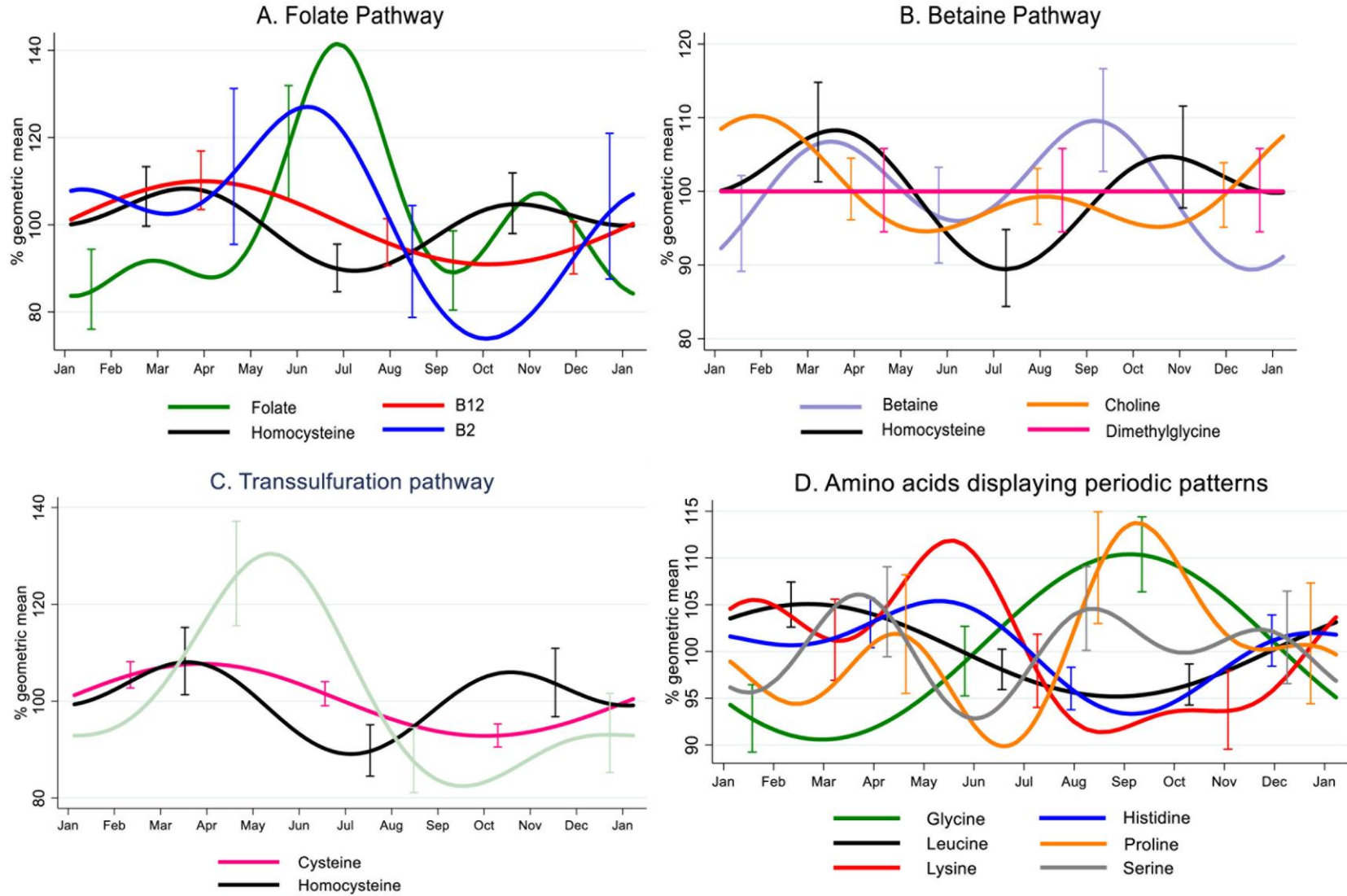
4.4 Discussion

I have explored the extent to which plasma biomarker concentrations of one-carbon metabolism and amino acids display periodic patterns over the year in a cross-sectional population of 350 women enrolled onto the ENID trial. All core one-carbon biomarkers apart from methionine and DMG demonstrated periodic variation, along with six of the other amino acids. The Fourier term regression models fitted to these biomarkers helped capture their covariation in different metabolic pathways, and helped quantify the extent to which periodic variation explained overall variation in the biomarkers. Timing of periodic peaks and troughs, as well as the amplitude of the periodic variations, varied greatly between the biomarkers.

4.4.1 Covariation in biomarkers by metabolic pathways

The covariation of biomarkers in the folate pathway largely followed what was expected from folate metabolism. Dietary folates and folic acid are reduced to form tetrahydrofolate (THF), which in turn is reduced to methylene-THF and then to 5-methyl-THF. B2 is required in this final reduction step since it is a precursor to flavin adenine dinucleotide, which is a cofactor for the enzyme 5-methyl tetrahydrofolate reductase to reduce methylene-THF to methyl-THF[24]. Methyl-THF donates its methyl group to homocysteine under the action of the enzyme methionine synthase, which requires cobalamin (vitamin B12) as a co-factor[25]. This final step converts homocysteine into methionine. These reactions explain why as folate and B2 increase so homocysteine decreases. We would also expect homocysteine to decrease as B12 increases, however, B12 has a modest periodic variation (CCV of 6.7%) compared to the larger periodic variations of folate and B2 (CCVs of 16.2% and 15.1% respectively). Furthermore, the women are more replete in B12 compared to folate or B2, suggesting that B12 status is not low enough to majorly disrupt homocysteine metabolism.

Covariation of the biomarkers in the betaine pathway do not follow exactly what is expected. Betaine is formed through the oxidation of choline[26,27], hence I expected to see very similar trends in variation of choline and betaine. The lack of synchronicity here may be because rates of conversion to betaine are dependent on the relative metabolic demand in comparison with other products that choline can form, such as acetylcholine, phosphocholine and phosphatidylcholine. Furthermore choline and betaine are found in different food products (e.g. wheat bran, wheat germ, spinach, beets for betaine and red meat, poultry, milk, eggs and fish for choline[28]), which may follow different seasonal dietary

intake patterns[1]. I expected homocysteine to decrease as choline and betaine increase because the methyl group from betaine can be donated to homocysteine via betaine-homocysteine methyl transferase[29], forming an alternative pathway for the remethylation of homocysteine independent from the B12-dependent pathway described above. Given DMG is a product of the reaction just described I would have expected DMG to increase as homocysteine decreases if the betaine pathway is being relied on to re-methylate homocysteine. The absence of such a relationship could be because the folate/B12 pathway is taking prominence, or potentially because plasma does not adequately reflect the betaine remethylation pathway, which occurs more in the kidney and liver.

In the transsulfuration pathway homocysteine is irreversibly degraded to cystathionine and then to cysteine[30]. Both of the enzymes involved in these steps require PLP as a cofactor. Whilst homocysteine does demonstrate an inverse pattern to PLP, they are not synchronised exactly and the peak of PLP comes two months before the trough of homocysteine. Similarly, although cysteine shows a similar pattern to homocysteine (as would be expected since it is a product of homocysteine metabolism), the corresponding peaks and troughs of the periodic patterns are about two months apart. Knowledge of the reactions involved in the transsulfuration pathway therefore help provide hypotheses as to the nature of covariation to expect, but do not fully explain what is actually found.

Covariation of the amino acids is not easy to interpret. Natural groupings to explore are a) serine, glycine and cysteine, b) histidine and tryptophan, c) tyrosine and phenylalanine and d) glutamate, proline and arginine.

Serine can be used to form both glycine and cysteine. Regarding glycine formation, serine donates a 1-carbon unit at the stage of converting THF to methylene-THF through the action of serine hydroxymethyltransferase and PLP[31]. In the process serine is converted to glycine, a reaction that is reversible. Regarding cysteine, serine donates its carbon skeleton in the transsulfuration pathway, where it is condensed with homocysteine to form cystathionine, which is then hydrolysed to form cysteine. However, there was not a pattern of covariation between serine and glycine, nor serine and cysteine. This may be because all three are classified as non-essential amino acids and therefore can be produced endogenously as per requirements. I also expected glycine to show correlations with folate, since it can be formed from carbon dioxide, ammonium and methylene-THF. In the reverse reaction in the glycine cleavage system in the mitochondria, the catabolism of glycine generates methylene-THF (with the involvement of THF), which is then used as a carbon donor in the one-carbon

pathways[32]. Identifying these interrelationships would have required a more specific folate assay that provides a breakdown of the different forms of THF. The assay used primarily measured the amount of 5-methyl-THF, given this is the dominant form of THF circulating in plasma.

The catabolism of amino acids tryptophan and histidine produces formate[2,33]. Formate is produced in the mitochondria and then released into the cytosol, forming formyl-THF by condensation with THF. Formyl-THF can then either be used in purine synthesis or be interconverted into other THF oxidation states as part of folate metabolism[34]. Unfortunately the assay used for formate measurement failed, but otherwise may have provided the missing link to better interpret covariation between these metabolites.

The other amino acids, whilst important for health in numerous ways[35,36], are less clearly related to one-carbon metabolism. I expected tyrosine and phenylalanine to co-vary, because tyrosine is formed from phenylalanine (an essential amino acid) under the action of phenylalanine hydroxylase. There was no evidence for a periodic pattern in phenylalanine, however. Similarly, I expected glutamate, proline and arginine to co-vary since glutamate is used as the precursor for the formation of proline and arginine[37]. However, these three amino acids are non-essential and therefore may better reflect metabolic demands rather than true periodic variations. Finally, leucine and lysine, both essential amino acids, had their lowest concentrations in August-September. Whilst both are needed for numerous functions (e.g. regulation of protein turnover, antiviral activity, protein methylation[35]), potential links with one-carbon metabolism are not well documented.

4.4.2   Comparison with previous studies

Comparing the timings of the troughs and peaks of the core one-carbon metabolites between this study and the Indicator group and the MDEG main group, there was generally a good replication of trends, with the exception of DMG. Only four biomarkers had an identical matching trend, however, and there are many possible reasons for this. The Indicator group was composed of non-pregnant women whereas the MDEG-2 and MDEG main study cohorts were comprised of pregnant women, mostly in their first trimester. Furthermore, I presented MDEG-2 periodic patterns adjusted for BMI, age, inflammation and gestational age for reasons outlined above (section 4.2.3), but the MDEG main study trends used unadjusted values. In the MDEG-2 data homocysteine, cysteine and B12 were all positively associated with inflammation (data not shown). Gestational age was positively

associated with homocysteine, cysteine and choline, and inversely associated with methionine, betaine, folate, PLP and B2. BMI was inversely associated with B12 and PLP and age was inversely associated with methionine and PLP. It is possible these confounders may partially explain some of the discrepancies found between the datasets.

The MDEG main study and MDEG-2 datasets are comprised of cross-sectional samples, and therefore will contain greater inter-individual variation than the Indicator group, which aimed to follow the same women all year (although replacements were made as some Indicator group participants withdrew due to becoming pregnant, reaching menopause, self-withdrawal, or moving away from the study area). The reduced inter-individual variation in the Indicator group may help explain why most of the CCVs were larger than in the MDEG-2 data.

Broadly, however, the comparison was useful in suggesting that despite the limitations of the subsample being composed of cross-sectional data, it was still able to capture similar periodic patterns found in the Indicator group, composed of longitudinal data. This suggests that intra-individual variation over the year is greater than inter-individual variation, and gives additional confidence that the existing MDEG-2 data can be used to assess potential timings for a supplementation trial (see below). It also suggests that once matching infant DNA methylation data becomes available for the subsample then this dataset will be helpful for looking at the nutritional predictors of methylation. **Annex 4.2** provides an example of how the Fourier term regression models from this chapter will be used to back-extrapolate the biomarker concentrations to the time of conception, for example, so that the research group can further interrogate periconceptional predictors of methylation to build on previous work in Dominguez-Salas *et al.* (2014)[6].

4.4.3   Fit of the periodic patterns from Fourier term regression models

Obtaining evidence for a periodic pattern does not guarantee that the trend provides a strong overall fit for the data, nor that there are clear seasonal differences. Obtaining a p<0.05 in the likelihood ratio test comparing the null model with the model with one pair of Fourier terms simply suggests that the Fourier term model fits better than having no terms. Whilst some biomarkers clearly show non-overlapping confidence intervals between peaks and troughs over the year (e.g. folate, PLP, cysteine), many demonstrate more subtle variation. Furthermore, for some biomarkers the extent of the periodic variation does not become apparent until all biomarkers are compared on the same scale (such as the

percentage of geometric mean graphs or the analysis of the CCVs). For example, looking at the B2 graph on its original scale of nmol/L (Fig. 4.4C) the periodic pattern appears fairly flat and the adjusted R^2 is only 3.7%. This is because, as with many of the biomarkers, the original distribution is right-tailed, meaning that the y-axis scale needs to be expanded to capture these high concentrations. However, looking instead at variation around the geometric mean (Fig. 4.6A), alongside calculating a CCV of 16.2%, shows the underlying periodic variation is much stronger than first thought. Fulford (2014)[14] explains that relying on R^2 estimates to explain the fit of periodic models can be problematic when imprecision in the biomarker estimates accounts for much of the variability (see laboratory CVs in **Supplementary Table 4.1**). That said, given the measurement imprecision and inter-individual variation it is encouraging that underlying periodic patterns can still be detected at all, including some biomarkers with a relatively high adjusted R^2 value of over 10% (folate, PLP, lysine).

The CCV places observation of periodic variation on a uniform scale that enables comparison cross-biomarker and cross-dataset. Observing the CCVs then, folate, B2, PLP in particular would seem to have a stronger periodic variation than consideration of the adjusted R^2 alone would indicate. Many of the biomarkers had both a low R^2 and a low CCV (e.g. homocysteine, B12, choline, histidine, leucine, proline, serine). In these cases it may be important to consider that small variations in individual biomarkers may combine to create much larger impacts e.g. on methyl donor supply and methylation potential, when viewed at the metabolic pathway or even metabolome level.

4.4.4   Peak seasonal windows

Previous analyses have used the peak rainy and dry seasonal windows for comparison[1,6]. The current MDEG-2 analyses enable a reflection on whether these windows are the most appropriate ones to consider when defining windows of vulnerability to micronutrient deficiencies, or lower methylation potential, for example. Clearly, biomarker variation over the year does not fall neatly into the peak rainy and dry seasons, and restricting analyses to these windows could miss seasonal differences at other time points. For example, comparing peak rainy and dry season concentrations for B2, choline and betaine suggests there is no difference (Table 4.1), yet Fourier term regression does justify a seasonal trend, and viewing variation over the whole year shows the peaks and troughs of their trends falling at periods outside the peak seasons.

A potential new window to consider is September to October, which is the end of the rainy

season and the start of the dry season. Here there are troughs for B2, folate, B12, choline, cysteine, PLP, lysine, histidine and leucine. There is not the same clear alignment of peak concentrations, but June may be another window to consider, when there are higher levels of folate, B2, PLP, lysine and histidine.

4.4.5   Limitations

There were a few planned metabolites that were not able to be measured. As discussed above, one of these was formate. Formate (CHO_2^-) is a very small molecule and it is possible that the Sigma-Aldrich ELISA-based assay I attempted to use was simply unable to detect the low levels of formate in this population (speculatively, formate assays may be better at capturing abnormally high levels to help diagnose methanol poisoning). I had also planned to measure S-adenosyl methionine (SAM) and S-adenosyl homocysteine (SAH) to investigate a more direct measure of methylation potential. However, these metabolites had degraded either with processing or storage time (with SAM converting into SAH) and were no longer reliable estimates. Despite not having SAM and SAH, homocysteine is a very useful measure of methylation potential (see Chapter 6 for a discussion on this). I had also wanted to measure decarboxylated SAM to see whether there might be an indication at certain times of the year that SAM was being used preferentially in the polyamine pathway over the transmethylation pathway.

In these analyses the additional metabolites considered in comparison to the Indicator group were the amino acids. As described above, their periodic variation was generally not as profound as for the core one one-carbon biomarkers, and the interpretation of their covariation is complex. Including information on amino acids other than those thought to influence one-carbon metabolism (methionine, cysteine, serine, glycine, tryptophan and histidine) may risk making the picture unnecessarily complex. However, as discussed in the introduction, the original aim was to collect as much metabolite information as possible to feed into a systems biology model, which then may have helped identify metabolic pathways more distal to core one-carbon ones that we might want to consider. The overall amino acid information may yet still be prove to be important when interpreting offspring DNA methylation data, when it comes in.

Performing longitudinal analyses on data made up of cross-sectional measurements throughout the year will be susceptible to much greater variability and reduced precision compared to repeated measurements in the same individual. However, the comparison of

the general shapes of the periodic patterns with the Indicator group was helpful in confirming the validity of this approach in the MDEG-2 dataset. These biomarker trends will be used to assess associations with infant DNA methylation data across the whole year when the offspring data is available. The Indicator group followed 30 women across the year and was a time- and resource-intensive approach that would not have been possible with the sample size used in the MDEG-2 dataset.

The dates of enrolment of women in the subsample onto the ENID trial ranged from February 2010 to July 2013. Over this period the onset of Ramadan was 11th August in 2010 and 20th July in 2012. It is possible that trends over the months of July and August therefore may show a 'Ramadan effect' brought on by the effects on fasting and increased intake of certain foods to break the fast. However, this window does not appear to coincide with any major peaks and troughs of the periodic patterns. Nonetheless, until nutritional status is assessed in a year where Ramadan lies outside of the July-August window it remains difficult to disentangle the relative contribution of Ramadan food intake practices versus more consistent underlying seasonal trends.

Plasma concentrations are imperfect measures of nutrient status for many reasons (see section 9.3.1 for more details), and these need to be considered at the time of interpretation. **Annex 4.3** details some of these factors for the core set of one-carbon-related biomarkers where information is available, including the relationship between dietary intake and plasma concentration, known changes throughout pregnancy, and issues related to storage and free-thaw cycles.

4.5   Conclusion

The periodic patterns obtained from this dataset, composed of cross-sectional data, generally replicated the trends seen for core one-carbon biomarkers in the longitudinal dataset of the Indicator group. Fitted periodic patterns using Fourier term regression models explained relatively little of the overall variation of the biomarker concentrations, mostly less than 10%. However, a closer inspection of the trends using the variation about the geometric mean and the CCV showed that some metabolites did show significant periodic variation, especially for folate, B2 and PLP. Small fluctuations in individual biomarkers may combine to exert larger effects on methylation potential, which should be explored more formally when matching infant DNA methylation data becomes available. The inclusion of additional amino

acid measurements in the expanded biomarker set did not provide easily interpretable information to further our understanding of the covariation of core one-carbon biomarkers. However, this may be a limitation of using plasma concentration data in the absence of kinetic data, which makes it difficult to explore the complicated dynamics of metabolism (including circular pathways and allosteric inhibition, for example). It may also be because some of the other crucial metabolites that help link the amino acid pools to the one-carbon pathways, such as formate, were not measured.

Looking at the overall picture of covariation of the metabolites there is no clear-cut season of deficiency and adequacy. The shape of the periodic patterns greatly differ, and the peaks and troughs of each biomarker rarely perfectly coincide. However, there does seem to be a rationale for exploring a new window of September-to-October for when infant methylation data is available, as this seems to be a period when several biomarkers are at (or near) their nadirs. The analyses in this chapter are helpful not only to validate previous findings of seasonal fluctuations in one-carbon biomarkers, but also in laying the groundwork for future analyses that will match this data with offspring DNA methylation data.

4.6   References

1.   Dominguez-Salas P, Moore SE, Cole D, et al. (2013) DNA methylation potential: dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women. *Am. J. Clin. Nutr.* **97**, 1217–27.

2.   Brosnan ME, MacMillan L, Stevens JR, et al. (2015) Division of labour: how does folate metabolism partition between one-carbon metabolism and amino acid oxidation? *Biochem. J.* **472**, 135–46.

3.   Rios-Avila L, Coats B, Chi Y-Y, et al. (2015) Metabolite profile analysis reveals association of vitamin B-6 with metabolites related to one-carbon metabolism and tryptophan catabolism but not with biomarkers of inflammation in oral contraceptive users and reveals the effects of oral contraceptives. *J. Nutr.* **145**, 87–95.

4.   Rios-Avila L, Coats B, Ralat M, et al. (2015) Pyridoxine supplementation does not alter in vivo kinetics of one-carbon metabolism but modifies patterns of one-carbon and tryptophan metabolites in vitamin B-6-insufficient oral contraceptive users. *Am. J. Clin. Nutr.* **102**, 616–25.

5.   da Silva VR, Rios-Avila L, Lamers Y, et al. (2013) Metabolite profile analysis reveals functional effects of 28-day vitamin B-6 restriction on one-carbon metabolism and tryptophan catabolic pathways in healthy men and women. *J. Nutr.* **143**, 1719–27.

6.   Dominguez-Salas P, Moore SE, Baker MS, et al. (2014) Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. *Nat. Commun.* **5**, 3746.

7.   D'Agostino RB, Belanger A & D'Agostino RB (1990) A Suggestion for Using Powerful and Informative Tests of Normality. *Am. Stat.* **44**, 316.

8.   Abbenhardt C, Miller JW, Song X, et al. (2014) Biomarkers of one-carbon metabolism are associated with biomarkers of inflammation in women. *J. Nutr.* **144**, 714–21.

9.   Forster S & Gariballa S (2005) Age as a determinant of nutritional status: A cross sectional study. *Nutr. J.* **4**, 28.

10.   Leslie W & Hankey C (2015) Aging, Nutritional Status and Health. *Healthc.* **3**, 648–58.

11.   Mathews F, Yudkin P, Smith RF, et al. (2000) Nutrient intakes during pregnancy: the influence of smoking status and age. *J. Epidemiol. Community Health* **54**, 17–23.

12.   Kimmons JE, Blanck HM, Tohill BC, et al. (2006) Associations between body mass index and the prevalence of low micronutrient levels among US adults. *MedGenMed* **8**, 59.

13.   Bjørke-Monsen A-L, Ulvik A, Nilsen RM, et al. (2016) Impact of Pre-Pregnancy BMI on B Vitamin and Inflammatory Status in Early Pregnancy: An Observational Cohort Study. *Nutrients* **8,** 776.

14.   Fulford AJ (2014) The coefficient of cyclic variation: a novel statistic to measure the magnitude of cyclic variation. *Emerg. Themes Epidemiol.* **11**, 15.

15.   Midttun Ø, Kvalheim G & Ueland PM (2013) High-throughput, low-volume, multianalyte quantification of plasma metabolites related to one-carbon metabolism using HPLC-MS/MS. *Anal. Bioanal. Chem.* **405**, 2009–17.

16.   Dhonukshe-Rutten RAM, de Vries JHM, de Bree A, et al. (2009) Dietary intake and status of folate and vitamin B12 and their association with homocysteine and cardiovascular disease in European populations. *Eur. J. Clin. Nutr.* **63**, 18–30.

17.   Eussen SJPM, Vollset SE, Hustad S, et al. (2010) Plasma vitamins B2, B6, and B12, and related genetic variants as predictors of colorectal cancer risk. *Cancer Epidemiol. Biomarkers Prev.* **19**, 2549–61.

18.   Ueland PM, Ulvik A, Rios-Avila L, et al. (2015) Direct and Functional Biomarkers of Vitamin B6 Status. *Annu. Rev. Nutr.* **35**, 33–70.

19.   Green R & Miller JW (2007) Vitamin B12. In *Handbook of Vitamins*, 4th ed. [ZempleniJ, Rucker R, McCormick D, et al., editors]. Boca Ralton, FL: Taylor & Francis Group.

20.   Lepage N, McDonald N, Dallaire L, et al. (1997) Age-specific distribution of plasma amino acid concentrations in a healthy pediatric population. *Clin. Chem.* **43**, 2397–402.

21.   Thurnham DI, McCabe LD, Haldar S, et al. (2010) Adjusting plasma ferritin concentrations to remove the effects of subclinical inflammation in the assessment of iron deficiency : *Am. J. Clin. Nutr.* **92**, 546–555.

22.   Midttun Ø, Theofylaktopoulou D, McCann A, et al. (2017) Circulating concentrations of biomarkers and metabolites related to vitamin status, one-carbon and the kynurenine pathways in US, Nordic, Asian, and Australian populations. *Am. J. Clin. Nutr.* **105**, 1314-1326.

23.   Dominguez-Salas P (2012) Epigenesis in humans: Can maternal methyl-donor-deficient diets induce epigenetic alterations in their offspring? Phd Thesis. London School of Hygiene and Tropical Medicine, University of London.

24.   Mason JB (2003) Biomarkers of Nutrient Exposure and Status in One-Carbon (Methyl) Metabolism. *J. Nutr.* **133**, 941S–947.

25.   Scott JM & Weir DG (1998) Folic acid, homocysteine and one-carbon metabolism: a review of the essential biochemistry. *J. Cardiovasc. Risk* **5**, 223–7.

26.   Velzing-Aarts F V, Holm PI, Fokkema MR, et al. (2005) Plasma choline and betaine and their relation to plasma homocysteine in normal pregnancy. *Am. J. Clin. Nutr.* **81**, 1383–9.

27.   McKeever MP, Weir DG, Molloy A, et al. (1991) Betaine-homocysteine methyltransferase: organ distribution in man, pig and rat and subcellular distribution in the rat. *Clin. Sci.* **81**, 551–6.

28.   Cho E, Zeisel SH, Jacques P, et al. (2006) Dietary choline and betaine assessed by food-frequency questionnaire in relation to plasma total homocysteine concentration in the Framingham Offspring Study. *Am. J. Clin. Nutr.* **83**, 905–11.

29.   Ueland PM (2011) Choline and betaine in health and disease. *J. Inherit. Metab. Dis.* **34**, 3–15.

30.   Selhub J (1999) Homocysteine metabolism. *Annu. Rev. Nutr.* **19**, 217–46.

31.   Davis SR, Stacpoole PW, Williamson J, et al. (2004) Tracer-derived total and folate-dependent homocysteine remethylation and synthesis rates in humans indicate that serine is the main one-carbon donor. *Am. J. Physiol. Endocrinol. Metab.* **286**, E272-9.

32.   Nijhout HF, Reed MC, Lam S-L, et al. (2006) In silico experimentation with a model of hepatic mitochondrial folate metabolism. *Theor. Biol. Med. Model.* **3**, 40.

33.   Morrow GP, MacMillan L, Lamarre SG, et al. (2015) In vivo kinetics of formate metabolism in folate-deficient and folate-replete rats. *J. Biol. Chem.* **290**, 2244–50.

34.   Lamarre SG, Morrow G, Macmillan L, et al. (2013) Formate: an essential metabolite, a biomarker, or more? *Clin. Chem. Lab. Med.* **51**, 571–8.

35.   Wu G (2009) Amino acids: metabolism, functions, and nutrition. *Amino Acids* **37**, 1–17.

36.   Wu G (2013) Functional amino acids in nutrition and health. *Amino Acids* **45**, 407–411. Springer Vienna.

37.   Wu G, Bazer FW, Burghardt RC, et al. (2011) Proline and hydroxyproline metabolism: implications for animal and human nutrition. *Amino Acids* **40**, 1053–1063.

Supplemental Table 4.1: Laboratory information and outlier justifications

Biomarker	Assay type	Intra-assay CV%	Inter-assay CV%	Outliers excluded	Outlier Justification	Insufficient plasma	Total included in analysis
Homocysteine	LC-MS	3.9	4.1	1	55.1µmol/L (11.1 z-scores)	0	349
Folate	Abbott Architect i1000	<4.0	<4.7	0	N/A	0	350
B2	LC-MS	1.3	7.4	2	>188nmol/L (>9.1 z-scores)	1	347
PLP	LC-MS	3.5	4.7	0	N/A	1	349
B12	Abbott Architect i1000	<5.6	<6.8	0	N/A	0	350
Choline	LC-MS	2.5	3.4	2	>23.7µmol/L (>8 z-scores)	0	348
Betaine	LC-MS	2.2	3.9	0	N/A	0	350
Methionine	LC-MS	4.1	3.9	0	N/A	0	350
Cysteine	LC-MS	2.6	2.9	0	N/A	0	350
DMG	LC-MS	2.4	3.6	4	>14.6µmol/L (>6.5 z-scores)	0	346
Alanine	Hitachi L-8900		5.8	0	N/A	4	346
Arginine	Hitachi L-8900		3.6	0	N/A	4	346
Aspartate	Hitachi L-8900		27.5	1	20.8µmol/L (7.0 z-scores)	4	345
Glutamate	Hitachi L-8900		12.0	2	>273.3µmol/L (>6.3 z-scores)	4	344
Glycine	Hitachi L-8900		4.4	0	N/A	4	346
Histidine	Hitachi L-8900		4.2	0	N/A	4	346
Isoleucine	Hitachi L-8900		4.9	0	N/A	4	346
Leucine	Hitachi L-8900		4.0	0	N/A	4	346
Lysine	Hitachi L-8900		3.6	0	N/A	4	346
Phenylalanine	Hitachi L-8900		7.4	0	N/A	4	346
Proline	Hitachi L-8900		4.0	1	711µmol/L (14.1 z-scores)	4	345
Serine	Hitachi L-8900		2.8	0	N/A	4	346
Threonine	Hitachi L-8900		3.1	0	N/A	4	346
Tyrosine	Hitachi L-8900		5.8	0	N/A	4	346
Valine	Hitachi L-8900		3.1	0	N/A	4	346
AGP	Cobas Integra 400plus	<16.0	<10.2	0	N/A	4	346

Abbreviations: AGP, Alpha-1-acid glycoprotein; CV, coefficient of variation; DMG, dimethylglycine; LC-MS, liquid chromatography mass spectrometry; PLP, pyrodixol-5'-phosphate.

Chapter 5 Maternal one-carbon metabolism and infant DNA methylation between contrasting seasonal environments: A case study from The Gambia

_____

Summary of chapter

Background: The periconceptional period is a time in which environmentally-induced changes to the epigenome could have significant consequences for offspring health. Metastable epialleles (MEs) are genomic loci demonstrating inter-individual variation in DNA methylation with intra-individual cross-tissue correlation, suggesting that methylation states are established in the very early embryo prior to gastrulation. In our previous Gambian studies we have shown that ME methylation states in the offspring are predicted by maternal levels of certain nutritional biomarkers around the time of conception.

Objective: We assess whether the profile of maternal biomarker predictors of offspring methylation differs between rainy and dry seasons in a population of rural Gambians, using a larger set of 50 recently identified MEs.

Methods: We measured 1-carbon biomarkers in maternal plasma back-extrapolated to conception, and cytosine-phosphate-guanine (CpG) methylation at 50 ME loci in their infants' blood at mean age 3.3 months (N=120 mother-child pairs). We tested for interactions between seasonality and effects of biomarker concentrations on mean ME methylation z-score. We used backwards stepwise linear regression to select the profile of nutritional predictors of methylation in each season, and repeated this analysis with biomarker principal components (PCs) to capture biomarker co-variation.

Results: We found preliminary evidence of seasonal differences in biomarker-methylation associations for folate, choline and homocysteine (interaction p values ≤0.03). Furthermore, in stratified analyses biomarker predictors of methylation changed between seasons. In the dry season B2 and methionine were positive predictors. In the rainy season, however, choline and B6 were positive predictors, and folate and B12 were negative ones. PC1 captured co-variation in the folate metabolism cycle and predicted methylation in dry season conceptions. PC2 represented the betaine remethylation pathway and predicted rainy

season methylation.

Conclusions: Underlying nutritional status may modify the association between nutritional biomarkers and methylation, and should be considered in future studies.

Notes

The chapter contains the final proofs sent to me before final publication. The narrative will not change from the version included here.

RESEARCH PAPER COVER SHEET

SECTION A – Student Details

Student	Philip James
Principal Supervisor	Dr Matt Silver
Thesis Title	Maternal nutrition, seasonality and epigenetics: an exploration of one-carbon metabolism and a novel nutritional supplement design in The Gambia.

SECTION B – Paper already published

Where was the work published?	Current Developments in Nutrition		
When was the work published?	12th October 2018		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N/A		
Have you retained the copyright for the work?*	No	Was the work subject to academic peer review?	Yes

If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

Multi-authored Work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper.

My supervisor originally thought of the idea of exploring seasonal interactions with nutrient exposure-methylation data. He provided me with the DNA methylation data to explore. I decided on the statistical approach and performed the analyses. I wrote the first draft of all sections of the paper. All co-authors provided comments on the narrative, which I incorporated into the final version.

Student Signature:

Date: 17th October 2018

Supervisor Signature:

Date: 17th October 2018

Maternal One-Carbon Metabolism and Infant DNA Methylation Between Contrasting Seasonal Environments: A Case Study from The Gambia

Philip T James [ID],[1] Paula Dominguez-Salas [ID],[2] Branwen J Hennig,[3] Sophie E Moore,[1,4] Andrew M Prentice,[1] and Matt J Silver [ID][1]

[1]Medical Research Council Unit The Gambia at the London School of Hygiene and Tropical Medicine, London, United Kingdom; [2]Department of Production and Population Health, Royal Veterinary College, London, United Kingdom; [3]Population Health, Science Division, Wellcome Trust, London, United Kingdom; and [4]Department of Women and Children's Health, King's College London, London, United Kingdom

ABSTRACT

Background: The periconceptional period is a time in which environmentally induced changes to the epigenome could have significant consequences for offspring health. Metastable epialleles (MEs) are genomic loci demonstrating interindividual variation in DNA methylation with intraindividual crosstissue correlation, suggesting that methylation states are established in the very early embryo before gastrulation. In our previous Gambian studies, we have shown that ME methylation states in the offspring are predicted by maternal concentrations of certain nutritional biomarkers around the time of conception.

Objective: We aimed to assess whether the profile of maternal biomarker predictors of offspring methylation differs between rainy and dry seasons in a population of rural Gambians, using a larger set of 50 recently identified MEs.

Methods: We measured 1-carbon biomarkers in maternal plasma back-extrapolated to conception, and cytosine-phosphate-guanine (CpG) methylation at 50 ME loci in their infants' blood at a mean age of 3.3 mo ($n = 120$ mother-child pairs). We tested for interactions between seasonality and effects of biomarker concentrations on mean ME methylation z score. We used backward stepwise linear regression to select the profile of nutritional predictors of methylation in each season and repeated this analysis with biomarker principal components (PCs) to capture biomarker covariation.

Results: We found preliminary evidence of seasonal differences in biomarker-methylation associations for folate, choline, and homocysteine (interaction P values ≤ 0.03). Furthermore, in stratified analyses, biomarker predictors of methylation changed between seasons. In the dry season, vitamin B-2 and methionine were positive predictors. In the rainy season, however, choline and vitamin B-6 were positive predictors, and folate and vitamin B-12 were negative predictors. PC1 captured covariation in the folate metabolism cycle and predicted methylation in dry season conceptions. PC2 represented the betaine remethylation pathway and predicted rainy season methylation.

Conclusions: Underlying nutritional status may modify the association between nutritional biomarkers and methylation, and should be considered in future studies. *Curr Dev Nutr* 2018;0:nzy082.

Introduction

The Developmental Origins of Health and Disease (DOHaD) hypothesis suggests that early-life environmental exposures affect lifelong health and disease risk (1–4). For example, exposure to the Dutch Hunger Winter famine in 1944–1945 across different stages of prepregnancy and pregnancy has been associated with lower birthweight (5), increased adult blood pressure

and obesity (6–8), and increased risk of schizophrenia (9). One plausible mechanism for these associations is through epigenetic modifications to the genome (10–12). Epigenetic processes encompass mitotically heritable changes to the genome that can alter gene expression without changing the underlying DNA sequence (13), and include DNA methylation (predominantly at cytosine-phosphate-guanine (CpG) sites), histone modifications, and RNA-based mechanisms (14).

Times of increased cell turnover, such as during fetal development, may be particularly susceptible to epigenetic errors or to adaptive modifications designed to capture early environmental cues (15, 16). Early embryonic development is a period of complex epigenetic remodeling and cell differentiation (17–19), and thus represents a critical window in which changes to the epigenetic program could have significant consequences for offspring health (20).

Metastable epialleles (MEs) are genomic loci whose (nongenetically determined) methylation state varies between individuals, but in whom variation is correlated across tissues originating from all 3 germ layers in a single individual (16, 21). This suggests the establishment of stochastic methylation states in the first few days after conception before separation into germ layers around gastrulation. ME methylation therefore provides a useful measure for studying the potential influence of the periconceptional environment on selected regions of the offspring epigenome (22, 23). ME methylation status in humans has been associated with obesity, immune function, and certain cancers (24–26).

A variety of nutritional and other environmental factors can impact the infant epigenome in utero through maternal exposure (20, 27, 28), including 1-carbon metabolites in the periconceptional period and during embryonic development (29). One-carbon metabolism refers to the interlinking reactions of the folate, choline, methionine, homocysteine, transsulfuration and transmethylation metabolic pathways (30, 31). DNA methylation is one of the numerous transmethylation reactions made possible by the donation of a methyl group from *S*-adenosylmethionine (SAM), forming *S*-adenosyl homocysteine (SAH) in the process (32). The SAM:SAH ratio has therefore been used as a proxy indicator of methylation potential (33). The 1-carbon pathways that enable transmethylation to occur rely on nutritional inputs in the form of methyl donors (e.g., folate, choline, betaine) and essential cofactors (e.g., vitamins B-2, B-6, and B-12) (30, 34). Nutritional status of the mother can therefore influence DNA methylation, and this is most clearly exemplified in animal models. In Agouti mouse experiments, pregnant dams fed a diet rich in vitamin B-12, folic acid, choline, and betaine gave birth to pups exhibiting increased methylation at the locus influencing the expression of the *Agouti* gene compared with controls. This resulted in changes to offspring fur color, appetite, adiposity, and glucose tolerance (27, 35). In humans, there is also evidence linking maternal nutrition to offspring DNA methylation, explored either as individual micronutrients or as proxy measures of nutrition such as famine and seasonality (36, 37). Although there is also much evidence linking DNA methylation to later phenotype (12, 38), studies fully exploring the continuum of maternal nutrient exposure, offspring DNA methylation, and later phenotype are relatively rare (39).

In a series of studies in rural Gambia, we have been able to exploit a seasonal "natural experiment," whereby a cycling pattern of rainy and dry seasons imposes strikingly different environmental, especially nutritional, exposures on the population. We have shown that plasma collected in nonpregnant women of child-bearing age contains higher concentrations of methyl donors and has a higher methylation potential in the peak rainy season (July to September) than in the peak dry season (February to April) (40). Furthermore, we found that seasonal differences in maternal periconceptional nutritional status are associated with offspring methylation at multiple MEs. Increased concentrations of vitamin B-2 and decreased concentrations of vitamin B-6, homocysteine, and cysteine predicted increased offspring mean methylation across 6 MEs (23), whereas offspring conceived in the rainy season had consistently higher concentrations of ME methylation in peripheral blood monocytes than those conceived in the dry season (22–26). However, our previous analyses did not explicitly test for an interaction with season for the associations between biomarker predictors and methylation.

Here, by exploring nutrient-season interactions, we extended our previous analyses to investigate whether the profile of maternal nutritional predictors of ME methylation varies between rainy and dry seasons. In doing so, we used a recently identified larger set of MEs associated with Gambian season of conception (SoC)-associated MEs (26) and explored in greater detail how covariation in the nutritional biomarkers can be captured in a principal components (PCs) model.

Methods

This paper utilizes data from 2 parallel studies: the Methyl Donors and Epigenetics (MDEG) study (23) and the Early Nutrition & Immune Development (ENID) Trial (41), both conducted in the rural West Kiang region of The Gambia.

Study population: The MDEG study

The MDEG study investigated the effects of periconceptional maternal biomarkers on infant DNA methylation at 6 candidate MEs (23). Women of reproductive age (18–45 y) were invited to participate and were followed monthly until pregnancy confirmation. Consenting women who conceived in the peak of the rainy season (July to September 2009) and the peak of the dry season (February to April 2010) were enrolled. Women provided a 10-mL fasting venous blood sample at the point they reported their first missed menses [mean (SD) 8.6 ± 4 weeks of gestation]. The following maternal 1-carbon biomarkers were analyzed: plasma folate, vitamin B-12, active vitamin B-12, choline, betaine, dimethylglycine (DMG), methionine, SAM, SAH, homocysteine (Hcy), cysteine, 4-pyridoxic acid (PA), pyridoxal (PL), pyridoxal-5′-phosphate (PLP), and erythrocyte riboflavin (vitamin B-2), as described previously (23). All biomarkers were back-extrapolated to the time of conception using seasonal trends from a cohort of 30 nonpregnant women from the same district, who provided fasted venous blood samples every month for a year, as previously detailed (40). Infant DNA was obtained from a 3-mL venepuncture taken 2–8 mo after delivery. In this analysis, we use a subset of 120 infants for whom we had analyzed genome-wide DNA methylation data (Gene Expression Omnibus accession GSE59592), obtained using the Illumina Infinium HumanMethytlation450 array ("450k array") (25, 42).

Selection of season of conception-associated ME loci from the ENID study

ME loci were identified using data from the ENID trial. Participants in ENID partially overlap with those in the MDEG study, although in the analysis described here, individuals from MDEG and ENID form distinct, nonoverlapping groups. $n = 50$ SoC-associated ME loci were identified as the intersection between loci identified in a recent screen for MEs on the 450k array and 2171 CpGs showing SoC-associated differential methylation using 450k data from 128 ENID blood samples from infants aged 24 mo (26). Selection of loci demonstrating both metastability and sensitivity to the periconceptional environment, each in independent samples, strengthens evidence that they are established in the early embryo (16, 25, 26). We provide details on the locations and genomic context of the 50 CpGs used in this analysis in **Supplementary Table 1.** Our use of ENID samples to identify SoC-associated MEs in this analysis carries a number of advantages. First, it offers an opportunity to validate observations of increased rainy SoC ME methylation across independent ENID and MDEG 450k methylation datasets. Second, annual patterns of Gambian seasonality mean that potential confounding due to the relation between SoC and season of sample collection is different between ENID (median age of collection 24 mo) and MDEG (median age 3 mo) cohorts, enabling more robust inference (43). Third, SoC effects identified using ENID infant 24 mo DNA are by definition more persistent than those identified in younger MDEG samples, making them potentially more robust candidates for use as biomarkers or mediators of later health outcomes.

Statistical analyses

Outcome: infant DNA methylation at 50 CpGs. The 50 SoC-associated ME loci on the 450k array identified using ENID methylation data (see above) were carried forward for use as candidates in the current analysis with 450k methylation data from 120 infants in the MDEG dataset, for which we had matching maternal plasma biomarker concentration data.

DNA methylation β values were adjusted for batch effects (25). CpG methylation across the 50 ME loci was highly correlated (Cronbach's α test reliability coefficient of 0.908). We therefore derived a univariate measure of ME methylation by converting methylation at each CpG into a z score [(individual observation – CpG mean)/CpG SD] and taking the mean of the methylation z scores across all 50 CpGs as our primary outcome measure.

Exposure variables: nutritional biomarkers. After removing variables demonstrating colinearity, the final list of nutritional exposure variables was: folate, active vitamin B-12, vitamin B-2, choline, betaine, DMG, SAM:SAH, Hcy, methionine, PLP, and cysteine. All nutritional biomarkers were treated as continuous variables. All biomarkers were log-transformed to improve normality, apart from SAM:SAH, which was left untransformed, and standardized before analyses.

In order to capture covariation of the 1-carbon biomarkers, we also conducted a PC analysis. Four PCs had an eigenvalue >1 and underwent orthogonal varimax rotation. We generated individual PC scores based on these loadings and used the resulting 4 PC variables in subsequent regression analyses.

Baseline characteristics. Because this study uses a subsample of 120 mother-child pairs from the original sample ($n = 166$) (23), we report the baseline characteristics again by SoC. Means (for continuous,

normal data) were compared using Student's t test; medians (of non-normal data) were compared using the Wilcoxon rank-sum test; and proportions were compared using a chi-squared test or Fisher exact test (for categories with sparse data).

Associations between nutritional exposures and infant DNA methylation: crude analyses. To validate the findings from the original MDEG study using this larger set of MEs, we ran linear regression models to assess the crude association between maternal nutritional exposures and SoC with infant mean methylation z score. To assess the hypothesis that the profile of nutritional predictors might change between seasons, we then included season as an interaction term in the association between the nutritional exposures and methylation, assessing the interaction using the likelihood ratio test.

Primary objective: Predictors of methylation by SoC. Given that the interaction tests justified stratifying the data by SoC, we then explored the predictors of methylation in each season separately using multivariable linear regression. We used an automatic backward stepwise approach for variable selection, using a P value of >0.2 as the criterion for removal from the model. Each regression model was run twice; first, using the individual 1-carbon biomarkers as exposures ("biomarker model") and second using the 4 PCs model.

All regression model residuals were checked for normality and met the assumptions of linear regression models. All models included 11 a priori confounders: maternal age, BMI, and gestational age at time of sample collection; infant sex and infant age at time of sample collection; and 5 methylation-derived white blood cell counts (25). All of these have previously been associated with methylation (44–47). We report likelihood ratio test results comparing the full model against the baseline model including a priori confounders only. Stata 14.0 (Stata Corporation) was used for all statistical analyses.

Ethical considerations

Ethical approvals for the ENID trial and MDEG study were given by The Gambia Government/MRC Joint Ethics Committee (SCC1126v2 and SCC1151, respectively). Consent was gained by signature or thumb print from mothers for their own participation and that of their child. All data were anonymized before analyses.

Results

Four PCs with an eigenvalue >1 explained 65.0% of the total variation seen in the 11 biomarkers (**Supplementary Table 2**). After rotation, PC1 was associated positively with folate and SAM:SAH, and inversely with Hcy. PC2 was strongly correlated positively with choline and betaine, and inversely with active vitamin B-12. PC3 was positively correlated with the amino acids methionine and cysteine, and PC4 was strongly correlated with PLP and active vitamin B-12. These 4 PCs explained more than half the variability of all biomarkers apart from active vitamin B-12 and B-2, which still had 54.3% and 60.0% unexplained respectively. **Figure 1** shows the correlation between the 1-carbon biomarkers and the PC loadings as a heat map.

Baseline maternal and infant characteristics are summarized in **Table 1**, detailed for the overall sample and by SoC. There was no difference in maternal age, gestational age, maternal BMI, or infant sex

Variable*	Rotated (orthogonal varimax) PCs†				Explained proportion of variance
	PC1	PC2	PC3	PC4	
Folate	0.51	0.15	0.11	0.22	0.729
Active B12	-0.03	-0.35	0.11	0.45	0.458
B2	0.22	-0.06	0.35	-0.14	0.400
Choline	-0.08	0.62	-0.04	0.11	0.751
Betaine	0.11	0.61	0.10	-0.02	0.717
Dimethylglycine	-0.29	0.24	-0.26	0.20	0.563
SAM:SAH ratio	0.47	-0.06	0.00	-0.04	0.591
Homocysteine	-0.57	0.02	0.27	-0.04	0.834
Methionine	0.06	0.18	0.54	-0.20	0.612
PLP	0.06	0.08	0.01	0.77	0.748
Cysteine	-0.16	-0.06	0.64	0.22	0.755

FIGURE 1 Heat map showing correlation between 1-carbon biomarkers and PCs after orthogonal rotation. *All log-transformed, standardized variables, apart from SAM:SAH, which is standardized but untransformed. †Cells shaded green for positive correlations and red for inverse correlations. Intensity of shading is proportional to strength of correlation. B12, vitamin B-12; B2, vitamin B-2; PC, principal component; PLP, pyridoxal-5′-phosphate; SAH, S-adenosyl homocysteine; SAM, S-adenosyl methionine.

by SoC. Infants conceived in the dry season were ∼0.5 mo younger at their DNA blood draw than were those conceived in the rainy season. Women conceiving in the rainy season had higher concentrations of folate, vitamin B-2, betaine, cysteine, and SAM:SAH, and lower concentrations of DMG, PLP, and Hcy than did those conceiving in the dry season. There were higher scores for PC1 and PC3, and lower scores for PC4 in the rainy season. We provide a detailed breakdown of the nutritional status of the population, stratified by season, in **Table 2.** Almost all women were deficient in vitamin B-2, >40% had low PLP status, ∼30% and 20% had low concentrations of betaine and choline respectively, 13% were folate-deficient, and participants were replete in methionine and cysteine. There was evidence to suggest that a greater proportion of women were folate-deficient in the dry season and that a higher proportion had low PLP concentrations in the rainy season.

The crude association between mean total CpG methylation (z scores) and each exposure (SoC, nutritional biomarkers and PCs) is shown in **Table 3**. Total mean methylation across the 50 CpG sites was 0.26 z scores higher in the rainy season than in the dry season (95% CI: 0.07, 0.45; $P = 0.008$). The SAM:SAH ratio was positively associated with methylation, and there was weak evidence to suggest that Hcy was inversely associated. Among the PCs, only PC1 was positively associated with methylation.

To justify stratified analyses, we first tested whether there was any interaction between the effect of the exposure on total mean methylation by SoC (Table 3). There was some evidence of an interaction with SoC for plasma folate, choline, and homocysteine (P value for interaction = 0.019, 0.030, and 0.030, respectively). There was no evidence of an interaction with season for any of the other biomarkers, and overall effect sizes remained small. For the PCs, only PC1 showed a different pattern of association with methylation by season (P value for interaction = 0.002).

TABLE 1 Maternal and infant characteristics, overall and stratified by season of conception[1]

Variable	Total		Dry season		Rainy season		P[2]
	n	Statistic	n	Statistic	n	Statistic	
Folate, nmol/L	117	15.4 (14.3–16.6)	59	12.6 (11.6–13.7)	58	18.9 (17.0–20.9)	<0.001
Active B$_{12}$, pmol/L	118	73.7 (68.1–79.8)	59	79.2 (71.4–87.8)	59	68.6 (60.7–77.6)	0.077
B$_2$, 1/EGRAC	113	0.45 (0.43–0.48)	57	0.41 (0.38–0.44)	56	0.51 (0.47–0.55)	<0.001
Choline, μmol/L	117	6.6 (6.2–6.70)	59	6.9 (6.3–7.5)	58	6.3 (5.9–6.8)	0.109
Betaine, μmol/L	118	18.8 (17.6–20.1)	59	17.5 (15.7–19.4)	59	20.2 (18.6–21.9)	0.033
Dimethylglycine, μmol/L	118	2.2 (1.9–2.4)	59	2.9 (2.5–3.3)	59	1.6 (1.4–1.8)	<0.001
SAM:SAH ratio	118	7.7 (7.2–8.2)	59	6.5 (6.0–7.1)	59	9.1 (8.4–9.7)	<0.001
Homocysteine, μmol/L	118	6.7 (6.3–7.2)	59	7.4 (6.8–8.1)	59	6.1 (5.7–6.6)	<0.001
Methionine, μmol/L	118	24.4 (23.5–25.3)	59	22.9 (21.8–24.1)	59	26.0 (24.7–27.3)	<0.001
Pyridoxal phosphate, nmol/L	118	21.9 (20.2–23.7)	59	23.7 (20.9–27.0)	59	20.1 (18.4–22.1)	0.041
Cysteine, μmol/L	118	197.3 (192.9–201.9)	59	192.6 (186.8–198.5)	59	202.2 (195.6–209.0)	0.032
PC1	111	0.00 ± 1.58	57	−0.89 ± 1.24	54	0.94 ± 1.36	<0.001
PC2	111	0.00 ± 1.34	57	−0.03 ± 1.46	54	0.04 ± 1.21	0.788
PC3	111	0.00 ± 1.26	57	−0.52 ± 1.14	54	0.55 ± 1.15	<0.001
PC4	111	0.00 ± 1.13	57	0.35 ± 1.18	54	−0.37 ± 0.95	<0.001
Maternal age, y	117	29.2 ± 6.7	58	28.2 ± 6.1	59	30.3 ± 7.2	0.089
Gestational age, d	120	60.7 ± 28.5	60	63.8 ± 27.8	60	57.6 ± 28.9	0.237
BMI, kg/m^2	120		60		60		0.866
Underweight		16.7 ± 20		15.0 ± 9		18.3 ± 11	
Normal		72.5 ± 87		73.3 ± 44		71.7 ± 43	
Overweight		10.8 ± 13		11.7 ± 7		10.0 ± 6	
Infant sex, % (n)	120		60		60		0.534
Female		48.3 (58)		51.7 (31)		45.0 (27)	
Male		51.7 (62)		48.3 (29)		55.0 (33)	
Infant age, mo	112	3.3 [3.11–3.77]	55	3.2 [3.07–3.21]	57	3.7 [3.31–4.0]	<0.001

[1]Values are geometric means (95% CIs) for maternal biomarkers; medians [IQRs] for infant age; means ± SDs for PCs, maternal age, gestational age, and BMI; and % (n) for infant sex. B$_{12}$, vitamin B-12; B$_2$, vitamin B-2; EGRAC, erythrocyte glutathione reductase activity coefficient; PC, principal component; SAH, S-adenosyl homocysteine; SAM, S-adenosyl methionine.
[2]Testing difference by season: Wilcoxon rank-sum test for nonnormal data, Student's t-test for normal data, chi-squared test for proportion.

TABLE 2 Maternal plasma biomarker status, overall and stratified by season of conception[1]

Variables	Cutoff for low/abnormal status	Overall (both seasons)		Dry season		Rainy season		P[2]
		n below cutoff	%	n	%	n	%	
Homocysteine	>15 μmol/L ([48])	2/118	1.7	2/59	3.4	0/59	0.0	0.496
Folate	<10 nmol/L ([49])	15/117	12.8	12/59	20.3	3/58	5.2	0.024
B$_2$	<0.77 (1/EGRAC) ([50])	109/113	96.5	57/57	100.0	52/56	92.9	0.057
PLP	<20 nmol/L ([51])[3]	50/118	42.4	19/59	32.2	31/59	52.5	0.025
Active B$_{12}$	<37 pmol/L ([52])	6/118	5.1	1/59	1.7	5/59	8.5	0.207
Choline	<5 μmol/L ([48])[3]	24/117	20.5	12/59	20.3	12/58	20.7	0.963
Betaine	<16 μmol/L ([48])[3]	35/118	29.7	22/59	37.3	13/59	22.0	0.070
Methionine	<20 μmol/L ([53])[4]	19/118	16.1	12/59	20.3	7/59	11.9	0.210
Cysteine	<36 μmol/L ([53])[4]	0/118	0.0	0/59	0.0	0/59	0.0	—

[1]B$_{12}$, vitamin B-12; B$_2$, vitamin B-2; EGRAC, erythrocyte glutathione reductase activity coefficient; PLP, pyridoxal-5′-phosphate.
[2]Test for seasonal difference in biomarker status. P values from chi-squared test, or Fisher's exact test (if any numerator <5).
[3]There are no clearly defined plasma cutoffs for deficiency. The suggested cutoffs indicate below the normal range and can be considered "low status."
[4]The amino acid cutoffs represent the 10th percentile of a healthy population age >16 y in Canada ([53]). Note that these cutoffs do not necessarily represent low status or deficiency.

In stratified analyses using backward stepwise regression, Hcy, vitamin B-2, methionine, and SAM:SAH were retained in the dry season multivariable biomarker model (**Table 4**). Of these selected variables, methionine was the strongest positive predictor of methylation, followed by SAM:SAH. Hcy was associated with decreasing methylation as in crude analyses, as was vitamin B-2. The full model explained 27.0% of total variance in methylation (adjusted R^2, model $P = 0.001$). In the dry season, PC model PC1 was the only covariate retained, and the model explained 18.7% of methylation variance (model $P = 0.009$).

In the rainy season biomarker model, a different profile of predictors was retained. SAM:SAH, choline, and PLP were associated with increasing methylation, whereas folate and active vitamin B-12 were associated with decreasing methylation (**Table 5**). The rainy season model explained 9.4% of methylation variance (adjusted R^2, model $P = 0.004$). In the rainy season PC model, PC2 was positively associated with methylation. PC1 and PC3 were also retained and showed weak inverse associations. The model, however, fitted poorly.

A graphical summary of the associations between predictors of methylation retained in the multivariable models by SoC is shown in **Figure 2**. This figure simplifies the above results by focusing on the PC associations, showing the switch of positive predictors of methylation from the folate pathway in the dry season to the choline/betaine pathway in the rainy season.

Discussion

This study extends our understanding of previously reported associations between 1-carbon biomarkers in mothers at the time of conception and DNA methylation at MEs in their infants. We have validated previous findings of increased methylation at MEs for rainy season conceptions, but found that maternal plasma biomarkers back-extrapolated to the time of conception generally demonstrate little individual effect on infant ME methylation, whether in crude analyses or in multivariable predictive models. There was some preliminary evidence to suggest an interaction between SoC and the association of maternal 1-carbon biomarkers with infant methylation.

PCs and metabolic pathways

The PC approach is useful for exploring covariation in biomarkers and their joint influence on methylation, although their biological interpretation can be difficult. However, our findings suggested the strongest loadings of each PC mapped onto different metabolic pathways. The major loadings for PC1 are involved in the folate metabolism cycle. The major form of folate in plasma is 5-methyl-tetrahydrofolate ([54]), which donates its methyl group to Hcy via methylene tetrahydrofolate reductase using vitamin vitamin B-12 as a coenzyme ([55]). The remethylation of Hcy forms methionine, which is then used to form SAM, thus explaining why the SAM:SAH ratio loading is also positively correlated with this PC along with folate. The inverse correlation of Hcy is expected because it is held in equilibrium with SAH, a buildup of which can impede the SAM to SAH reaction via product inhibition of methyltransferases ([56]). In contrast, PC2 loadings are positively associated with the betaine remethylation pathway. Choline is the precursor to betaine, which is formed via a 2-step oxidation reaction ([57]). Betaine donates its methyl group to homocysteine, catalyzed by betaine-homocysteine methyl transferase ([57]). Methionine and cysteine provide the major loadings for PC3. This could represent the transsulfuration pathway, because methionine provides the sulfur atom for cysteine synthesis, via the irreversible degradation of Hcy ([31]). It could also reflect that methionine and cysteine are dietary components found in similar food sources. The PC4 primary loading comes from PLP, which is particularly involved in 1-carbon metabolism as a coenzyme in the transsulfuration pathway, as well as being required to reduce THF to methylene-THF ([31]).

Crude analyses between SoC, 1-carbon related exposures, and methylation

Using an expanded set of 50 ME CpGs associated with SoC in samples from older infants ([26]), we validated our previous finding ([23]) of increased ME methylation in rainy season conceptions in the younger cohort analyzed here. Biomarker concentrations differed by season in ways that have been previously described ([23]), forming a profile with higher methylation potential in the rainy season than in the dry season. In the original MDEG study, we found that periconceptional concentrations of vitamin B-2 were positively associated with offspring

TABLE 3 Crude association between exposures and total mean CpG methylation *z* score, overall and stratified by season using linear regression[1]

Variable[2]	Overall (both seasons)			Stratified analysis by season			
	n	Coefficient (95% CI)[3]	P[4]	Season	Coefficient (95% CI)[3]	P[4]	P interaction[5]
Season[6]	109	0.26 (0.07, 0.45)	0.008				
Log folate	108	0.02 (−0.07, 0.11)	0.623	Dry	0.10 (−0.06, 0.26)	0.237	0.019
				Rainy	−0.14 (−0.26, −0.01)	0.034	
Log active B$_{12}$	109	−0.08 (−0.17, 0.02)	0.102	Dry	−0.04 (−0.18, 0.09)	0.549	0.758
				Rainy	−0.07 (−0.18, 0.05)	0.264	
Log B$_2$	105	0.04 (−0.05, 0.14)	0.347	Dry	−0.01 (−0.16, 0.14)	0.916	0.986
				Rainy	−0.01 (−0.14, 0.13)	0.924	
Log choline	108	0.02 (−0.07, 0.11)	0.642	Dry	−0.02 (−0.13, 0.09)	0.705	0.030
				Rainy	0.16 (0.02, 0.3)	0.023	
Log betaine	109	0.05 (−0.04, 0.14)	0.236	Dry	0.02 (−0.09, 0.13)	0.661	0.661
				Rainy	0.06 (−0.08, 0.2)	0.397	
Log dimethylglycine	109	−0.03 (−0.12, 0.06)	0.496	Dry	−0.01 (−0.15, 0.12)	0.869	0.227
				Rainy	0.10 (−0.05, 0.26)	0.175	
SAM:SAH ratio	109	0.12 (0.03, 0.20)	0.010	Dry	0.16 (0.02, 0.31)	0.029	0.086
				Rainy	0.00 (−0.14, 0.14)	0.986	
Log homocysteine	109	−0.09 (−0.18, 0.00)	0.054	Dry	−0.13 (−0.25, −0.01)	0.039	0.030
				Rainy	0.06 (−0.08, 0.21)	0.389	
Log methionine	109	0.05 (−0.04, 0.15)	0.277	Dry	0.09 (−0.03, 0.22)	0.132	0.062
				Rainy	−0.07 (−0.21, 0.07)	0.341	
Log PLP	109	0.05 (−0.04, 0.15)	0.277	Dry	0.00 (−0.11, 0.11)	0.999	0.762
				Rainy	−0.03 (−0.18, 0.13)	0.727	
Log cysteine	109	−0.05 (−0.14, 0.04)	0.303	Dry	−0.06 (−0.19, 0.07)	0.352	0.955
				Rainy	−0.06 (−0.18, 0.07)	0.371	
PC1	103	0.06 (0.00, 0.12)	0.037	Dry	0.13 (0.03, 0.22)	0.015	0.002
				Rainy	−0.07 (−0.16, 0.02)	0.139	
PC2	103	0.04 (−0.03, 0.11)	0.299	Dry	0.01 (−0.07, 0.10)	0.763	0.255
				Rainy	0.09 (−0.02, 0.19)	0.123	
PC3	103	0.00 (−0.09, 0.08)	0.909	Dry	−0.02 (−0.13, 0.09)	0.762	0.399
				Rainy	−0.08 (−0.21, 0.04)	0.191	
PC4	103	−0.06 (−0.14, 0.02)	0.126	Dry	−0.03 (−0.13, 0.08)	0.628	0.837
				Rainy	−0.04 (−0.18, 0.09)	0.533	

[1]B$_{12}$, vitamin B-12; B$_2$, vitamin B-2; CpG, cytosine-phosphate-guanine; PC, principal component; PLP, pyridoxal-5'-phosphate; SAH, *S*-adenosyl homocysteine; SAM, *S*-adenosyl methionine.
[2]All 1-carbon biomarkers log-transformed and standardized, apart from SAM:SAH (standardized only).
[3]Adjusted for maternal BMI at time of bled, gestational age at time of bleed, maternal age, infant sex, infant age, white blood cell composition.
[4]Two-tailed *t*-test for coefficient slope.
[5]Likelihood ratio test comparing models with and without interaction term.
[6]Season is coded 0 = dry 1 = rainy.

methylation and Hcy, whereas vitamin B-6 and cysteine were inversely associated (23). In these current analyses, we found the same association with Hcy, but not with vitamin B-2, vitamin B-6, or cysteine. Instead, in crude analyses, we found that SAM:SAH was positively associated with methylation, in line with the expected effect of these intermediary metabolites on methylation potential (33). The differences between the current and previous analyses could reflect the reduced sample size in this updated analysis (due to the smaller number of samples with Illumina 450k array data), additional adjustment covariates used, or the larger panel of MEs used to derive a univariate methylation score in the current study. The associations reported here help explain why there is evidence of a crude association between PC1 and methylation. Homocysteine has been associated with decreased methylation in several cross-sectional studies (58). Taken in isolation, folate did not show any association with methylation in this study, and this has also been the case in other studies (59–61). However, folate did load strongly onto PC1, which showed a positive association with methylation in the

dry season model. Increased maternal periconceptional folate status has been associated with increased methylation in infants at a differentially methylated region of *RXRA* (62). However, this pattern is not consistent, and inverse associations between fetal periconceptional folate exposure and methylation have also been found at *STX11*, *OTX2*, *TFAP2A*, *CYS1*, and *LEP* (39, 62, 63).

Predictors of methylation by SoC from multivariable analyses

In the dry season, the predictors broadly indicate that increasing methylation potential (increasing SAM:SAH and decreasing Hcy, most likely through the folate pathway looking at the PC1 loadings) contributes to higher levels of DNA methylation. However, in the rainy season, when there is higher plasma folate and lower plasma Hcy than in the dry season, the folate pathway unexpectedly switches to an inverse association, and we can hypothesize that the betaine remethylation pathway takes prominence. Although these simple regression models

TABLE 4 Multivariable linear regression: predictors of methylation (dry season)[1]

	Biomarker model			PC model	
Variable[2]	Coefficient (95% CI)[3]	P[4]	Variable	Coefficient (95% CI)[3]	P[4]
Log homocysteine	−0.16 (−0.31, −0.01)	0.040	PC1	0.15 (0.03, 0.26)	0.012
Log methionine	0.17 (0.04, 0.30)	0.011			
Log B$_2$	−0.10 (−0.25, 0.06)	0.214			
SAM:SAH ratio	0.13 (−0.05, 0.31)	0.164			
n	—	52			52
Overall model P[5]	—	0.001			0.009
R-squared	—	0.471			0.362
Adjusted R-squared	—	0.270			0.187

[1]B$_2$, vitamin B-2; PC, principal component; SAH, *S*-adenosyl homocysteine; SAM, *S*-adenosyl methionine.
[2]All 1-carbon biomarkers log-transformed and standardized, apart from SAM:SAH (standardized only).
[3]Adjusted for maternal BMI at time of bleed, gestational age at time of bleed, maternal age, infant sex, infant age, white blood cell composition.
[4]Two-tailed *t* test for coefficient slope.
[5]Likelihood ratio test comparing the final model with the model only including a priori confounders.

cannot address the specific molecular mechanisms involved, we can speculate on a few different possibilities. In the rainy season, we could be seeing the effect of feedback loops attenuating the influence of increased plasma folate. One-carbon metabolism is governed by intricately controlled feedback loops, which help protect the flux of metabolites through key reactions over a range of nutrient and cofactor concentrations (64, 65). Alternatively, it could be that the rainy season folate metabolism is at saturation, and the system can then only enhance SAM:SAH through the betaine remethylation pathway. This highlights the complexity of 1-carbon metabolism in human populations and suggests that the potential for effect modifiers, for example season in our Gambian setting, should be considered when modeling methyl donor pathways. Given the exploratory nature of our analyses and the small effect sizes reported, our findings need to be replicated in confirmatory studies.

Limitations

There are a number of unmeasured exposures that may follow a seasonal pattern and could contribute to differences in offspring methylation, thus confounding our results. These include other nutrition-related exposures [e.g., vitamin A (66), vitamin D (67, 68), and dietary polyphenols (69)], as well as other potential exposures such as maternal stress (70), toxin exposure (71), intrauterine growth restriction (72–74), maternal hyperglycemia (75), infection (76), and seasonal differences in the microbiome (77). Future nutritional intervention studies will help establish whether there is a causal association between differences in diet, nutritional biomarkers, and methylation.

Linear regression models rest on certain assumptions and have a number of limitations. For example, when modeling the effects of nutritional factors, minimum detection thresholds and saturation effects will introduce nonlinear effects that cannot be captured by linear regression. Stepwise regressions allow large numbers of predictors to be evaluated, but have been criticized for producing inflated coefficients, and for the fact that after the strongest predictor has been considered, there is little additional explanatory power for any correlated predictor (78). They are also known to be relatively unstable in that small changes in the data can cause one variable to be selected over another, which can then alter subsequent variable selection (79). It is therefore possible that some of the differences we see between the seasons are related to model instability rather than reflecting real changes. This point is exacerbated by the fact that our sample size was small, meaning stratified analyses by season may not have had adequate power to distinguish true differences.

TABLE 5 Multivariable linear regression: predictors of methylation (rainy season)[1]

	Biomarker model			PC model	
Variable[2]	Coefficient (95% CI)[3]	P[4]	Variable	Coefficient (95% CI)[4]	P[4]
Log folate	−0.20 (−0.35, −0.06)	0.008	PC1	−0.06 (−0.16, 0.03)	0.188
Log active B$_{12}$	−0.08 (−0.20, 0.04)	0.201	PC2	0.09 (−0.01, 0.20)	0.080
Log PLP	0.11 (−0.08, 0.30)	0.236	PC3	−0.08 (−0.20, 0.04)	0.169
Log choline	0.18 (0.03, 0.32)	0.018			
SAM:SAH ratio	0.14 (−0.02, 0.30)	0.093			
n	—	51			51
Overall model P[5]	—	0.004			0.052
R-squared	—	0.366			0.243
Adjusted R-squared	—	0.094			−0.023

[1]B$_{12}$, vitamin B-12; PC, principal component; PLP, pyridoxal-5′-phosphate; SAH, *S*-adenosyl homocysteine; SAM, *S*-adenosyl methionine.
[2]All 1-carbon biomarkers log-transformed and standardized, apart from SAM:SAH (standardized only).
[3]Adjusted for maternal BMI at time of bleed, gestational age at time of bleed, maternal age, infant sex, infant age white blood cell composition.
[4]Two-tailed *t*-test for coefficient slope.
[5]Likelihood ratio test comparing the final model with the model only including a priori confounders. PC, principal component.
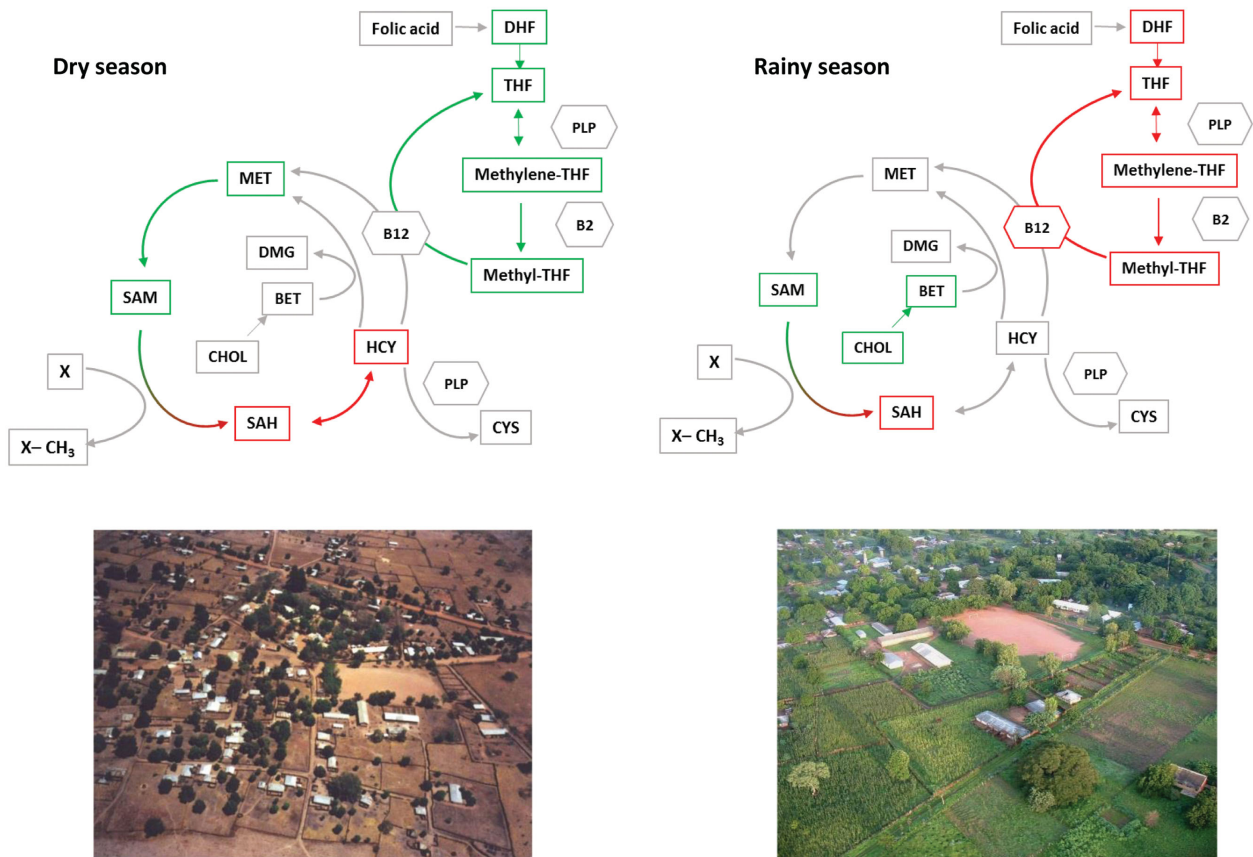
FIGURE 2 Simplified summaries of metabolic pathways acting as positive (green) and negative (red) predictors of methylation by season of conception, drawn from the retained principal-component coefficients in the multivariate models. Photos show Keneba, West Kiang district, The Gambia. Photo credit: Andrew Prentice. BET, betaine; B12, vitamin B-12; B2, vitamin B-2; CHOL, choline; CYS, cysteine; DHF, dihydrofolate; DMG, dimethyl glycine; Hcy, homocysteine; MET, methionine; methylene-THF, $N^{5,10}$-methylene tetrahydrofolate; methyl-THF, N^5-methyl tetrahydrofolate; PLP, pyridoxal-5′-phosphate; SAH, S-adenosyl homocysteine; SAM, S-adenosyl methionine; THF, tetrahydrofolate; X, acceptor.

The use of PCs gives some further insight into the joint effect of correlated biomarkers, offering an analysis strategy that lies conceptually between the consideration of biomarkers in isolation, and more sophisticated approaches that attempt to model the full complexity of metabolic networks. PC regression models are, however, hard to interpret. There are other models that are designed to help generate an understanding of how 1-carbon pathways interact (often in nonlinear ways), for example by estimating fluxes of metabolites through the pathways under given scenarios, and within specific cellular compartments (50, 80–82). Although these models are mathematically sophisticated, many are based on kinetic data that can be difficult to obtain at the population level. Furthermore, there is a need to generate models that can integrate plasma concentration data, the most common and accessible type of experimental data used for human in vivo studies. A promising way forward is within the field of systems biology, an integrative discipline that analyses complex datasets to help generate hypotheses, which can be experimentally validated and used to improve computer modeling in an iterative fashion (83). However, despite the limitations of the linear regression models we used, they can still play a role in hypothesis generation.

Conclusions

Before this study, we had observed that methylation at 6 MEs is higher among infants conceived in the rainy season than in those conceived in the dry season, and this trend has been seen again in a larger set of 50 MEs in the current analysis. However, we had not previously investigated whether the same combination of methyl donors and cofactors were consistently associated with methylation, or whether there was an interaction with season. In this current analysis, we find preliminary evidence to suggest that the rainy and dry seasons in The Gambia have a different set of maternal nutritional predictors of infant methylation. However, larger sample sizes and more sophisticated ways of modeling the complex nonlinear interrelations of metabolites are needed to further our understanding of what might trigger a switch between different methylation pathways at the molecular level.

Although there is still much work to do to complete our understanding of underlying mechanisms, our findings highlight potential considerations for future study design. If underlying nutritional status (partially captured in this study by the observed seasonal variations in plasma biomarker concentrations) influences the predictors of DNA methylation, then this would be applicable to populations with

heterogeneous patterns of dietary intake, whether seasonally driven or otherwise. This suggests that studies would benefit from collecting detailed information on nutritional status to assess if underlying nutritional status acts as an effect modifier. In observational studies, this information may help to explain contradicting associations between nutrition and other environmental exposures and DNA methylation or, in the case of trials, between nutritional interventions and DNA methylation. Such considerations might also inform the timing of future studies if there are seasonal dietary intake variations, or the targeting of subgroups in the context of populations with broad variation in nutritional status. In summary, the underlying nutritional status could be an essential piece of information to help disentangle the often complex and contradictory findings from nutritional epigenetics studies.

Acknowledgments

References

1. Barker DJ, Osmond C, Golding J, Kuh D, Wadsworth ME. Growth in utero, blood pressure in childhood and adult life, and mortality from cardiovascular disease. BMJ 1989;298:564–7.
2. Barker DJ, Gluckman PD, Godfrey KM, Harding JE, Owens JA, Robinson JS. Fetal nutrition and cardiovascular disease in adult life. Lancet 1993;341:938–41.
3. Curhan GC, Chertow GM, Willett WC, Spiegelman D, Colditz GA, Manson JE, Speizer FE, Stampfer MJ. Birth weight and adult hypertension and obesity in women. Circulation 1996;94:1310–5.
4. Eriksson JG, Forsen T, Tuomilehto J, Winter PD, Osmond C, Barker DJP. Catch-up growth in childhood and death from coronary heart disease: longitudinal study. BMJ 1999;318:427–31.
5. Smith CA. The effect of wartime starvation in Holland upon pregnancy and its product. Am J Obstet Gynecol 1947;53:599–608.
6. Roseboom TJ, van der Meulen JH, van Montfrans GA, Ravelli AC, Osmond C, Barker DJ, Bleker OP. Maternal nutrition during gestation and blood pressure in later life. J Hypertens 2001;19:29–34.
7. Stein AD, Zybert PA, van der Pal-de Bruin K, Lumey LH. Exposure to famine during gestation, size at birth, and blood pressure at age 59 y: Evidence from the Dutch Famine. Eur J Epidemiol 2006;21:759–65.
8. Ravelli GP, Stein ZA, Susser MW. Obesity in young men after famine exposure in utero and early infancy. N Engl J Med 1976;295:349–53.
9. Susser E, Neugebauer R, Hoek HW, Brown AS, Lin S, Labovitz D, Gorman JM. Schizophrenia after prenatal famine. Further evidence. Arch Gen Psychiatry 1996;53:25–31.
10. Heijmans BT, Tobi EW, Stein AD, Putter H, Blauw GJ, Susser ES, Slagboom PE, Lumey LH. Persistent epigenetic differences associated with prenatal exposure to famine in humans. Proc Natl Acad Sci USA 2008;105:17046–9.
11. Tobi EW, Slieker RC, Stein AD, Suchiman HED, Slagboom PE, van Zwet EW, Heijmans BT, Lumey L. Early gestation as the critical time-window for changes in the prenatal environment to affect the adult human blood methylome. Int J Epidemiol 2015;44:1211–23.
12. Lillycrop KA, Burdge GC. Epigenetic mechanisms linking early nutrition to long term health. Best Pract Res Clin Endocrinol Metab 2012;26:667–76.
13. Jaenisch R, Bird A. Epigenetic regulation of gene expression: How the genome integrates intrinsic and environmental signals. Nat Genet 2003;33 Suppl:245–54.
14. Fuks F. DNA methylation and histone modifications: teaming up to silence genes. Curr Opin Genet Dev 2005;15:490–5.
15. Langley-Evans SC. Nutrition in early life and the programming of adult disease: a review. J Hum Nutr Diet 2015;28:1–14.
16. Kessler NJ, Waterland RA, Prentice AM, Silver MJ. Establishment of environmentally sensitive DNA methylation states in the very early human embryo. Sci Adv 2018;4:eaat2624.
17. Seisenberger S, Peat JR, Hore TA, Santos F, Dean W, Reik W. Reprogramming DNA methylation in the mammalian life cycle: building and breaking epigenetic barriers. Philos Trans R Soc Lond B Biol Sci 2013;368:20110330.
18. Messerschmidt DM, Knowles BB, Solter D. DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. Genes Dev 2014;28:812–28.
19. Smallwood SA, Kelsey G. De novo DNA methylation: a germ cell perspective. Trends Genet 2012;28:33–42.
20. Perera F, Herbstman J. Prenatal environmental exposures, epigenetics, and disease. Reprod Toxicol 2011;31:363–73.
21. Rakyan VK, Blewitt ME, Druker R, Preis JI, Whitelaw E. Metastable epialleles in mammals. Trends Genet 2002;18:348–51.
22. Waterland RA, Kellermayer R, Laritsky E, Rayco-Solon P, Harris RA, Travisano M, Zhang W, Torskaya MS, Zhang J, Shen L, et al. Season of conception in rural Gambia affects DNA methylation at putative human metastable epialleles. PLoS Genet 2010;6:e1001252.
23. Dominguez-Salas P, Moore SE, Baker MS, Bergen AW, Cox SE, Dyer RA, Fulford AJ, Guan Y, Laritsky E, Silver MJ, et al. Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. Nat Commun 2014;5:3746.
24. Kühnen P, Handke D, Waterland RA, Hennig BJ, Silver M, Fulford AJ, Dominguez-Salas P, Moore SE, Prentice AM, Spranger J, et al. Interindividual variation in DNA methylation at a putative POMC metastable epiallele is associated with obesity. Cell Metab 2016;24:502–9.
25. Silver MJ, Kessler NJ, Hennig BJ, Dominguez-Salas P, Laritsky E, Baker MS, Coarfa C, Hernandez-Vargas H, Castelino JM, Routledge MN, et al. Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. Genome Biol 2015;16:118.
26. Van Baak TE, Coarfa C, Dugué P-A, Fiorito G, Laritsky E, Baker MS, Kessler NJ, Dong J, Duryea JD, Silver MJ, et al. Epigenetic supersimilarity of monozygotic twin pairs. Genome Biol 2018;19:2.
27. Waterland RA, Michels KB. Epigenetic epidemiology of the developmental origins hypothesis. Annu Rev Nutr 2007;27:363–88.
28. Tammen S, Friso S, Choi SW. Epigenetics: The link between nature and nurture. Mol Aspects Med 2013;34:753–64.
29. Steegers-Theunissen RPM, Twigt J, Pestinger V, Sinclair KD. The periconceptional period, reproduction and long-term health of offspring: The importance of one-carbon metabolism. Hum Reprod Update 2013;19:640–55.
30. Fox JT, Stover PJ. Folate-mediated one-carbon metabolism. Vitam Horm 2008;79:1–44.
31. Selhub J. Homocysteine metabolism. Annu Rev Nutr 1999;19:217–46.
32. Fontecave M, Atta M, Mulliez E. S-adenosylmethionine: Nothing goes to waste. Trends Biochem Sci 2004;29:243–9.
33. Mason JB. Biomarkers of nutrient exposure and status in one-carbon (methyl) metabolism. J Nutr 2003;133:941S–947.
34. Bertolo RF, McBreairty LE. The nutritional burden of methylation reactions. Curr Opin Clin Nutr Metab Care 2013;16:102–8.
35. Waterland RA, Jirtle RL. Transposable elements: Targets for early nutritional effects on epigenetic gene regulation. Mol Cell Biol 2003;23:5293–300.
36. Jiménez-Chillarón JC, Díaz R, Martínez D, Pentinat T, Ramón-Krauel M, Ribó S, Plösch T. The role of nutrition on epigenetic modifications and their implications on health. Biochimie 2012;94:2242–63.
37. Lee H-S. Impact of maternal diet on the epigenome during in utero life and the developmental programming of diseases in childhood and adulthood. Nutrients 2015;7:9492–507.

38. Godfrey KM, Costello PM, Lillycrop KA. The developmental environment, epigenetic biomarkers and long-term health. J Dev Orig Health Dis 2015;6:399–406.

39. James P, Sajjadi S, Tomar AS, Saffari A, Fall CHD, Prentice AM, Shrestha S, Issarapu P, Yadav DK, Kaur L, et al. Candidate genes linking maternal nutrient exposure to offspring health via DNA methylation: A review of existing evidence in humans with specific focus on one-carbon metabolism. Int J Epidemiol 2018;dyy153.

40. Dominguez-Salas P, Moore SE, Cole D, da Costa K-A, Cox SE, Dyer RA, Fulford AJC, Innis SM, Waterland RA, Zeisel SH, et al. DNA methylation potential: Dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women. Am J Clin Nutr 2013;97:1217–27.

41. Moore SE, Fulford AJ, Darboe MK, Jobarteh ML, Jarjou LM, Prentice AM. A randomized trial to investigate the effects of pre-natal and infant nutritional supplementation on infant immune development in rural Gambia: The ENID trial: Early Nutrition and Immune Development. BMC Pregnancy Childbirth 2012;12:107.

42. Hernandez-Vargas H, Castelino J, Silver MJ, Dominguez-Salas P, Cros M-P, Durand G, Le Calvez-Kelm F, Prentice AM, Wild CP, Moore SE, et al. Exposure to aflatoxin B1 in utero is associated with DNA methylation in white blood cells of infants in The Gambia. Int J Epidemiol 2015;44:1238–48.

43. Munafò MR, Davey Smith G. Repeating experiments is not enough. Nature 2018;553(7689):399–401.

44. Schroeder JW, Conneely KN, Cubells JC, Kilaru V, Newport DJ, Knight BT, Stowe ZN, Brennan PA, Krushkal J, Tylavsky FA, et al. Neonatal DNA methylation patterns associate with gestational age. Epigenetics 2011;6:1498–504.

45. Sharp GC, Lawlor DA, Richmond RC, Fraser A, Simpkin A, Suderman M, Shihab HA, Lyttleton O, McArdle W, Ring SM, et al. Maternal pre-pregnancy BMI and gestational weight gain, offspring DNA methylation and later offspring adiposity: Findings from the Avon Longitudinal Study of Parents and Children. Int J Epidemiol 2015;44:1288–304.

46. Herbstman JB, Wang S, Perera FP, Lederman SA, Vishnevetsky J, Rundle AG, Hoepner L, Qu L, Tang D. Predictors and consequences of global DNA methylation in cord blood and at three years. PLoS One 2013;8: e72824.

47. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. Genome Biol 2014;15:R31.

48. Midttun Ø, Kvalheim G, Ueland PM. High-throughput, low-volume, multianalyte quantification of plasma metabolites related to one-carbon metabolism using HPLC-MS/MS. Anal Bioanal Chem 2013;405: 2009–17.

49. Dhonukshe-Rutten RAM, de Vries JHM, de Bree A, van der Put N, van Staveren WA, de Groot LCPGM. Dietary intake and status of folate and vitamin vitamin B-12 and their association with homocysteine and cardiovascular disease in European populations. Eur J Clin Nutr 2009;63:18–30.

50. Hill MHE, Bradley A, Mushtaq S, Williams EA, Powers HJ. Effects of methodological variation on assessment of riboflavin status using the erythrocyte glutathione reductase activation coefficient assay. Br J Nutr 2009;102:273–8.

51. Ueland PM, Ulvik A, Rios-Avila L, Midttun Ø, Gregory JF. Direct and functional biomarkers of vitamin vitamin B-6 status. Annu Rev Nutr 2015;35:33–70.

52. Loikas S, Löppönen M, Suominen P, Møller J, Irjala K, Isoaho R, Kivelä S-L, Koskinen P, Pelliniemi T-T. RIA for serum holo-transcobalamin: method evaluation in the clinical laboratory and reference interval. Clin Chem 2003;49:455–62.

53. Lepage N, McDonald N, Dallaire L, Lambert M. Age-specific distribution of plasma amino acid concentrations in a healthy pediatric population. Clin Chem 1997;43:2397–402.

54. Bailey L. Folate in health and disease. 2nd edition. Boca Raton, FL: CRC Press, Taylor & Francis Group. 2009.

55. Scott JM, Weir DG. Folic acid, homocysteine and one-carbon metabolism: a review of the essential biochemistry. J Cardiovasc Risk 1998;5:223– 7.

56. Scotti M, Stella L, Shearer EJ, Stover PJ. Modeling cellular compartmentation in one-carbon metabolism. Wiley Interdiscip Rev Syst Biol Med 2013;5:343–65.

57. Ueland PM. Choline and betaine in health and disease. J Inherit Metab Dis 2011;34:3–15.

58. Zhou S, Zhang Z, Xu G. Notable epigenetic role of hyperhomocysteinemia in atherogenesis. Lipids Health Dis 2014;13:134.

59. van Mil NH, Bouwland-Both MI, Stolk L, Verbiest MMPJ, Hofman A, Jaddoe VW V, Verhulst FC, Eilers PHC, Uitterlinden AG, Steegers EAP, et al. Determinants of maternal pregnancy one-carbon metabolism and newborn human DNA methylation profiles. Reproduction 2014;148:581–92.

60. Ba Y, Yu H, Liu F, Geng X, Zhu C, Zhu Q, Zheng T, Ma S, Wang G, Li Z, et al. Relationship of folate, vitamin vitamin B-12 and methylation of insulin-like growth factor-II in maternal and cord blood. Eur J Clin Nutr 2011;65:480–5.

61. Fryer AA, Emes RD, Ismail KMK, Haworth KE, Mein C, Carroll WD, Farrell WE. Quantitative, high-resolution epigenetic profiling of CpG loci identifies associations with cord blood plasma homocysteine and birth weight in humans. Epigenetics 2011;6:86–94.

62. Pauwels S, Ghosh M, Duca RC, Bekaert B, Freson K, Huybrechts I, Langie SAS, Koppen G, Devlieger R, Godderis L. Maternal intake of methyl-group donors affects DNA methylation of metabolic genes in infants. Clin Epigenetics 2017;9:16.

63. Gonseth S, Roy R, Houseman EA, de Smith AJ, Zhou M, Lee S-T, Nusslé S, Singer AW, Wrensch MR, Metayer C, et al. Periconceptional folate consumption is associated with neonatal DNA methylation modifications in neural crest regulatory and cancer development genes. Epigenetics 2015;10:1166–76.

64. Reed MC, Gamble M V, Hall MN, Nijhout HF. Mathematical analysis of the regulation of competing methyltransferases. BMC Syst Biol 2015;9:69.

65. Nijhout HF, Best J, Reed MC. Escape from homeostasis. Math Biosci 2014;257:104–10.

66. Feng Y, Zhao L-Z, Hong L, Shan C, Shi W, Cai W. Alteration in methylation pattern of GATA-4 promoter region in vitamin A-deficient offspring's heart. J Nutr Biochem 2013;24:1373–80.

67. Pereira F, Barbáchano A, Singh PK, Campbell MJ, Muñoz A, Larriba MJ. Vitamin D has wide regulatory effects on histone demethylase genes. Cell Cycle 2012;11:1081–9.

68. Harvey NC, Sheppard A, Godfrey KM, McLean C, Garratt E, Ntani G, Davies L, Murray R, Inskip HM, Gluckman PD, et al. Childhood bone mineral content is associated with methylation status of the RXRA promoter at birth. J Bone Miner Res 2014;29:600–7.

69. Fang M, Chen D, Yang CS. Dietary polyphenols may affect DNA methylation. J Nutr 2007;137:223S–228.

70. Babenko O, Kovalchuk I, Metz GAS. Stress-induced perinatal and transgenerational epigenetic programming of brain development and mental health. Neurosci Biobehav Rev 2014;48:70–91.

71. Anway MD, Skinner MK. Epigenetic transgenerational actions of endocrine disruptors. Endocrinology 2006;147:S43–9.

72. Bouwland-Both MI, van Mil NH, Stolk L, Eilers PHC, Verbiest MMPJ, Heijmans BT, Tiemeier H, Hofman A, Steegers EAP, Jaddoe VW V, et al. DNA methylation of IGF2DMR and H19 is associated with fetal and infant growth: The generation R study. PLoS One 2013;8:e81731.

73. Einstein F, Thompson RF, Bhagat TD, Fazzari MJ, Verma A, Barzilai N, Greally JM. Cytosine methylation dysregulation in neonates following intrauterine growth restriction. PLoS One 2010;5:e8887.

74. Toure DM, Baccaglini L, Opoku ST, Barnes-Josiah D, Cox R, Hartman T, Klinkebiel D. Epigenetic dysregulation of insulin-like growth factor (IGF)-related genes and adverse pregnancy outcomes: A systematic review. J Matern Fetal Neonatal Med 2016;18:1–11.

75. El Hajj N, Schneider E, Lehnen H, Haaf T. Epigenetics and life-long consequences of an adverse nutritional and diabetic intrauterine environment. Reproduction 2014;148:R111–20.

76. Claycombe KJ, Brissette CA, Ghribi O. Epigenetics of inflammation, maternal infection, and nutrition. J Nutr 2015;145:1109S–15S.

77. Davie JR. Inhibition of histone deacetylase activity by butyrate. J Nutr 2003;133:2485S–2493.

78. Tibshirani R, Johnstone I, Hastie T, Efron B. Least angle regression. Ann Stat 2004;32:407–99.

79. Hesterberg T, Choi NH, Meier L, Fraley C. Least angle and ℓ 1 penalized regression: A review. Stat Surv 2008;2:61–93.

80. Nijhout HF, Reed MC, Lam S-L, Shane B, Gregory JF, Ulrich CM. In silico experimentation with a model of hepatic mitochondrial folate metabolism. Theor Biol Med Model 2006;3:40.

81. Ulrich CM, Reed MC, Nijhout HF. Modeling folate, one-carbon metabolism, and DNA methylation. Nutr Rev 2008;66 Suppl 1:S27–30.

82. Obeid R, Hartmuth K, Herrmann W, Gortner L, Rohrer TR, Geisel J, Reed MC, Nijhout HF. Blood biomarkers of methylation in Down syndrome and metabolic simulations using a mathematical model. Mol Nutr Food Res 2012;56:1582–9.

83. Thiele I, Swainston N, Fleming RMT, Hoppe A, Sahoo S, Aurich MK, Haraldsdottir H, Mo ML, Rolfsson O, Stobbe MD, et al. A community-driven global reconstruction of human metabolism. Nat Biotechnol 2013;31:419–25.

Supplementary Table 5.1: Location and genomic context of the 50 CpGs used in analyses*

CpG ID	Chromo-some	Position	Gene name (UCSC)	Gene region feature category (UCSC)	Relationship to Canonical CpG Island[†]	Enhancer element
cg08787039	1	182921857	C1orf14	Body	Island	
cg15025200	1	152161237			North Shore	
cg18388270	1	149212246				
cg17598339	2	170624727			Island	Yes
cg00817731	3	46759475	PRSS50	TSS200	Island	
cg01788113	3	46759472	PRSS50	TSS200	Island	
cg02834909	3	46759438	PRSS50	TSS200	Island	
cg16975614	3	105601834				
cg19668234	3	46759449	PRSS50	TSS200	Island	
cg20814179	4	940893	TMEM175	5'UTR	Island	
cg23206463	5	561769			North Shore	
cg02735058	6	16181250				
cg09894276	6	169977394	WDR27	Body	Island	
cg10218546	6	32729823	HLA-DQB2	Body	Island	
cg15189031	6	34499314	PACSIN1	Body	Island	
cg18609891	6	32729647	HLA-DQB2	Body	Island	
cg23414861	6	74064064	DPPA5	TSS200	Island	
cg08767820	7	51539583				Yes
cg10614809	7	51539190				Yes
cg14329783	7	75779857			Island	
cg06565641	8	141359674	TRAPPC9;TRAPPC9	Body;Body	South Shore	Yes
cg07092985	8	55533887	RP1	Body	Island	
cg07234876	8	600039			Island	
cg23958373	8	599963			Island	
cg06159404	10	43846376			Island	
cg01794156	11	397077	PKP3	Body	Island	
cg05940452	11	118842484	FOXR1;FOXR1	5'UTR;1stExon	Island	
cg23188684	11	67383651			Island	Yes
cg11258452	12	119594454	SRRM4	Body	South Shelf	
cg23887609	12	130822674	PIWIL1;PIWIL1	1stExon;5'UTR	Island	

CpG ID	Chromo-some	Position	Gene name (UCSC)	Gene region feature category (UCSC)	Relationship to Canonical CpG Island[†]	Enhancer element
cg24229701	12	130821962	PIWIL1	TSS1500	North Shore	
cg24838063	12	130822603	PIWIL1	TSS200	Island	
cg26677194	12	130822605	PIWIL1	TSS200	Island	
cg14089267	13	23412409			Island	
cg14395298	13	23412250			Island	
cg25576711	14	24780557	LTB4R2;LTB4R2;LTB4R2CIDEB;CIDEB;LTB4R	Body;1stExon;Body;1stExon;5'UTR;TSS200	Island	
cg26310551	14	24780540	LTB4R2;LTB4R2;LTB4R2CIDEB;CIDEB;LTB4R	Body;1stExon;Body;1stExon;5'UTR;TSS200	Island	
cg01232511	16	2867446	PRSS21;PRSS21;PRSS21	Body;Body;Body	Island	
cg08288433	17	20799511	CCDC144NL	TSS200	Island	
cg13590055	18	77917647	LOC100130522;LOC100130522;PARD6G	Body;Body;3'UTR	Island	
cg21237861	18	77918142	PARD6G;LOC100130522;LOC100130522	Body;Body;Body	Island	
cg26211724	18	57636670			Island	
cg02452966	19	39056216	RYR1;RYR1	Body;Body	Island	
cg05345154	19	2250901	AMH	Body	Island	
cg15035382	19	40728255	CNTD2	3'UTR	Island	
cg17434634	19	49522954			Island	Yes
cg19605788	19	40729386	CNTD2	Body	Island	Yes
cg19712189	19	33210851	TDRD12;TDRD12	5'UTR;1stExon	Island	
cg24923931	19	40904691	PRX;PRX	Body;Body	Island	
cg13403462	20	32256071	NECAB3;NECAB3;C20orf134;C20orf134	Body;Body;1stExon;3'UTR	South Shore	

*All details from Infinium HumanMethylation450K Manifest File v1.2 using coordinates from Genome Build 37.

[†]Shores - 0-2 kb from CpG island; Shelves - 2-4 kb from CpG island.

Supplementary Table 5.2: Principal components explaining variation in 11 biomarkers of 1-carbon metabolism

Principal Component (PC)	Eigenvalue	Proportion of variance explained	Cumulative proportion of variance explained
PC1	2.70	0.25	0.25
PC2	1.99	0.18	0.43
PC3	1.40	0.13	0.55
PC4	1.06	0.10	0.65
PC5	0.93	0.08	0.74
PC6	0.78	0.07	0.81
PC7	0.73	0.07	0.87
PC8	0.47	0.04	0.92
PC9	0.40	0.04	0.95
PC10	0.33	0.03	0.98
PC11	0.21	0.02	1.00

Chapter 6 Nutritional predictors of plasma homocysteine in three datasets from West Kiang region, The Gambia.

_____

Summary of Chapter

Background: In rural Gambia dry season plasma homocysteine concentrations are higher than in the rainy season amongst women of reproductive age. Previous studies have suggested that children born to women conceiving in the dry season have patterns of DNA methylation suggestive of loss of regular imprinting. Furthermore, maternal homocysteine is inversely associated with DNA methylation at several genomic loci. The hypothesis is that a pre-conception supplement lowering homocysteine will help enable epigenetic processes to function unhindered during fetal development. This chapter identifies potential ingredients for a novel homocysteine-lowering supplement, tailoring the design to the requirements of the population in West Kiang, rural Gambia.

Methods: Linear regression models from three datasets from rural Gambia are used to identify nutritional predictors of homocysteine in plasma. Recognising that predictive models can be unstable in their variable selection, three regression approaches are adopted (two-step approach, backwards stepwise and Lasso) to obtain a qualitative overview of the most consistent patterns of predictors of homocysteine. Predictive equations for homocysteine from each dataset are tested in the other two datasets in order to understand how generalizable the results are.

Results: Four predictors were consistently inversely associated with homocysteine across the datasets and regression approaches: folate, B12, B2 and betaine. Cysteine and dimethylglycine were consistently positively associated with homocysteine. The patterns for choline and B6 were inconsistent, either not being retained as predictors in the models or being retained with conflicting directions of association with homocysteine. Predictive equations for homocysteine performed well cross-dataset (all adjusted $R^2 > 0.39$).

Conclusions: The analyses result in the selection of four nutritional inputs to consider at the supplement design stage: betaine, vitamin B12, vitamin B2 and folate. These findings are generalizable to different cohorts from the same region in The Gambia.

6.1 Introduction

Prior work from the MRC International Nutrition Group is summarised in Chapter 2. In brief, the existing evidence suggests that women conceiving in the dry season have a profile of nutrients related to one-carbon metabolism that correspond to a lower methylation potential compared to those conceiving in the rainy season. In turn, offspring conceived in the dry season display DNA methylation patterns suggestive of loss of imprinting at the *VTRNA2-1* gene[1]. Our group's long-term goal is therefore to design a periconceptional nutritional intervention for women that improves the regulation of the infant epigenome by providing micronutrients in the quantity necessary for optimal one-carbon metabolism all year round. To achieve this we first require a proof-of-concept trial showing that a nutritional supplement can optimise the metabolome in non-pregnant women by correcting nutritional imbalances and increasing the methylation potential by reducing homocysteine. Should the supplement work it could be a promising candidate for future pregnancy trials investigating offspring epigenetic outcomes.

The primary end point in the proof-of-concept trial is to reduce plasma homocysteine (Hcy), since maternal plasma Hcy is strongly inversely associated with the ratio of S-adenosyl methionine (SAM) to S-adensosyl homocysteine (SAH)[2], a measure of methylation potential[3], and inversely associated with infant methylation at ME loci (as shown in Chapter 5 and Dominguez-Salas *et al*. 2014[4]). The hypothesis here is that reducing homocysteine will help improve the methylation potential and enable one-carbon metabolic pathways to function unhindered.

Whilst the ratio of SAM:SAH is the most common and direct measure of methylation potential, there are a number of reasons why investigating plasma Hcy is preferable. Firstly, plasma Hcy is stable in EDTA at room temperature for 8 days and at -25°C for 29 years[5]. This is in contrast to SAM, which rapidly converts into SAH and is strongly affected by freeze-thaw cycles. For example, unpublished data from the University of British Columbia suggests that the concentration of SAH after the second freeze-thaw cycle is 16.7% higher than after the first free-thaw cycle (details in **Annex 4.3**). Investigating Hcy is therefore more reliable, particularly when using stored samples. Secondly, Hcy assays are cheaper and more commonly available than those for SAM and SAH. Thirdly, there is already a wealth of existing evidence to show how nutritional interventions have been successful in reducing homocysteine through targeting one-carbon metabolic pathways[6–8].

In the target population, women of reproductive age of West Kiang district, there were three available datasets of plasma metabolites involved in one-carbon metabolism (MDEG main study, Indicator group, MDEG-2). In this chapter I investigate the nutritional biomarker predictors of homocysteine in these three datasets with the aim of identifying the micronutrients consistently demonstrating inverse associations with homocysteine as candidates to take forwards into the supplement design stage. These analyses therefore provide an opportunity to tailor the supplement more specifically to the population in rural Gambia rather than rely on a more generic formula.

6.2   Methods

The three datasets had different characteristics, for example, varying in which months of the year were captured, pregnant versus non-pregnant participants and different years of data collection. Comparing predictors of homocysteine across these three different datasets was intended to improve the generalisability of results. Full details of the three datasets are presented in Chapter 2 (an overview of each) and Chapter 4 (details on MDEG-2).

6.2.1   *Selection of 1-carbon biomarkers as potential predictors*

I assessed the original list of biomarkers in each dataset for multicollinearity by inspecting the variance inflation factors. Amongst the B6 vitamers I dropped pyridoxic acid and pyridoxal from all datasets since they were highly correlated with pyridoxal-5'-phospahte (PLP), which is the main functional marker of vitamin B6. Since the active B12 data was missing in the MDEG-2 dataset (see Chapter 4) I used plasma total B12 from all datasets instead. As previously mentioned, SAM and SAH from the MDEG-2 dataset had degraded over storage time and/or freeze-thaw processing and were unusable. I dropped them from the list of potential predictors for that reason, and also because I wanted to prioritise the biomarkers that could be considered for a supplement rather than the intermediary metabolites. For the MDEG main study and the Indicator group the final list of potential predictors was therefore: B2, PLP, B12, cysteine, methionine, dimethylglycine (DMG), folate, choline and betaine. MDEG-2 used the same list but also included the amino acid panel.

Cysteine and DMG are the products of metabolic reactions involving homocysteine. Cysteine is formed in the transsulfuration pathway after Hcy is irreversibly degraded, and DMG is formed during the remethylation pathway involving betaine as the methyl donor. As such, they are not so useful for consideration of supplementation design and so I also repeated all analyses excluding these two metabolites. Although methionine is also a product of Hcy

remethylation it sits in a direct circular pathway that regenerates Hcy via SAM (see **Annex 4.4** figure B), and therefore methionine was kept in all models.

6.2.2 *Predictive models: three approaches*

Since regression models are not always stable in their variable selection (see discussion in Chapter 5) I wanted to compare three different methods of linear regression modelling: a two-step approach, backwards stepwise and least angle regression using Lasso ('least absolute shrinkage and selection operator'). The aim was to obtain a qualitative overview in order to identify consistent patterns of predictors of homocysteine in different datasets using different variable selection approaches.

Log Hcy was used as the dependent variable in all the models. All potential nutritional predictors were standardised and log-transformed before entry into the models as independent variables. All effect sizes therefore represent the change in log Hcy for a one SD increase in the log-transformed independent variable. Some *a priori* confounders were forced into all models, and these varied according to the information available for each dataset (age for Indicator group; age, BMI and gestational age for MDEG main study; age, BMI, gestational age and inflammation using α-1-acid glycoprotein (AGP) for MDEG-2). The Indicator group dataset contained repeated measures on the same individual. Both the Breusch-Pagan test (for random effects versus ordinary least squares regression, $p<0.001$) and the Hausman test (for fixed versus random effects, $p=0.24$), confirmed the use of random effects. This longitudinal dataset, however, could only use the two-step approach, which was the only method allowing for the modelling of random effects. The MDEG main study and MDEG-2 datasets, being cross-sectional, could use all three modelling approaches.

In the two-step approach log Hcy was regressed against the nutritional predictors in univariable linear regression. Any variables with $p<0.1$ were then taken forwards to a multivariable linear regression model. In the final results I report any retained independent variables with $p<0.05$. The Wald test was used for the statistical tests in this approach. In the backwards stepwise approach I used automatic selection with $p>0.2$ as the criteria for removal from the model. Again, in the final results I only report the retained independent variables with $p<0.05$ using the Wald test. In the Lasso model the number of variables retained in the final model corresponds to the lowest Cp statistic found. No individual p values are reported and the output contains the coefficients only. Given it was not possible to force confounders into the Lasso models I pre-adjusted log Hcy by regressing it against the

confounders prior to model entry.

I ran the three models using all the data available ('combined season' model), but then also stratified the models by peak of the dry season (February – April) and the peak of the rainy season (July – September). This was because the aim was to implement the supplementation trial in the dry season since, as mentioned previously, this is the time of year associated with lower methylation potential and possible loss of imprinting at *VTRNA 2-1*. Table 6.1 summarises the different datasets used, detailing the list of potential nutritional predictors, the confounders, the type of statistical models used and the sample sizes.

6.2.3   *Generalisability of predictors across datasets*

For supplement design considerations it can be useful to know how generalizable results are from one dataset to another by assessing how well the predictors of homocysteine in one cohort work in other datasets. To do this I generated predictive equations for homocysteine in each of the three datasets and then tested how well these equations worked when applied to the other two datasets. I considered the set of plasma nutritional biomarkers that were common across all three datasets: betaine, choline, DMG, B12, B2, folate, methionine, cysteine and PLP.   To obtain the predictive linear equations I ran a multivariable linear regression with log Hcy as the outcome and the plasma biomarkers all entered as exposures. The constant and the coefficients from this regression model provided the equation for predicted Hcy, and I recorded how well this model predicted actual Hcy within the same dataset. Then I applied the predictive equation to the other datasets to generate the new predicted Hcy variable, and regressed this against actual Hcy to assess its predictive power ('cross-dataset' prediction). In all these cross-dataset predictive models the predictors were all log-transformed and standardised.   All models were adjusted for age. Two additional variables, BMI and gestational age, were considered when comparing MDEG main study and MDEG-2 datasets (the Indicator group did not have this information). All regression models performed using the Indicator group dataset used random effects to account for the longitudinal data design. When describing model fit in the MDEG main study and MDEG-2 datasets the adjusted R^2 is reported.

Supplementary Material 6.1 provides a more detailed example of the methodology I followed to generate the equations for these cross-dataset predictions.

Table 6.1 Summary of datasets and models used

Dataset	Total N	Dry Season N	Rainy Season N	Biomarkers included in model	A priori confounders forced into all models	Models used	Notes
Indicator group	48 women with 288 total observations	34 women with 79 observations	28 women with 63 observations	Hcy, B2, PLP, B12, folate, methionine, choline, betaine, DMG, cysteine	Age	Linear regression with random effects. 2-step approach only	Non-pregnant women, roughly 30 followed for one year, longitudinal study with monthly blood samples[2].
MDEG main study	167	83	84	Hcy, B2, PLP, B12, folate, methionine, choline, betaine, DMG, cysteine	Maternal age, BMI, gestational age	Ordinary linear regression. 2-step, backwards stepwise and Lasso	Cross-sectional data, pregnant women selected from peak of rainy and peak of dry season[4].
MDEG-2	350	87	100	Hcy, B2, PLP, B12, folate, methionine, choline, betaine, DMG, cysteine, Asp, Thr, Ser, Glu, Gly, Ala, Val, Ile, Leu, Tyr, Phe, Lys, His, Arg, Pro	Maternal age, BMI, gestational age, inflammation (AGP)	Ordinary linear regression. 2-step, backwards stepwise and Lasso	Cross sectional data, pregnant women selected from each month of year.

Abbreviations: Hcy, homocysteine; PLP, pyridoxal 5'-phosphate; DMG, dimethylglycine; BMI, body mass index; Asp, aspartic acid; Thr, threonine; Ser, serine; Glu, glutamic acid; Gly, glycine; Ala, alanine; Val, valine; Ile, isoleucine; Leu, leucine; Tyr, tyrosine; Phe, phenylalanine; Lys, lysine; His, histidine; Arg, arginine; Pro, proline.

6.3   Results

Tables 6.2, 6.3 and 6.4 detail the predictors of Hcy retained in final multivariable models in the MDEG main study, MDEG-2 and Indicator group datasets respectively. A summary of the retained coefficients in all models is provided in Figure 6.1.

There were four predictors consistently inversely associated with Hcy in most model approaches and seasons: folate, B2, B12 and betaine. Of these folate had the strongest effect sizes, particularly in the dry season. Cysteine was positively associated with Hcy in all models. DMG was also positively associated with Hcy in the combined and dry season models.

The pattern for choline and PLP was inconsistent. In several models they were not retained as predictors, and where they did feature the direction of the association varied. Choline was inversely associated with Hcy in the MDEG-2 Lasso dry and rainy season models, yet was positively associated in the MDEG main study Lasso models in all seasons. PLP was positively associated with Hcy in the MDEG-2 dataset but showed different directions of association by season in the MDEG main study dataset.

The amino acids were only measured in the MDEG-2 dataset. Glycine was a positive predictor of Hcy in the combined models but was not retained when models were stratified by season. None of the other amino acids showed a consistent pattern, and when they were retained in the final models they had small effect sizes.

When the analyses were repeated with the by-products DMG and cysteine removed from the models (Supplementary Tables 6.1-6.3, Supplementary Figure 6.1) the underlying patterns were largely similar. Folate, B2, B12 and betaine, where they were retained in the models, remained negative predictors of Hcy. The positive association between PLP and Hcy became stronger, but the associations between choline and Hcy became weaker (although still demonstrating both positive and inverse directions). For the amino acids glycine remained a positive predictor of Hcy in the MDEG-2 combined season models, however, this time proline was also identified as a positive predictor and alanine as a negative predictor.

Table 6.5 summarises how well the set of nutritional biomarkers common to all three datasets predicted homocysteine within their own dataset. For example, in the MDEG main study dataset the predicted Hcy using the equation from nine biomarkers plus age explained 57.4% of the variation of actual homocysteine. In the MDEG-2 dataset predicted Hcy using these ten predictors explained 52.5% of the variance of actual Hcy. In both cases adding the

two extra co-variates of BMI and gestational age improved the adjusted R^2 only modestly (<1%). In the Indicator group the predicted Hcy performed better at explaining the variance of actual Hcy within each subject (47.8%) compared to between subjects (35.1%).

Table 6.6 summarises how well the predictive equations worked cross-dataset. The Indicator group equation predicted 54.4% of the variance of Hcy in the MDEG main study, which was only 3% less than the MDEG main study's own set of predictors. It performed almost as well in MDEG-2, explaining just 7.4% less than the original set of predictors. For predicting Hcy in the Indicator group, the MDEG main study and MDEG-2 equations predicted similar levels of within-group and overall variance. They achieved very close to the performance using the Indicator group's own set of predictors (difference of <2%). The cross dataset comparisons between the MDEG main study and MDEG-2 did not perform quite so well, explaining between 14-18.4% less variance than the same-dataset predictions. However, even in the worst performing comparison (MDEG-2 predictive model from 12 variables tested in the MDEG main study) the overall variance explained was 39.6%, which is still considerable. Whilst adding the two extra variables (gestational age and BMI) slightly improved model fit when prediction occurred within the same dataset, in cross-dataset comparisons these two extra variables actually decreased model fit by around 3%.

Table 6.2 Nutritional Biomarker Independent Predictors of Homocysteine in MDEG Main Study Dataset

Seasonal Model	Statistical approach*	N	R^2	Positive predictors[‡] (β, 95%CI, p value)		Negative predictors[‡] (β, 95%CI, p value)	
Combined Seasons	2-step	150	0.582	Cysteine DMG	0.14 (0.10, 0.18), p<0.001 0.05 (0.01, 0.09), p=0.015	Folate B2	-0.12 (-0.16, -0.09), p<0.001 -0.05 (-0.08, -0.01), p=0.009
	Backwards Stepwise	149	0.610	Cysteine	0.16 (0.12, 0.20), p<0.001	Folate B2 B12	-0.14 (-0.17, -0.10), p<0.001 -0.04 (-0.07, 0.00), p=0.029 -0.05 (-0.09, -0.02), p=0.005
	Lasso	152	0.422	Cysteine DMG Choline Methionine	0.14 0.04 0.04 0.02	Folate Betaine B2 B12 PLP	-0.11 -0.08 -0.04 -0.03 -0.01
Dry season (Feb-Apr)	2-step	79	0.575	Cysteine	0.16 (0.10, 0.22), p<0.001	Folate B2	-0.17 (-0.23, -0.11), p<0.001 -0.07 (-0.13, -0.01), p=0.018
	Backwards Stepwise	79	0.542	Cysteine	0.16 (0.10, 0.23), p<0.001	Folate	-0.18 (-0.24, -0.12), p<0.001
	Lasso	80	0.423	Cysteine DMG Choline Methionine	0.13 0.05 0.06 0.03	Folate B2 B12 Betaine PLP	-0.11 -0.06 -0.02 -0.11 -0.02
Rainy season (Jul-Oct)	2-step	77	0.622	Cysteine	0.13 (0.08, 0.17), p<0.001	None	
	Backwards stepwise	70	0.732	Cysteine	0.17 (0.12, 0.21), p<0.001	Folate B12	-0.10 (-0.14, -0.06), p<0.001 -0.05 (-0.09, -0.01), p=0.010
	Lasso	72	0.504	Cysteine Choline DMG PLP	0.14 0.04 0.03 0.03	Folate Betaine B2 B12	-0.09 -0.05 -0.02 -0.04

*Lasso model fitted to variables demonstrating the smallest Cp value. Note no confidence intervals are reported for Lasso models.

[‡]Predictors are adjusted for maternal age, gestational and BMI. Only variables with a final p<0.05 reported.

[‡]β Coefficients represent the change in log homocysteine for a one SD increase in the log-transformed independent variable.

Abbreviations: CI, confidence interval; DMG, dimethylglycine; PLP, pyridoxal 5'-phosphate

Table 6.3 Nutritional Biomarker Independent Predictors of homocysteine in MDEG-2 Dataset

Seasonal Model	Statistical approach*	N	R^2	Positive predictors[ᵻ] (β, 95%CI, p value)		Negative predictors[ᵻ] (β, 95%CI, p value)	
Year-round	2-step	339	0.588	Cysteine	0.21 (0.18, 0.24), p<0.001	Folate	-0.10 (-0.13, -0.08), p<0.001
				Glycine	0.05 (0.02, 0.08), p<0.001	B12	-0.10, (-0.13, -0.08), p<0.001
				DMG	0.03 (0.01, 0.06), p=0.019	Betaine	-0.07 (-0.10, -0.04), p<0.001
	Backwards Stepwise	338	0.610	Cysteine	0.2 (0.18, 0.23), p<0.001	Folate	-0.10 (-0.13, -0.08), p<0.001
				Glycine	0.06 (0.04, 0.09), p<0.001	B12	-0.10 (-0.12, -0.07), p<0.001
				DMG	0.04 (0.01, 0.06), p=0.004	Betaine	-0.07 (-0.10, -0.04), p<0.001
				Proline	0.04 (0.01, 0.07), p=0.009	B2	-0.05 (-0.07, -0.02), p<0.001
				PLP	0.04 (0.00, 0.07), p=0.027		
				Aspartate	0.03 (0.00, 0.06), p=0.046		
	Lasso	342	0.540	Cysteine	0.18	Folate	-0.09
				Glycine	0.06	B12	-0.08
				PLP	0.04	Betaine	-0.06
				DMG	0.03	B2	-0.04
				Proline	0.03	Tyrosine	-0.02
				Glutamate	0.02	Threonine	-0.02
				Aspartate	0.02	Alanine	-0.01
				Isoleucine	0.01	Methionine	-0.01
				Serine	0.01		
Dry season (Feb-Apr)	2-step	82	0.470	Cysteine	0.18 (0.13, 0.23), p<0.001	B12	-0.08 (-0.13, -0.03), p=0.001
						Betaine	-0.07 (-0.12, -0.03), p=0.003
	Backwards Stepwise	81	0.624	Cysteine	0.18 (0.13, 0.23), p<0.001	Betaine	-0.08 (-0.13, -0.04), p<0.001
				PLP	0.05 (0.01, 0.10), p=0.021	Folate	-0.08 (-0.13, -0.02), p=0.008
				Aspartate	0.05 (0.00, 0.10), p=0.042	B12	-0.07 (-0.12, -0.03), p=0.003
						B2	-0.05 (-0.10, -0.01), p=0.013
						Threonine	-0.04 (-0.08, 0.00), p=0.045
	Lasso	85	0.572	Cysteine	0.16	Betaine	-0.06
				PLP	0.07	Folate	-0.06
				Proline	0.04	B12	-0.06
				Serine	0.03	Choline	-0.04

134

Seasonal Model	Statistical approach*	N	R^2	Positive predictors[‡] (β, 95%CI, p value)		Negative predictors[‡] (β, 95%CI, p value)	
				Glycine	0.02	B2	-0.03
				Aspartate	0.02	Threonine	-0.03
				DMG	0.01		
	2-step	98	0.636	Cysteine	0.23 (0.17, 0.29), p<0.001	B12	-0.09 (-0.14, -0.04), p<0.001
						Folate	-0.08 (-0.12, -0.03), p=0.001
						B2	-0.06 (-0.11, -0.01), p=0.022
	Backwards stepwise	98	0.700	Cysteine	0.24 (0.18, 0.29), p<0.001	Leucine	-0.10 (-0.19, -0.01), p=0.035
				Valine	0.11 (0.03, 0.19), p=0.005	B12	-0.08, (-0.12, -0.04), p<0.001
Rainy season				Proline	0.07 (0.01, 0.13), p=0.028	Folate	-0.07 (-0.11, -0.03), p=0.002
				Aspartate	0.04, p=0.022	Alanine	-0.07 (-0.12, -0.01), p=0.028
(Jul-Oct)						B2	-0.05 (-0.10, 0.00), p=0.036
	Lasso	98	0.529	Cysteine	0.18	B12	-0.06
				Glutamate	0.03	Folate	-0.06
				Proline	0.02	B2	-0.04
				Aspartate	0.01	Betaine	-0.02
				Valine	0.01	Tyrosine	-0.02
						Histidine	-0.01
						Choline	-0.01

*Lasso model fitted to variables demonstrating the smallest Cp value. Note no confidence intervals are reported for Lasso models.
[‡]Predictors are adjusted for maternal age, BMI, gestational age and inflammation (AGP). Only variables with a final p<0.05 reported.
[‡]β Coefficients represent the change in log homocysteine for a one SD increase in the log-transformed independent variable.
Abbreviations: AGP, Alpha-1-acid glycoprotein; CI, confidence interval; DMG, dimethylglycine; PLP, pyridoxal 5'-phosphate

Table 6.4 Nutritional Biomarker Independent Predictors of homocysteine in the Indicator group dataset

Seasonal Model	No. observations	No. women	Overall R2*	Positive predictors[‡] (β[‡], 95% CI, p value)		Negative predictors[‡] (β[‡], 95% CI, p value)	
Combined seasons	288	48	0.306	Cysteine	0.14 (0.11, 0.17), p<0.001	Folate	-0.11 (-0.13, -0.08), p<0.001
				DMG	0.08 (0.04, 0.12), p<0.001	Betaine	-0.05 (-0.08, -0.02), p=0.003
Dry (Feb-Apr)	79	34	0.295	Cysteine	0.17 (0.12, 0.21), p<0.001	Folate	-0.15 (-0.22, -0.08), p<0.001
Rainy (Jul-Oct)	63	28	0.549	Cysteine	0.16 (0.11, 0.22), p<0.001	Folate	-0.13 (-0.18, -0.07), p<0.001
						B12	-0.09 (-0.16, -0.02), p=0.009
						B2	-0.09 (-0.17, -0.01), p=0.026

*Two-step linear regression using random effects

[‡] Predictors are those variables retained in the multivariable model with p<0.05, adjusted for age.

[‡]β Coefficients represent the change in log homocysteine for a one SD increase in the log-transformed independent variable.

Abbreviations: CI, confidence interval; DMG, dimethylglycine

Figure 6.1 Summary of retained coefficients of multivariable linear regression models in three datasets, by season*.

Biomarker	MDEG2 2-step	MDEG2 Stepwise	MDEG2 LASSO	MDEG1 2-step	MDEG1 Stepwise	MDEG1 LASSO	Indicator 2-step
				Combined seasons			
Folate	-0.1	-0.1	-0.09	-0.12	-0.14	-0.11	-0.11
B2		-0.05	-0.04	-0.05	-0.04	-0.04	
PLP		0.04	0.04			-0.01	
B12	-0.1	-0.1	-0.08		-0.05	-0.03	
Betaine	-0.07	-0.07	-0.06			-0.08	-0.05
Choline						0.04	
Methionine			-0.01			0.02	
Cysteine	0.21	0.2	0.18	0.14	0.16	0.14	0.14
DMG	0.03	0.04	0.03	0.05		0.04	0.08
Glycine	0.05	0.06	0.06				
Glutamate			0.02				
Proline		0.04	0.03				
Aspartate		0.03	0.02				
Tyrosine			-0.02				
Alanine			-0.01				
Isoleucine			0.01				
Serine			0.01				
Threonine			-0.02				

Biomarker	MDEG2 2-step	MDEG2 Stepwise	MDEG2 LASSO	MDEG1 2-step	MDEG1 Stepwise	MDEG1 LASSO	Indicator 2-step
				Dry season			
Folate		-0.08	-0.06	-0.17	-0.18	-0.11	-0.15
B2		-0.05	-0.03	-0.07		-0.06	
PLP		0.05	0.07			-0.02	
B12	-0.08	-0.07	-0.06			-0.02	
Betaine	-0.07	-0.08	-0.06			-0.11	
Choline			-0.04			0.06	
Methionine						0.03	
Cysteine	0.18	0.18	0.16	0.16	0.16	0.13	0.17
DMG			0.01			0.05	
Glycine			0.02				
Glutamate							
Proline			0.04				
Aspartate		0.05	0.02				
Tyrosine							
Alanine							
Isoleucine							
Serine			0.03				
Threonine		-0.04	-0.03				

Biomarker	MDEG2 2-step	MDEG2 Stepwise	MDEG2 LASSO	MDEG1 2-step	MDEG1 Stepwise	MDEG1 LASSO	Indicator 2-step
				Rainy season			
Folate	-0.08	-0.07	-0.06		-0.1	-0.09	-0.13
B2	-0.06	-0.05	-0.04			-0.02	-0.09
PLP						0.03	
B12	-0.09	-0.08	-0.06		-0.05	-0.04	-0.09
Betaine			-0.02			-0.05	
Choline			-0.01			0.04	
Methionine						0.02	
Cysteine	0.23	0.24	0.18	0.13	0.17	0.14	0.16
DMG						0.03	
Glycine							
Glutamate			0.03				
Proline		0.07	0.02				
Aspartate		0.04	0.01				
Tyrosine			-0.02				
Alanine		-0.07					
Histidine			-0.01				
Valine		0.11	0.01				
Leucine		-0.1					

*Note: Coefficients represent change in log homocysteine for each SD increase in the log-transformed independent variable. Only variables with p<0.05 shown. Those inversely associated with homocysteine are shaded in green, and those positively associated are shaded in red.

Abbreviations: DMG, dimethylglycine; MDEG1, MDEG main study

137

Table 6.5 Performance of predictors of homocysteine within their own dataset

Dataset	Variables used	N	Root MSE	R²	Adj R²	p
				Regress actual against predicted*		
MDEG main study	Bet, chol, DMG, B12, fol, met, cys, B2, PLP, age (10)	152	0.203	0.602	0.574	<0.001
MDEG main study	Bet, chol, DMG, B12, fol, met, cys, B2, PLP, age, BMI, gestational age (12)	149	0.202	0.614	0.580	<0.001
Indicator group	Bet, chol, DMG, B12, fol, met, cys, B2, PLP, age (10)	288 Obs, 48 women		0.478 (within) 0.351 (between) 0.359 (overall)		<0.001
MDEG-2	Bet, chol, DMG, B12, fol, met, cys, B2, PLP, age (10)	346	0.238	0.539	0.525	<0.001
MDEG-2	Bet, chol, DMG, B12, fol, met, cys, B2, PLP, age, BMI, gestational age (12)	346	0.236	0.547	0.531	<0.001

*Results from a univariable linear regression with actual log homocysteine as the dependent variable and predicted log homocysteine as the independent variable.

Abbreviations: Bet, betaine; chol, choline, DMG, dimethylglycine; fol, folate; met, methionine; cys, cysteine, PLP, pyridoxal 5'-phosphate; MSE, mean square error; Adj, adjusted; Obs, observations.

Table 6.6 Performance predictors of homocysteine cross-dataset

Original Dataset	Test dataset	Variables used	N	Root MSE	R^2	Adj R^2	p	Difference in R^2 compared to same dataset prediction (Table 6.5)
					Regress actual against predicted*			
MDEG main study	Indicator	Bet, chol, DMG, B12, fol, met, cys, B2, PLP age (10)	288 Obs, 48 women		0.421 (within) 0.414 (between) 0.387 (overall)		<0.001	-0.057 (within) 0.063 (between) 0.028 (overall)
MDEG main study	MDEG-2	Bet, chol, DMG, B12, fol, met, cys, B2, PLP age (10)	346	0.263	0.418	0.417	<0.001	-0.121
MDEG main study	MDEG-2	Bet, chol, DMG, B12, fol, met, cys, B2, PLP age, BMI, gestational age (12)	346	0.269	0.383	0.391	<0.001	-0.164
Indicator group	MDEG main study	Bet, chol, DMG, B12, fol, met, cys, B2, PLP age (10)	152	0.210	0.546	0.544	<0.001	-0.056
Indicator group	MDEG-2	Bet, chol, DMG, B12, fol, met, cys, B2, PLP age (10)	346	0.256	0.452	0.451	<0.001	-0.087
MDEG-2	Indicator group	Bet, chol, DMG, B12, fol, met, cys, B2, PLP age (10)	288 Obs 48 women		0.401 (within) 0.305 (between) 0.350 (overall)		<0.001	-0.077 (within) -0.046 (between) -0.009 (overall)
MDEG-2	MDEG main study	Bet, chol, DMG, B12, fol, met, cys, B2, PLP age (10)	152	0.236	0.431	0.427	<0.001	-0.171
MDEG-2	MDEG main study	Bet, chol, DMG, B12, fol, met, cys, B2, PLP age, BMI, gestational age (12)	149	0.243	0.400	0.396	<0.001	-0.214

*Results from a univariable linear regression with actual log homocysteine as the dependent variable and predicted log homocysteine as the independent variable.

Abbreviations: Bet, betaine; chol, choline, DMG, dimethylglycine; fol, folate; met, methionine; cys, cysteine, PLP, pyridoxal 5'-phosphate; MSE, mean square error; Adj, adjusted; Obs, observations.

6.4   Discussion

In order to assess which nutritional components could be considered for a supplement to reduce plasma homocysteine I analysed three different datasets using three different linear regression approaches to see which nutritional biomarkers were consistently inversely associated with Hcy. This approach built up an overview of the most reliable predictors of homocysteine. Folate, B12, B2 and betaine were the most consistent negative predictors of Hcy. Predictive equations for Hcy generated in one dataset worked well in the other datasets, indicating that results from these models could be generalizable to different cohorts from the same region in The Gambia.

6.4.1   Negative predictors of homocysteine

The findings that plasma folate, B12, B2 and betaine were inversely associated with Hcy are consistent with the literature[10–13]. Supplementation with these components have been successful in previous trials with the aim of reducing homocysteine[6,14–17].

Hcy can be metabolised via a process involving remethylation or through the transsulfuration pathway. In the former, Hcy accepts a methyl group to form methionine[18], which can then in turn be condensed with ATP to form S-adenosyl methionine, a methyl donor involved in numerous transmethylation reactions. The remethylation of Hcy to methionine uses two distinct pathways. The major one is the vitamin B12- dependent reaction involving folate metabolic pathways[19], chiefly the donation of a methyl group from N^5-methyl tetrahydrofolate ('methyl-THF'). The alternative pathway for the methylation of Hcy, predominantly used in the liver and kidneys, uses the methyl group from betaine, a product formed through the oxidation of choline[20,21]. In the transsulfuration pathway Hcy is metabolised through its irreversible degradation to cystathionine and cysteine, requiring PLP (vitamin B6)[19]. Annex 4.4 provides more details on the context of these pathways.

Dietary folates and folic acid therefore contribute to the removal of homocysteine via methyl-THF. They are first reduced to form tetrahydrofolate (THF), which is in turn is reduced to methylene-THF, then to methyl-THF. This is the form that donates its methyl group to Hcy using vitamin B12. In human plasma and the cytosol the predominant form of vitamin B12 is methylcobalamin[22]. This is the form that is used as a coenzyme for methionine synthase, the enzyme responsible for adding the methyl group from 5-methyl-THF to Hcy to form methionine. The cobalt atom in the methylcobalamin cycles between different oxidation states in the reaction catalysed by methionine synthase. Firstly cob(I)alamin accepts the

methyl group from 5-methyl-THF to form methylcob(I)alamin and THF. Methylcob(I)alamin then transfers the methyl group to homocysteine to form methionine, and generating cob(I)alamin again. Every 100-2000 cycles the cob(I)alamin is oxidised to cob(II)alamin (it loses an electron)[23,24]. This is oxidation state is catalytically inactive and the compound needs to be reduced back to the +1 oxidation state. In humans methionine synthase reductase catalyses the reaction whereby the cob(II)abalamin is firstly reduced to cob(I)alamin and then SAM provides a methyl donor to regenerate methylcob(I)alamin[23,24].

The above metabolic descriptions explain why folate, betaine and B12 are required to recycle Hcy. In order to understand where B2 fits in we need to re-visit the step where dietary folates are converted to methylene-THF. B2 is required as a precursor to flavin mononucleotide (FMN), a reaction catalysed by riboflavin kinase. FMN is converted to flavin adenine dinucleotide (FAD) by FAD synthase[25,26]. FAD is a cofactor required by methylenetetrahydrofolate reductase (MTHFR) to reduce methylene-THF to methyl-THF[3].

Threonine was retained as a negative predictor of Hcy in some models, as was alanine in the analyses excluding cysteine and DMG. The relationship between plasma Hcy and amino acids is not well documented in the literature. However, in a study investigating plasma metabolites in healthy controls no association between Hcy and amino acids was found[27]. Without further justification from the literature it is not possible to say whether these are genuine or chance findings of associations, especially as amino acids were only measured in one of the datasets.

6.4.2   *Positive predictors of homocysteine*

Cysteine was the strongest positive predictor of Hcy, which was to be expected given that Hcy can be converted into cystathionine by Cystathionine β-synthase (CBS) and then onto cysteine by Cystathionine γ-lyase[19]. The positive association of DMG with homocysteine can be explained by the reaction catalysed by betaine-homocysteine methyltransferase (BHMT), which methylates Hcy using a methyl group from betaine and in the process forms DMG[28]. Thus describing cysteine and DMG as 'positive predictors' can be misleading since they are by-products of Hcy catabolism, and this is the reason the analyses were repeated without these two metabolites.

There were no strong associations between methionine and Hcy, most likely explained by its position on a circular metabolic pathway and therefore it can be viewed as both a substrate for Hcy production and as a by-product of Hcy catabolism. It can be seen as a substrate for

Hcy production since it is condensed with ATP to form SAM, which is then demethylated to SAH, which in turn is hydrolysed to form Hcy and adenosine[19,29]. Methionine can also be viewed as the product resulting from Hcy catabolism through Hcy remethylation using the methyl group from 5-methyl-THF or betaine. Linear regression models are not designed to model such circular pathways and therefore it is perhaps not surprising that no clear associations between methionine and Hcy can be identified. We might only expect homocysteine to significantly rise in response to a methionine load (e.g. straight after a protein-rich meal), whereas the samples in these datasets are fasted blood samples.

Given that choline is oxidised to betaine[30], and that in the datasets betaine is strongly inversely associated with Hcy, an inverse association between choline and homocysteine was expected. Whilst this was the case in the MDEG-2 dataset, in the MDEG main study choline had a positive association with Hcy. Despite seeming counter intuitive, this positive association has been found previously in pregnant women[10]. Molloy *et al.* (2005) hypothesise that increased fetal requirements for choline during pregnancy upregulates choline production via phosphatidylcholine (PC) synthesis[10]. To produce one molecule of PC using the PEMT pathway three methyl groups from SAM are required. Once SAM donates its methyl group it forms SAH, which is hydrolysed to Hcy[31], thereby increasing Hcy concentrations. It is therefore unlikely that choline is associated with increased Hcy in a causal relationship *per se,* but that the association is due to the physiological adaptation in pregnancy to increased endogenous production of choline. This remains speculative since we cannot distinguish between dietary and endogenous sources of choline in the plasma samples, nor was the direction of the association between choline and Hcy consistent within the datasets.

Vitamin B6 (in the active form of PLP) is required to reduce THF to methylene-THF, and is also a cofactor in the transsulfuration pathway converting homocysteine to cysteine[19]. It was therefore expected to be a negative predictor, considering its role in pathways responsible for Hcy's remethylation and transsulfuration. Indeed, B6 has been inversely associated with Hcy in previous studies[12,32,33]. However, there are also several studies showing that PLP has no effect on fasting Hcy levels[11,34,35]. It is more likely that the effect of PLP will be seen in reducing Hcy after a methionine load[19]. It is biologically very unlikely that PLP would cause an increase in plasma Hcy, and its retention in some of the models as a positive predictor is not easily explained.

The positive association between glycine and Hcy could be explained by the reaction

catalysed by glycine N-methyltransferase (GNMT). In this reaction glycine accepts a methyl group from SAM to form SAH and sarcosine. This reaction occurs especially when SAM concentrations are high[17]. The production of SAH from this reaction would therefore contribute to higher Hcy through the hydrolysis of SAH to Hcy. Surprisingly, serine was not retained in many models at all, and where it was retained the effect size was very small. Given that serine is used in the transsulfuration pathway to catabolise Hcy, and also used as a 1-carbon donor at the stage of converting THF to methylene-THF through the action of serine hydroxymethyltransferase and PLP[36], it may have been expected to have an inverse association with Hcy.

6.4.3   Predictive power and generalisability of predictors

In the two cross-sectional datasets (MDEG main study and MDEG-2) the set of predictors explained well over half the variability of plasma Hcy, with the predictors working slightly less well in the longitudinal dataset (Indicator group). However, the predictors were still able to explain over a third of the variability of Hcy in the Indicator group. The predictors also performed well in cross-dataset comparisons. This is particularly encouraging given that the predictive model from the MDEG main study dataset, derived from only 6 months of data, was still able to predict Hcy in datasets including year-round data (MDEG-2 and the Indicator group). Taken together these results affirmed the notion that nutritional predictors of Hcy are generalizable between different cohorts from West Kiang region, and can therefore be taken forwards as fairly robust predictors of Hcy when it comes to supplement design considerations.

The unexplained variability in plasma Hcy could perhaps be partially explained by the effect of SNPs in enzymes involved in one-carbon metabolism (e.g. MTHFR, CBS) on homocysteine levels. The notion that relevant SNPs influence plasma Hcy levels has been well documented in other populations[37–43]. However, the effect of these genetic variants on total plasma Hcy variability is likely to be less significant than the nutritional biomarker predictors, with the former often explaining less than 6% of total Hcy variability[37,42]. There are also several other potential predictors of Hcy which were not measured (e.g. fat-free mass, renal function, estradiol and creatinine concentrations, caffeine consumption, smoking, alcohol intake etc.[12,35]). The overall aim of these analyses was to identify nutritional inputs that could be modified through a supplement, rather than an exercise in identifying all potential predictors of Hcy. The high proportion of Hcy variability explained by the profile of nutritional predictors provided confidence that a nutritional approach to modifying Hcy should work well in the

target population.

6.4.4 Limitations

All linear regression approaches have their own set of limitations in the way they select and report the effect of a sub-set of explanatory variables. These limitations have been explored in Chapter 5 and elsewhere in detail[44–47]. The potential differences that the type of linear regression approach can make to variable selection was the main reason that three different approaches were compared. In a traditional research study on predictors of homocysteine this would clearly be open to criticisms of multiple testing and an incoherent statistical approach. However, for a supplement design aim it was useful to assess to what extent similar trends emerged across multiple datasets and linear regression approaches that could help identify consistent negative predictors of Hcy.

The datasets carry the limitations of utilising cross-sectional plasma information. Whilst useful for identifying predictors of Hcy they cannot go much further in guiding the doses of ingredients to include in a supplement, nor what the expected effect of a supplement might be on plasma Hcy. To achieve this would require the integration of dietary intake data and kinetic data to be used in more formal pharmacodynamics approaches, which go beyond the scope of this PhD. A plasma sample taken at one point in time can provide clues as to what some of the underlying determinants of plasma Hcy are, but without further clinical tests it can be difficult to disentangle the components. For example, a high Hcy concentration could be due to increased transmethylation, decreased transsulfuration, decreased remethylation, decreased uptake of Hcy by the kidney, genetic defects in key enzymes, or a combination of these factors[19,29,48]. Thus the utility of these analyses is constrained to selection of potential ingredients to take forwards to the supplement design stage. The outputs of this analysis may therefore be seen as more qualitative in nature rather than quantitative, but they still help to design a supplement that is better tailored to the West Kiang population rather than rely solely on a generic, commercially available supplement.

6.5   Conclusion

In the context of West Kiang, the four nutritional biomarkers consistently inversely associated with plasma Hcy were betaine, B12, B2 and folate. These findings were consistent with the literature. Vitamin B6 and choline, contrary to expectations, were positively associated with Hcy and therefore dropped from the list of potential supplement ingredients. Betaine, B12, B2 and folate were taken forwards into the supplement design stage, which is explored in more detail in the following chapter.

6.6   References

1. Silver MJ, Kessler NJ, Hennig BJ, et al. (2015) Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. *Genome Biol.* **16**, 118.

2. Dominguez-Salas P, Moore SE, Cole D, et al. (2013) DNA methylation potential: dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women. *Am. J. Clin. Nutr.* **97**, 1217–27.

3. Mason JB (2003) Biomarkers of Nutrient Exposure and Status in One-Carbon (Methyl) Metabolism. *J. Nutr.* **133**, 941S–947.

4. Dominguez-Salas P, Moore SE, Baker MS, et al. (2014) Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. *Nat. Commun.* **5**, 3746.

5. Hustad S, Eussen S, Midttun Ø, et al. (2012) Kinetic modeling of storage effects on biomarkers related to B vitamin status and one-carbon metabolism. *Clin. Chem.* **58**, 402–10.

6. Homocysteine Lowering Trialists' Collaboration. (1998) Lowering blood homocysteine with folic acid based supplements: meta-analysis of randomised trials. *BMJ* **316**, 894–8.

7. Huang T, Zheng J, Chen Y, et al. (2011) High consumption of Ω-3 polyunsaturated fatty acids decrease plasma homocysteine: a meta-analysis of randomized, placebo-controlled trials. *Nutrition* **27**, 863–7.

8. McRae MP (2013) Betaine supplementation decreases plasma homocysteine in healthy adult participants: a meta-analysis. *J. Chiropr. Med.* **12**, 20–25.

9. Moore SE, Fulford AJ, Darboe MK, et al. (2012) A randomized trial to investigate the effects of pre-natal and infant nutritional supplementation on infant immune development in rural Gambia: the ENID trial: Early Nutrition and Immune Development. *BMC Pregnancy Childbirth* **12**, 107.

10. Molloy AM, Mills JL, Cox C, et al. (2005) Choline and homocysteine interrelations in umbilical cord and maternal plasma at delivery. *Am. J. Clin. Nutr.* **82**, 836–42.

11. Holm PI, Ueland PM, Vollset SE, et al. (2005) Betaine and folate status as cooperative determinants of plasma homocysteine in humans. *Arterioscler. Thromb. Vasc. Biol.* **25**, 379–85.

12. Jacques PF, Bostom AG, Wilson PW, et al. (2001) Determinants of plasma total homocysteine concentration in the Framingham Offspring cohort. *Am. J. Clin. Nutr.* **73**, 613–621.

13. van Beynum IM, den Heijer M, Thomas CM, et al. (2005) Total homocysteine and its predictors in Dutch children. *Am. J. Clin. Nutr.* **81**, 1110–1116.

14. Steenge GR, Verhoef P & Katan MB (2003) Betaine supplementation lowers plasma homocysteine in healthy men and women. *J. Nutr.* **133**, 1291–5.

15. Olthof MR, van Vliet T, Verhoef P, et al. (2005) Effect of homocysteine-lowering nutrients on blood lipids: results from four randomised, placebo-controlled studies in healthy humans. *PLoS Med.* **2**, e135

16. Dawson SL, Bowe SJ & Crowe TC (2016) A combination of omega-3 fatty acids, folic acid and B-group vitamins is superior at lowering homocysteine than omega-3 alone: A meta-analysis. *Nutr. Res.* **36**, 499–508.

17. Olthof MR & Verhoef P (2005) Effects of betaine intake on plasma homocysteine concentrations and consequences for health. *Curr. Drug Metab.* **6**, 15–22.

18. Robinson JL & Bertolo RF (2016) The Pediatric Methionine Requirement Should Incorporate Remethylation Potential and Transmethylation Demands. *Adv. Nutr.* **7**, 523–34.

19. Selhub J (1999) Homocysteine metabolism. *Annu. Rev. Nutr.* **19**, 217–46.

20. Velzing-Aarts F V, Holm PI, Fokkema MR, et al. (2005) Plasma choline and betaine and their relation to plasma homocysteine in normal pregnancy. *Am. J. Clin. Nutr.* **81**, 1383–9.

21. McKeever MP, Weir DG, Molloy A, et al. (1991) Betaine-homocysteine methyltransferase: organ distribution in man, pig and rat and subcellular distribution in the rat. *Clin. Sci.* **81**, 551–6.

22. Green R & Miller JW (2007) Vitamin B12. In *Handbook of Vitamins*, 4th ed. [ZempleniJ, Rucker R, McCormick D, et al., editors]. Boca Ralton, FL: Taylor & Francis Group.

23. Ragsdale SW (2008) Catalysis of methyl group transfers involving tetrahydrofolate and B(12). *Vitam. Horm.* **79**, 293–324.

24. Banerjee R V & Matthews RG (1990) Cobalamin-dependent methionine synthase. *FASEB J.* **4**, 1450–9.

25. Hustad S, McKinley MC, McNulty H, et al. (2002) Riboflavin, Flavin Mononucleotide, and Flavin Adenine Dinucleotide in Human Plasma and Erythrocytes at Baseline and after Low-Dose Riboflavin Supplementation. *Clin. Chem.* **48**, 1571–1577.

26. Hustad S, Schneede J & Ueland PM (2013) Madame Curie Bioscience Database

[Internet]. In *Madame Curie Biosci. Database [Internet].* Austin (TX): Landes Bioscience.

27.   Hong SY, Yang DH & Chang SK (1998) The relationship between plasma homocysteine and amino acid concentrations in patients with end-stage renal disease. *J. Ren. Nutr.* **8**, 34–9.

28.   Ueland PM, Holm PI & Hustad S (2005) Betaine: a key modulator of one-carbon metabolism and homocysteine status. *Clin. Chem. Lab. Med.* **43**, 1069–75.

29.   Stipanuk MH (2004) Sulfur amino acid metabolism: pathways for production and removal of homocysteine and cysteine. *Annu. Rev. Nutr.* **24**, 539–77.

30.   Ueland PM (2011) Choline and betaine in health and disease. *J. Inherit. Metab. Dis.* **34**, 3–15.

31.   Pajares MA & Pérez-Sala D (2006) Betaine homocysteine S-methyltransferase: just a regulator of homocysteine metabolism? *Cell. Mol. Life Sci.* **63**, 2792–803.

32.   Selhub J, Jacques PF, Wilson PW, et al. (1993) Vitamin status and intake as primary determinants of homocysteinemia in an elderly population. *JAMA* **270**, 2693–8.

33.   Siri PW, Verhoef P & Kok FJ (1998) Vitamins B6, B12, and folate: association with plasma total homocysteine and risk of coronary atherosclerosis. *J. Am. Coll. Nutr.* **17**, 435–41.

34.   Andersson A, Brattström L, Israelsson B, et al. (1992) Plasma homocysteine before and after methionine loading with regard to age, gender, and menopausal status. *Eur. J. Clin. Invest.* **22**, 79–87.

35.   Dierkes J, Jeckel A, Ambrosch A, et al. (2001) Factors explaining the difference of total homocysteine between men and women in the European Investigation Into Cancer and Nutrition Potsdam study. *Metabolism.* **50**, 640–5.

36.   Davis SR, Stacpoole PW, Williamson J, et al. (2004) Tracer-derived total and folate-dependent homocysteine remethylation and synthesis rates in humans indicate that serine is the main one-carbon donor. *Am. J. Physiol. Endocrinol. Metab.* **286**, E272-9.

37.   van Meurs JB, Pare G, Schwartz SM, et al. (2013) Common genetic loci influencing plasma homocysteine concentrations and their effect on risk of coronary artery disease. *Am. J. Clin. Nutr.* **98**, 668–676.

38.   Cotlarciuc I, Malik R, Holliday EG, et al. (2014) Effect of genetic variants associated with plasma homocysteine levels on stroke risk. *Stroke* **45**, 1920–4.

39.   Zhu Y, Zhu R-X, He Z-Y, et al. (2015) Association of MTHFR C677T with total homocysteine plasma levels and susceptibility to Parkinson's disease: a meta-analysis. *Neurol. Sci.* **36**, 945–951.

40.   Nienaber-Rousseau C, Ellis SM, Moss SJ, et al. (2013) Gene–environment and gene–gene interactions of specific MTHFR, MTR and CBS gene variants in relation to homocysteine in black South Africans. *Gene* **530**, 113–118.

41.   STEFANO V, DEKOU V, NICAUD V, et al. (1998) Linkage disequilibrium at the cystathionine beta synthase (CBS) locus and the association between genetic variation at the CBS locus and plasma levels of homocysteine. *Ann. Hum. Genet.* **62**, 481–490.

42.   Dekou V, Gudnason V, Hawe E, et al. (2001) Gene-environment and gene-gene interaction in the determination of plasma homocysteine levels in healthy middle-aged men. *Thromb. Haemost.* **85**, 67–74.

43.   Williams SR, Yang Q, Chen F, et al. (2014) Genome-Wide Meta-Analysis of Homocysteine and Methionine Metabolism Identifies Five One Carbon Metabolism Loci and a Novel Association of ALDH1L1 with Ischemic Stroke. *PLoS Genet.* **10**, e1004214

44.   George EI (2000) The Variable Selection Problem. *J. Am. Stat. Assoc.* **95**, 1304–1308.

45.   Greenland S (1989) Modeling and variable selection in epidemiologic analysis. *Am. J. Public Health* **79**, 340–9.

46.   Walter S & Tiemeier H (2009) Variable selection: current practice in epidemiological studies. *Eur. J. Epidemiol.* **24**, 733–736.

47.   Morozova O, Levina O, Uusküla A, et al. (2015) Comparison of subset selection methods in linear regression in the context of health-related quality of life and substance abuse in Russia. *BMC Med. Res. Methodol.* **15**, 71.

48.   Hoffer LJ (2004) Homocysteine remethylation and trans-sulfuration. *Metabolism.* **53**, 1480–3.

Supplementary Material 6.1: Generating the cross-dataset predictive models for homocysteine

Example using the MDEG main study dataset:

Step 1: Perform a linear regression with log-transformed homocysteine as the outcome and all predictor variables and *a priori* confounders as explanatory variables. Note all explanatory variables are log-transformed and standardised prior to model entry.

Log Homocysteine	Coef.	Std. Err.	t	P>t	[95% Conf.	Interval]
Betaine	-.0225277	.0273018	-0.83	0.411	-.0765188	.0314633
Choline	.0270422	.0204959	1.32	0.189	-.0134897	.0675741
Dimethylglycine	.0428009	.0202828	2.11	0.037	.0026903	.0829115
B12	-.049664	.018733	-2.65	0.009	-.0867097	-.0126183
Folate	-.1267153	.0194482	-6.52	0.000	-.1651752	-.0882553
Methionine	.0142125	.0198103	0.72	0.474	-.0249635	.0533885
Cysteine	.1642254	.0206111	7.97	0.000	.1234657	.2049851
B2	-.0391662	.017476	-2.24	0.027	-.073726	-.0046064
Pyridoxal 5'-phophate	-.0154349	.0195357	-0.79	0.431	-.0540679	.023198
Age	.056085	.0187471	2.99	0.003	.0190113	.0931586
Body Mass Index	-.014398	.0182335	-0.79	0.431	-.0504558	.0216599
Gestational Age	-.039648	.0255307	-1.55	0.123	-.0901366	.0108405
Constant	1.928049	.0166372	115.89	0.000	1.895147	1.96095

Step 2: Obtain the predictive equation from the constant and coefficients from step 1

Predicted log Hcy = 1.928049 -0.0225277*betaine + 0.0270422*choline + 0.0428009*DMG -0.049664 *B12 - 0.1267153*folate + 0.0142125*methionine + 0.1642254*cysteine - 0.0391662*B2 - 0.0154349*PLP + 0.056085*age - 0.014398*BMI - 0.039648*gestational age

Step 3: In a new dataset (e.g. MDEG-2) generate the predicted homocysteine values by applying the equation from step 2 to the log transformed, standardised explanatory variables.

Step 4: Regress the actual homocysteine in the new dataset against the predicted homocysteine to assess how well the equation performs.

S1 Table 6.1: Nutritional Biomarker Independent Predictors of Homocysteine without Cysteine & Dimethylglycine in MDEG Main Study

Seasonal Model	Statistical approach*	N	R^2	Positive predictors[‡] (β, 95%CI, p value)			Negative predictors[‡] (β, 95%CI, p value)	
Combined Seasons	2-step	150	0.386	None			Folate	-0.09 (-0.13, -0.05), p<0.001
							B2	-0.06 (-0.10, -0.02), p=0.006
	Backwards Stepwise	149	0.391	None			Folate	-0.10 (-0.15, -0.06), p<0.001
							B2	-0.06 (-0.10, -0.02), p=0.007
	Lasso	149	0.161	Methionine	0.01		Folate	-0.07
				Choline	0.01		B2	-0.04
Dry season (Feb-Apr)	2-step	79	0.397	None			Folate	-0.16, (-0.23, -0.09), p<0.001
	Backwards Stepwise	79	0.414	None			Folate	-0.16 (-0.23, -0.09), p<0.001
							B2	-0.07 (-0.14, -0.01), p=0.032
	Lasso	79	0.216	None			Folate	-0.1
							B2	-0.03
Rainy season (Jul-Oct)	2-step	77	0.439	None			None	
	Backwards stepwise	70	0.481	None			None	
	Lasso	70	0.085	Choline	0.03		Folate	-0.02
							B2	-0.01

*Lasso model fitted to variables demonstrating the smallest Cp value. Note no confidence intervals are reported for Lasso models.

[‡]Predictors are adjusted for maternal age, gestational and BMI. Only variables with a final p<0.05 reported.

[‡]β Coefficients represent the change in log homocysteine for a one SD increase in the log-transformed independent variable.

Abbreviations: CI, confidence interval

S1 Table 6.2: Nutritional Biomarker Independent Predictors of Homocysteine without Cysteine & Dimethyl glycine in MDEG-2 Study

Seasonal Model	Statistical approach*	N	R^2	Positive predictors[†] (β, 95%CI, p value)		Negative predictors[†] (β, 95%CI, p value)	
Year-round	2-step	341	0.309	PLP	0.09 (0.05, 0.13), p<0.001	Folate	-0.1 (-0.14, -0.07), p<0.001
				Glycine	0.06 (0.03, 0.10), p<0.001	Betaine	-0.07 (-0.11, -0.04), p<0.001
						B12	-0.06 (-0.10, -0.03), p<0.001
	Backwards Stepwise	340	0.359	PLP	0.12 (0.08, 0.16), p<0.001	Folate	-0.10 (-0.14, -0.07), p<0.001
				Glycine	0.08 (0.05, 0.12), p<0.001	Alanine	-0.09 (-0.13, -0.06), p<0.001
				Proline	0.06 (0.03, 0.10), p<0.001	Betaine	-0.08 (-0.12, -0.05), p<0.001
				Aspartate	0.04 (0.01, 0.08), p=0.015	B12	-0.07 (-0.10, -0.04) p<0.001
						B2	-0.03 (-0.07, 0.00), p<0.037
	Lasso	344	0.297	PLP	0.10	Folate	-0.10
				Glycine	0.08	Alanine	-0.07
				Proline	0.05	Betaine	-0.07
				Aspartate	0.03	B12	-0.06
				Choline	0.02	B2	-0.03
				Glutamate	0.01	Tyrosine	-0.02
				Valine	0.01	Threonine	-0.01
				Histidine	0.01		
				Arginine	0.01		
				Serine	0.01		
				Phenylalanine	0.01		
Dry season (Feb-Apr)	2-step	82	0.119	None		None	
	Backwards Stepwise	82	0.402	PLP	0.11 (0.05, 0.17), p<0.001	Betaine	-0.10 (-0.16, -0.04), p<0.001
				Proline	0.11 (0.03, 0.19), p=0.005	Alanine	-0.09 (-0.16, -0.02), p=0.009
						Folate	-0.08 (-0.15, 0.00), p=0.039
						B2	-0.06 (-0.11, 0.00), p= 0.033
	Lasso	85	0.280	PLP	0.09	Betaine	-0.04
				Proline	0.05	B12	-0.03
				Serine	0.02	B2	-0.03

Seasonal Model	Statistical approach*	N	R^2	Positive predictors[‡] (β, 95%CI, p value)		Negative predictors[‡] (β, 95%CI, p value)	
						Choline	-0.02
						Threonine	-0.02
						Folate	-0.01
Rainy season (Jul-Oct)	2-step	98	0.382	None		B2	-0.07 (-0.13, 0.00), p=0.038
	Backwards stepwise	98	0.480	Valine	0.12 (0.02, 0.23), p=0.02	Alanine	-0.11 (-0.19, -0.03), p=0.006
				PLP	0.10 (0.02, 0.18), p=0.013	Betaine	-0.08 (-0.15, -0.01), p=0.019
				Proline	0.09 (0.01, 0.17), p=0.027		
	Lasso	98	0.216	Glutamate	0.05	Betaine	-0.06
				PLP	0.03	B2	-0.03
						B12	-0.02
						Folate	-0.02

*Lasso model fitted to variables demonstrating the smallest Cp value. Note no confidence intervals are reported for Lasso models.

[‡]Predictors are adjusted for maternal age, BMI, gestational age and inflammation (AGP). Only variables with a final p<0.05 reported.

[‡]β Coefficients represent the change in log homocysteine for a one SD increase in the log-transformed independent variable.

Abbreviations: AGP, Alpha-1-acid glycoprotein; CI, confidence interval; PLP, pyridoxal 5'-phosphate

S1 Table 6.3: Nutritional Biomarker Independent Predictors of Homocysteine without Cysteine & Dimethylglycine in the Indicator group

Seasonal Model	No. observations	No. women	Overall R2*	Positive predictors[ŧ] (β[‡], 95% CI, p value)		Negative predictors[ŧ] (β[‡], 95% CI, p value)	
Combined seasons	288	48	0.226	Methionine	0.05 (0.02, 0.07), p<0.001	Folate	-0.08 (-0.11, -0.05), p<0.001
						Betaine	-0.06 (-0.10, -0.02), p=0.002
						B2	-0.04 (-0.08, 0.00), p=0.042
Dry (Feb-Apr)	79	34	0.188	None		None	
Rainy (Jul-Oct)	63	28	0.264	None		Folate	-0.09 (-0.15, -0.02), p=0.006
						B12	-0.09 (-0.17, 0.00), p=0.046

*Two-step linear regression using random effects

[ŧ] Predictors are those variables retained in the multivariable model with p<0.05, adjusted for age.

[‡]β Coefficients represent the change in log homocysteine for a one SD increase in the log-transformed independent variable.

Abbreviations: CI, confidence interval

Supplementary Figure 6.1: Summary of retained coefficients of multivariable linear regression models without cysteine & DMG in three datasets*.

Combined seasons

Biomarker	MDEG2 2-step	MDEG2 Stepwise	MDEG2 LASSO	MDEG1 2-step	MDEG1 Stepwise	MDEG1 LASSO	Indicator 2-step
Folate	-0.1	-0.1	-0.1	-0.09	-0.1	-0.07	-0.08
B2		-0.03	-0.03	-0.06	-0.06	-0.04	-0.04
Choline			0.02				
B12	-0.06		-0.06				
PLP	0.09	0.12	0.1				
Betaine	-0.07	-0.08	-0.07				-0.06
Glycine	0.06	0.08	0.08				
Glutamate							
Proline		0.06	0.05				
Aspartate		0.04	0.03				
Tyrosine			-0.02				
Alanine		-0.09	-0.07				
Serine							
Threonine							
Methionine							0.05
Valine							

Dry season

	MDEG2 2-step	MDEG2 Stepwise	MDEG2 LASSO	MDEG1 2-step	MDEG1 Stepwise	MDEG1 LASSO	Indicator 2-step
Folate		-0.08	-0.01	-0.16	-0.16	-0.1	
B2		-0.06		-0.07	-0.07	-0.03	
Choline			-0.02				
B12			-0.03				
PLP		0.11	0.09				
Betaine		-0.1	-0.04				
Glycine							
Glutamate							
Proline		0.11	0.05				
Aspartate							
Tyrosine							
Alanine		-0.09					
Serine			0.02				
Threonine			-0.02				
Methionine							
Valine							

Rainy season

Biomarker	MDEG2 2-step	MDEG2 Stepwise	MDEG2 LASSO	MDEG1 2-step	MDEG1 Stepwise	MDEG1 LASSO	Indicator 2-step
Folate			-0.02			-0.02	-0.09
B2	-0.07		-0.03			-0.01	
Choline						0.03	
B12			-0.02				-0.09
PLP		0.1	0.03				
Betaine		-0.08	-0.06				
Glycine							
Glutamate			0.05				
Proline		0.09					
Aspartate							
Tyrosine							
Alanine		-0.11					
Serine							
Threonine							
Methionine							
Valine		0.12					

*Note: Coefficients represent change in log homocysteine for each SD increase in the log-transformed independent variable. Only variables with p<0.05 shown. Those inversely associated with homocysteine are shaded in green, and those positively associated are shaded in red.

Abbreviations: DMG, dimethylglycine; MDEG1, MDEG main study

Chapter 7 Design of the MDEG-2 Methyl Donor Supplement

_____

Summary of Chapter

In this chapter I summarise the development of the novel drink powder supplement designed for the clinical trial. This includes:

- An overview of the product characteristics of the new drink powder supplement.
- A section focussing especially on betaine, since this is the first known time betaine has been used in a supplement in The Gambia.
- A section overviewing the use of folic acid, vitamin B2 and vitamin B12, the other supplement ingredients that have been used several times before.

Previous trials that have used the supplement ingredients in interventions, especially pregnancy trials where possible, have been summarised in the chapter, with more detailed information on each trial provided in **Annex 7.1.**

7.1   Introduction

7.1.1   Supplement design rationale

Section 6.1 summarises the rationale for developing a novel nutritional supplement, explaining the long-term goal of shifting the maternal metabolome to optimise regulation of the infant epigenome by providing micronutrients in the ratio and quantity necessary for optimal one-carbon metabolism all year round. That section also gives the reasons why in the trial of the novel supplement amongst non-pregnant women the primary objective is to reduce plasma homocysteine. The hypothesis is that reducing homocysteine will help improve the methylation potential and enable one-carbon metabolic pathways to function unhindered.

7.1.2   Summary of research leading to design

Information presented in Chapters 4, 5 and 6 helped provide the ground work for the design of a novel nutritional supplement, called the 'MDEG-2 Methyl Donor Supplement'. Chapter 4 presented analysis from a sub-sample of women from West Kiang enrolled into the ENID trial (n=350). Geometric mean (95% CI) homocysteine concentration is higher in the peak dry season (February – April) compared to the peak rainy season (July-September) [8.06 (7.65, 8.50) µmol/L vs. 7.31 (6.82, 7.84) µmol/L; p=0.023]. These results corroborate longitudinal data from previous studies demonstrating similar observations in seasonal homocysteine concentrations[1,2]. The rationale was therefore to design a supplement for this pilot study that could decrease dry season homocysteine plasma concentrations to at least those found in the rainy season, if not lower. Chapter 4 also showed that over a quarter of women had a low plasma folate concentration (<10nmol/L), 39% of women had a low plasma B2 concentration (<10nmol/L), 9.4% had a low B12 concentration (<221 pmol/L) and a quarter had a betaine concentration below the normal range (<16µmol/L). Figure 7.1 presents a summary of the seasonal trends for the biomarkers of interest using the data presented in Chapter 4.

Recommendations from Chapter 5 suggested that future supplementation studies for epigenetic outcomes would need to consider the underlying nutritional status of the participants, because the maternal nutritional predictors of offspring DNA methylation vary between the dry and rainy seasons. Although the clinical trial described in Chapter 8 does not look at epigenetic outcomes, the plan was to conduct the study in the dry season as this is the period associated with lower maternal methylation potential.

Chapter 6 provided the analyses from three separate datasets from West Kiang to determine the consistent nutritional predictors of decreased plasma homocysteine. These were folate, B2, B12 and betaine, and provided the final list of components for supplement consideration. There is already a lot of evidence suggesting similar nutritional interventions have been successful in reducing homocysteine[3–5]. Whilst the ingredients comprising this supplement have been widely used before (see review below), this is the first time to our knowledge this specific combination and dosage of micronutrients have been used. The final supplement composition is described in section 7.2 below.
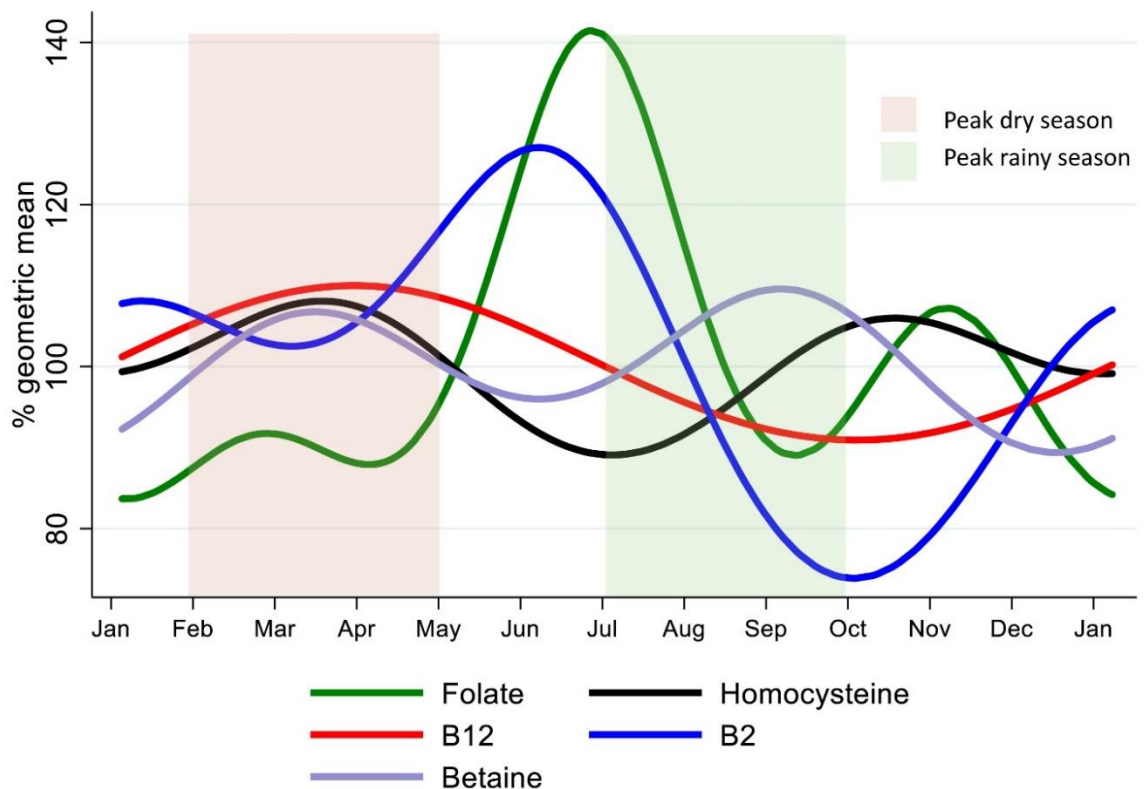
Figure 7.1: Summary of seasonal trends of folate, homocysteine, B2, B12 and betaine in first trimester pregnancy in West Kiang (N=350). Data presented in Chapter 4.

7.1.3   MDEG-2 Methyl Donor Supplement development

Literature reviews and data analysis for supplement ingredients took place from September 2016 to June 2017. Supplement design was finalised on 7th July 2017 following a final meeting between the PhD candidate and Prof Andrew Prentice (Nutrition theme leader, MRC Gambia), Dr Matt Silver (Senior Investigator Scientist, MRC Gambia), Prof Patrick Stover (external advisor, Cornell University), Dr Martha Field (external advisor, Cornell University) and Dr Sophie Moore (external advisor, Kings College London). The Joint Gambian

government / MRC Gambia Scientific Coordination Committee approved the clinical trial protocol for submission to the Ethics Committee on 11th October 2017. Full ethical approval was granted by the Joint Gambian government / MRC Gambia Ethics Committee on 11th December, 2017. Production of the supplement was started by Kendy Ltd. on 11th December, 2017. Import authorisation from the Medicines Control Agency in The Gambia was granted on 9th February 2018, and the supplements arrived in The Gambia ready for the trial described in Chapter 8 on 19th February, 2018.

The following sections provide a summary of the overall product characteristics of the MDEG-2 methyl donor supplement, followed by a section explicitly focussing on the characteristics of betaine, since this is the component that has not previously been used in nutritional supplements in The Gambia. The chapter finishes with a brief review of the use of vitamins B12, B2 and folic acid.

7.2   Overall characteristics of the MDEG-2 Methyl Donor Supplement

The novel supplement is a white, crystalline powdered, water-soluble nutritional supplement provided in daily dose sachets of net weight 4011mg. It is designed to be taken orally after being fully dissolved in 200ml of water. The dosage for adults 16 years and older is one sachet (4011mg) per day for 12 weeks. The supplement has been designed to reduce plasma homocysteine concentrations. In the planned clinical trial the hypothesis is it will reduce plasma homocysteine by at least 1µmol/L over 12 weeks of daily supplementation.

Table 7.1 details the product formulation. Given the poor nutritional status of women of West Kiang in B2, B12 and folate (Table 4.1) the dose is twice the dietary reference intake of these components (explored in more detail in section 7.4). The dose of betaine is 4g per day (justified in section 7.3). In order to ensure each sachet contains the appropriate quantity of the active ingredients a correction for overage is considered. Vitamin B2 uses a conversion factor of 1.4286 to ensure the appropriate amount of active ingredient is incorporated (2.80mg x 1.4286 x 10% overage = 4.40mg). Vitamin B12 is obtained from 0.1% titrated powder (5.2µg x 1000 x 10% overage = 5.70mg). The calculation for folic acid is more straight forwards (800µg x 12.5% overage = 0.9mg) and for the pure betaine powder no overage is necessary due to the large volume (4g). There are no excipients in the supplement.

Table 7.1 MDEG-2 Methyl Donor Supplement active ingredients

Active Ingredients	Generic Name	Anatomical Therapeutic Chemical code*	Quantity of pure active substance	Total quantity of chemical compound (mg)
Vitamin B2	Riboflavin 5-phosphate	A11HA04	2.80mg	4.40
Vitamin B9	Folic Acid	B03BB01	800µg	0.90
Vitamin B12	Cyanocobalamin	B03BA01	5.2 µg	5.70
Betaine	Betaine	A16AA06	4g	4000.00
TOTAL WEIGHT				**4011.0 mg**

* The Anatomical Therapeutic Chemical classification system is designated by the WHO Collaborating Centre for Drug Statistics Methodology: https://www.whocc.no/atc/structure_and_principles/

7.2.1   Potential drug interactions

Folic acid can decrease the effectiveness of the drugs used to prevent seizures: Fosphenytoin (Cerebyx), Phenobarbital (Luminal), Phenytoin (Dilantin), Primidone (Mysoline). It can also decrease the effectiveness of certain drugs for cancer treatment: Adrucil (Fluorouracil), Xeloda (capecitabine) and Methotrexate (MTX, Rheumatrex). Patients at risk of seizure or undergoing cancer treatment should not receive this supplement. Folic acid, in high doses, can reduce the effectiveness of antifolate antimalarials (such as Fansidar)[6]. The risk of this is low in the study population given folate status is low to start with and the supplement contains a low dose of folic acid, however, women taking antifolate antimalarials should be closely monitored for any signs of malaria.

7.2.2   Fertility

There is not yet epidemiological data on use of this supplement in pregnancy and lactation given it is the first time of usage. However, folic acid, B12 and B2 are commonly administered in pregnancy at the doses contained in this supplement. There have been no documented adverse effects of betaine given in pregnancy, although data is limited. Animal reproduction studies have not been conducted. Administering this supplement during pregnancy is compatible with good maternal and fetal outcomes based on available literature[7,8]. There is currently a lack of data on whether betaine anhydrous is excreted in human milk. Supplementation of lactating women should be done with careful monitoring until more data is available. No data is available on effects of the supplement on fertility.

7.2.3 Undesirable effects

Riboflavin may cause urine to be more yellow in colour. Betaine may increase plasma levels of methionine. There are no reported cases of overdose for any of the supplement ingredients.

7.2.4 Pharmacological properties

The supplement reduces levels of homocysteine by tackling the two pathways of methylating homocysteine to methionine. Chapter 6 describes the mechanisms by which plasma homocysteine can be decreased in detail. In brief, one way in which this can happen is by accepting a methyl group to form methionine[9], using two distinct pathways. The major one is the vitamin B12- dependent reaction involving folate metabolic pathways[10], chiefly the donation of a methyl group from N^5-methyl tetrahydrofolate. An alternative pathway, predominantly used in the liver and kidneys, uses the methyl group from betaine, a product formed through the oxidation of choline[11,12]. Homocysteine can also be removed through its irreversible degradation to cystathionine and cysteine in the transsulfuration pathway requiring vitamin B6[10].

7.2.5 Production, shipping and storage

The supplement is manufactured and packed by Kendy Ltd., 101 Sofia Street, 1320 Bankya, Bulgaria for MRC Unit The Gambia, Atlantic Boulevard, Faraja, The Gambia. It is packed into sachets containing one daily dose (4011mg). The supplement is shipped in boxes of 60 sachets. It should be stored in a dry place below 25°C. Expiry of the unopened supplement is two years after the date of manufacture.

Kendy Ltd. received its Good Manufacturing Practice certification from the Bulgarian Drug Agency (BG/GMP/2016/090). The Certificate of Analysis for the supplement was supplied by the Kendy Ltd. quality control manager (No. 1034/18.12.2017).

7.3   Characteristics of betaine

Table 7.2 provides some of the physical, chemical and pharmaceutical properties of the betaine dose used in the supplement.

Table 7.2 Overview of key properties of betaine

Daily dose:	4g
PubChem CID:	247
Chemical formula:	$C_5H_{11}NO_2$
Physical properties:	Solid, molecular weight 117.148 g/mol, sweet taste, water-soluble[13]
Synonyms:	Glycine betaine, trimethylglycine, betaine anhydrous, Cystadane®
Pharmaceutical properties:	Methylation agent used to lowers homocysteine, to increase low plasma methionine and S-adenosyl methionine (SAM) in patients with MTHFR deficiency, to treat homocystinuria[14].
Handling hazards:	The betaine will be pre-packaged in the supplement sachets, so no direct handling will be necessary. In the event of accidental spillage most companies (169/193) declare no hazards[15]. The remaining companies suggest avoiding contact with eyes and prolonged contact with skin due to irritant properties.

7.3.1   Metabolism and toxicology

Betaine can be directly sourced from the diet (e.g. wheat bran, wheat germ, spinach, beets[16]). Intake from food is estimated at 0.5-2 g/d[17]. It can also be formed endogenously from choline. Choline can be synthesised or obtained from the diet (good sources include red meat, poultry, milk, eggs and fish[16]). A two-step oxidation reaction converts it to betaine[7], a process occurring mainly in the liver and kidneys[18]. Betaine can donate its methyl group to homocysteine via betaine-homocysteine methyl transferase (BHMT)[7]. Betaine is absorbed into the enterocytes from the ileum and then released into the portal circulation, where it is transported to the liver[18]. Bioavailability of ingested betaine is thought to be almost 100%[19]. When betaine is metabolised its methyl group is transferred to homocysteine via BHMT, enabling the formation of methionine from homocysteine[7]. In losing its methyl group betaine is converted to dimethylglycine[7]. Urinary excretion of betaine is variable and appears to be under homeostatic control[7].

In a toxicology study of male and female Sprague-Dawley rats (n=50), 0, 1, 2, and 5% betaine (corresponding to 0-7143mg/kg) was added to a maintenance chow and the rats were followed for up to 90 days. No toxicity occurred[20]. It is reported that doses equal to or greater than 10,000 mg/kg in rats frequently caused death[21]. For human use, the manufacturers of Cystadane® (oral betaine anhydrous powder) state that whilst long-term carcinogenicity and reproductive toxicity studies have not been conducted, a 'standard battery of genotoxicity test reveals no specific hazard for humans'[14]. Betaine is categorised for hepatotoxicity as likelihood score E by the US National Library of Medicine, which

corresponds to an 'unlikely cause of clinically apparent liver injury' and that 'despite extensive use, there is no evidence that the drug has caused liver injury'[22].

7.3.2 Effects in humans

Betaine is a natural food product, but has been used therapeutically in humans to reduce homocysteine. Table 7.3 summarises betaine supplementation intervention trials in humans.

Table 7.3 Summary of trials assessing effect of oral betaine on homocysteine

Reference	Betaine daily dose	Duration of trial	Participants	Results	Serious adverse events?
Schwab et al. (2002)[23]	6g	12 weeks	N=42, mean age = 44 years, healthy adults	Reduction of Hcy by 1µmol/l	No
Steenge et al. (2003)[24]	6g	6 weeks	N=24, mean age = 44.5 years, healthy adults	Reduction of Hcy by 1.8µmol/l	No
Olthof et al. (2003)[25]	6g	6 weeks	N=38, healthy adults	Reduction of Hcy by 2.2µmol/l	No
Olthof et al. (2006)[26]	6g	6 weeks	N=39, mean age = 59 years, healthy adults	Reduction of Hcy by 1.2µmol/l	No
Schwab et al. (2011)[27]	4g	24 weeks	N=63, mean age = 27 years, healthy adults	No significant reduction of Hcy.	No

In a meta-analysis of the studies included the Table 7.3 the authors concluded that 'supplementation of betaine at 4 to 6 g/d significantly lowers plasma homocysteine concentration in healthy adults by 1.23 µmol/L or 11.8% of baseline values'[5].

There were no serious adverse effects in any of the trials listed in Table 7.3. In larger doses of 6g/day three small studies have reported a small rise of 10mg/dL low-density lipoprotein cholesterol, however, this effect has not been consistently seen in trials and is not considered to be a clinically significant side effect[5].

The Summary of Product Characteristics documentation for Cystadane® contains information on side effects. Cystadane® is used to treat homocystinuria and is recommended in a dose of 100mg/kg/day. For a 60kg adult this would be 6g total per day. Documented side effects for Cystadane® are provided in Table 7.4. Note that the side-effects of brain oedema and blood methionine increases in Table 7.4 are restricted to patients with CBS deficiency. In all cases symptoms were reversed and complete recovery attained after stopping

treatment[14]. The UK National Institute for Health and Care Excellence (NICE) report all side-effects related to oral betaine as uncommon[28]. No cases of overdose have been reported and no drug interaction studies have been performed.

Table 7.4 Side-effects for Cystadane®

Side-effect category	Frequency*
Metabolism and nutrition disorders	Uncommon: anorexia
Psychiatric disorders	Uncommon: agitation, irritability
Nervous system disorders	Uncommon: brain oedema
Gastrointestinal disorders	Uncommon: diarrhoea, glossitis, nausea, stomach discomfort, vomiting
Skin and subcutaneous tissue disorders	Uncommon: hair loss, hives, skin odour abnormal
Renal and urinary disorders	Uncommon: urinary incontinence
Investigations	Very common: blood methionine increased

*Frequency: 'uncommon' (≥ 1/1,000 to < 1/100); very common (≥1/10)
Source: Cystadane® Summary of Product Characteristics[14]

7.3.3   Overall summary of safety of betaine in humans

Betaine anhydrous is Generally Recognised As Safe (GRAS) and so is exempted from the usual Federal Food, Drug, and Cosmetic Act (FFDCA) food additive tolerance requirements. This means there are no official upper limits to its consumption. However, the UK National Institute for Health and Care Excellence advise a maximum of 20g per day[28]. With the proposed dosage of 4g/day there are no foreseen adverse effects related to the consumption of this supplement ingredient.

7.4   Characteristics of Vitamins B2, B12 and Folic Acid

Some of the physical, chemical and pharmaceutical properties of the B vitamins used in the supplement are shown in Table 7.5. All three vitamins can be found naturally in certain foods. Folate is found in green leafy vegetables, liver and wholegrain cereals. Vitamin B2 is found in milk, eggs, yeast and liver. Vitamin B12 is found in liver, meat and dairy foods.

Table 7.5 Overview of properties of folic acid, B12 and B2

	Vitamin B2[29]	Vitamin B12[30]	Folic Acid[31]
Daily dose:	2.80mg	5.2 µg	800µg
PubChem CID:	493570	5311498	6037
Chemical formula:	$C_{17}H_{20}N_4O_6$	$C_{63}H_{88}CoN_{14}O_{14}P$	$C_{19}H_{19}N_7O_6$
Molecular Weight:	376.369 g/mol	1355.388 g/mol	441.404 g/mol
Physical properties:	Water-soluble, yellow crystalline powder, solid, bitter	Water-soluble, red crystalline powder, solid, tasteless	Low water solubility, yellow crystalline powder, tasteless
Synonyms:	Riboflavin 5-phosphate, lactoflavin, riboflavin, Vitamin G, lactoflavine	Cyanocobalamin, cobalamin, 68-19-9, crystamine, anacobin	59-30-3; folate, pteroylglutamic acid, Vitamin M, Folacin
Pharmaceutical properties:	Co-enzyme role in energy production and 1-carbon metabolism	Hematopoiesis, neural metabolism, DNA and RNA production, macronutrient metabolism and 1-carbon metabolism	Hematopoiesis, one-carbon donor, purine & pyrimidine synthesis
Handling hazards:	No hazard	No hazard (34/35 companies). 1 company suggests avoid eye and skin contact (irritant)	Avoid eye and skin contact (irritant)

7.4.1 Metabolism and toxicology

Homocysteine concentrations can be lowered through one of two major pathways as described above. Firstly, the transsulfuration pathway causes its irreversible degradation when metabolising it to cysteine. Secondly, the methylation of homocysteine forms methionine, and is the pathway that this supplement targets.

Dietary folates and folic acid are reduced to form tetrahydrofolate (THF), which is a scaffold upon which one-carbon units can be attached and activated. THF in turn is reduced to methylene-THF, then to 5-methyl-THF, which donates its methyl group to homocysteine using vitamin B12[32]. The different forms of THF are interconvertible, except the methyl-THF form. In the absence of sufficient B12 to utilise the methyl group from methyl-THF in homocysteine methylation (a process which re-generates THF), methyl-THF can accumulate at the expense of other THF forms. This is termed the 'folate trap' or 'methyl trap' and means pathways dependent on other forms of THF (e.g. purine synthesis and the thymidylate pathway) can become impaired[33].

Riboflavin is combined with ATP to form two important coenzymes: flavin mononucleotide (FMN) in a reaction catalysed by riboflavin kinase, and flavin-adenine dinucleotide (FAD), which is formed from FMN using FAD synthase[34,35]. FAD is a cofactor required for MTHFR to reduce methylene-THF to methyl-THF[36]. Vitamin B12 is a cofactor in the methylation of homocysteine by 5-methyl-THF. In human plasma and the cytosol the predominant form is methylcobalamin[37]. The B12 (in methyl, adenosyl or cyano (synthetic) form) is transported in plasma bound to two proteins: haptocorrin or transcobalamin. The binding of B12 to transcobalamin forms holotranscobalamin, and this is the only form taken into cells. Once in the cell lysosome the transcobalamin is degraded and the B12 used for intracellular metabolism. If the B12 is in 5'-deoxyadenosylcobalamin form it is used as a coenzyme for methylmalonyl CoA mutase. If it is in the methylcobalamin form it is used as a coenzyme for methionine synthase.

Riboflavin (Vitamin B2) is mainly absorbed in the small intestine by active or facilitated transport when at low levels and by passive diffusion at high levels[38]. Some is also absorbed in the colon[39]. It is mainly transported in the plasma attached to proteins such as immunoglobulins[38]. Vitamin B12 is bound to Intrinsic Factor (IF) and absorbed in the ileum. The IF-B12 complex is taken into the intestinal enterocytes by receptor-mediated endocytosis[37]. In the enterocyte B12 is released from IF and is transported in plasma bound to transcobalamin or haptocorrin[37]. Folic acid is a synthetic version of folate, which does not occur naturally. It is fully oxidised and requires a two-step reduction process to get it to the active THF form. To be taken up and transported THF needs to be in monoglutamate form. Monoglutamate-THF is absorbed in the duodenum and jejenum by Slc46a1 (a proton-coupled transporter). It enters the circulation and is taken into cells by folate receptors 1&2 and by Slc19a1 reduced folate carrier. The circulating form in plasma is free or bound to albumin or other proteins.

Riboflavin intake in excess of requirements is excreted in urine[39]. Vitamin B12 is stored in the liver and mainly secreted in bile. Excess can be excreted in urine[30]. Excess folic acid is excreted in urine[31].

In animal studies riboflavin has proved toxic in doses of 50mg/kg in mice[29]. For vitamin B12 1.5-3 mg/kg body weight have caused convulsions then cardiac or respiratory failure in mice[30]. In humans there is no reported toxicity for riboflavin, B12 and folic acid given in oral doses.

7.4.2 Effects in humans

Folic acid, B12 and B2 are common supplements provided in research studies and public health campaigns. Table 7.6 summarises the previous doses given in different countries in supplementation trials for a variety of health outcomes. It shows how the doses provided in the MDEG-2 methyl donor supplement have been already been commonly utilised. Further information of each of the studies included in this table is provided in **Annex 7.1**, which details the trial population, location, doses, primary outcomes and whether any side effects were reported.

Table 7.6 Overview of previous doses given of folic acid, B2 and B12

Ingredient	Recommended daily intakes and biomarker levels	Amount in MDEG2 Methyl Donor Supplement	Amounts given in previous trials
Folate given as folic acid	UK Recommended Daily Allowance (RDA): 200 µg/d Pregnancy + 100 µg/d[40] USA Dietary Reference Intake (DRI): 400 µg/d in pregnancy[41]	800 µg/d (2x Dietary Reference Intake [DRI])	200 µg/d[42] 215 µg/d[43] 300 µg/d[44] 350 µg/d[45] 400 µg/d[46–60] 500 µg/d[61] 600 µg/d[62] 800 µg/d[24,46,63–65] Increments between 50–800µg/d[66] 150-2500 µg/d[67] 2500 µg/d[68] 4000 µg/d [69] 2800 µg/wk[70]
Vitamin B12 given as cyanocobalamin	UK RDA : 1.5µg/d[40] USA RDA: 2.6 µg/d in pregnancy[41]	5.2 µg/d (2x DRI)	2.0 µg/d[43] 2.2 µg/d[51] 2.6 µg /d[47–49,54–60,62] 3 µg/d[45] 4 µg/d[65,71] 5.2 µg/d[52] 6 µg/d[53] 20 µg/d[72] 50 µg/d[64] 400 µg/d[63,68]
Riboflavin (vitamin B2) given as Riboflavin-5'-phosphate	UK RDA for adult women: 1.1mg/d, +0.3mg in pregnancy[40] USA RDA: 1.4 mg/d in pregnancy[41]	2.8 mg/d (2x DRI)	1.4 mg/d[47–49,55–58,60,62] 1.5mg/d[69] 1.6 mg/d[42,71] 1.8 mg/d[54,65] 1.9 mg/d[43,59] 2 mg/d[53] 2.8 mg/d[52] 5 mg/d[44] 15 mg/d[45] 20 mg/d[64]

There are no reported serious adverse effects to taking riboflavin supplements[73], however, some individuals may find their urine is more yellow in colour than normal. There are no documented significant adverse effects of taking B12 supplement, even at doses several hundred times what is contained in this supplement[73]. There are no documented significant adverse effects of taking folic acid supplements at this dose.

7.4.3 Overall summary of safety in humans

Supplement doses for folic acid, B2 and B12 are proposed at twice the Dietary Reference Intake (DRI) set for the United States of America[41] in order to correct the micronutrient deficiencies in the dry season.

Folic acid (proposed supplement 800 µg/d, 2xDRI):

The proposed supplement dose has been used before in several large trials[63–65], and is well below the amount expected to cause adverse effects. The Expert Group on Vitamins & Minerals (2003) states that "in the general population a supplemental dose of 1 mg/day (equivalent to 0.017 mg/kg bw/day in a 60 kg adult) would not be expected to cause adverse effects"[74]. The European Agency for Food Safety (2006) states that the Lowest observed-adverse-effect level (LOAEL) is 5 mg per day[73], and current US government advice for women who have previously had a baby with a birth defect is to take up to 4000 µg/d[75]. Side effects have only been documented with doses in excess of 1mg/day.

Riboflavin (B2) (proposed supplement 2.8 mg/d, 2xDRI):

Riboflavin is Generally Recognised As Safe (GRAS), even when given in doses up to 14 times what is contained in the supplement[76]. GRAS Report 114 on Riboflavin states that "there is no evidence in the available information on riboflavin or riboflavin-5'-phosphate that demonstrates or suggest reasonable grounds to suspect, a hazard to the public when they are used at levels that are now current or that might reasonably be expected in the future"[76]. The Expert Group on Vitamins & Minerals (2003) also concur, stating that "supplemental intakes of 40 mg riboflavin/day (equivalent to 0.67 mg/kg bw for a 60 kg adult) would be unlikely to result in adverse effects. This is in addition to riboflavin provided by the diet"[74].

Cyanocobalamin (B12) (proposed supplement 5.2 µg/d, 2xDRI):

The GRAS Report 104 on Cyanocobalamin states that "there is no evidence in the available information on vitamin B12 that demonstrates, or suggests reasonable grounds to suspect,

a hazard to the public when it is used at levels that are now current and in the manner now practiced, or that might reasonably be expected in the future"[76]. The Expert Group on Vitamins & Minerals (2003) states that "it is generally accepted that ingested vitamin B12 (cobalamin) has a very low toxicity in humans...supplemental 2.0 mg cyanocobalamin/day should not produce any adverse effects and this intake can be used for guidance purposes"[74].

7.5   Overall conclusion on usage of MDEG-2 Methyl Donor Supplement

The ingredients included in this supplement are extremely unlikely to result in any adverse events when taken at the correct dosage of one sachet per day. There are many existing commercial formulations that contain doses several hundred to several thousand times more than the MDEG-2 methyl donor supplement (e.g. Bluebonnet Nutrition Homocysteine Formula, Thorne Methyl Guard Plus, Biocare Methyl Multinutrient). Considering the research project will be delivered under observed supplementation, and provided against the backdrop of micronutrient deficiencies associated with the Gambian dry season, it is considered a safe nutritional supplement tailored to the requirements of the target population.

7.6 References

1. Dominguez-Salas P, Moore SE, Cole D, et al. (2013) DNA methylation potential: dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women. *Am. J. Clin. Nutr.* **97**, 1217–27.

2. Dominguez-Salas P, Moore SE, Baker MS, et al. (2014) Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. *Nat. Commun.* **5**, 3746.

3. (1998) Lowering blood homocysteine with folic acid based supplements: meta-analysis of randomised trials. Homocysteine Lowering Trialists' Collaboration. *BMJ* **316**, 894–8.

4. Huang T, Zheng J, Chen Y, et al. (2011) High consumption of Ω-3 polyunsaturated fatty acids decrease plasma homocysteine: a meta-analysis of randomized, placebo-controlled trials. *Nutrition* **27**, 863–7.

5. McRae MP (2013) Betaine supplementation decreases plasma homocysteine in healthy adult participants: a meta-analysis. *J. Chiropr. Med.* **12**, 20–25.

6. Verhoef H, Veenemans J, Mwangi MN, et al. (2017) Safety and benefits of interventions to increase folate status in malaria-endemic areas. *Br. J. Haematol.* **177**, 905–918.

7. Ueland PM (2011) Choline and betaine in health and disease. *J. Inherit. Metab. Dis.* **34**, 3–15.

8. Shaw GM, Carmichael SL, Yang W, et al. (2004) Periconceptional dietary intake of choline and betaine and neural tube defects in offspring. *Am. J. Epidemiol.* **160**, 102–9.

9. Robinson JL & Bertolo RF (2016) The Pediatric Methionine Requirement Should Incorporate Remethylation Potential and Transmethylation Demands. *Adv. Nutr.* **7**, 523–34.

10. Selhub J (1999) Homocysteine metabolism. *Annu. Rev. Nutr.* **19**, 217–46.

11. Velzing-Aarts F V, Holm PI, Fokkema MR, et al. (2005) Plasma choline and betaine and their relation to plasma homocysteine in normal pregnancy. *Am. J. Clin. Nutr.* **81**, 1383–9.

12. McKeever MP, Weir DG, Molloy A, et al. (1991) Betaine-homocysteine methyltransferase: organ distribution in man, pig and rat and subcellular distribution in the rat. *Clin. Sci.* **81**, 551–6.

13. National Center for Biotechnology Information (2017) Compound Summary for CID 247 (Betaine). *PubChem Compd. Database*.

14. Orphan Europe S.A.R.L. (2007) *Cystadane Summary of Product Characteristics*. Puteaux, France.

15. ECHA Europe (2017) Summary of Classification and Labelling: Betaine. *Eur. Chem.*

Agency. https://echa.europa.eu/information-on-chemicals/cl-inventory-database/-/discli/details/33607.

16. Cho E, Zeisel SH, Jacques P, et al. (2006) Dietary choline and betaine assessed by food-frequency questionnaire in relation to plasma total homocysteine concentration in the Framingham Offspring Study. *Am. J. Clin. Nutr.* **83**, 905–11.

17. Olthof MR, van Vliet T, Verhoef P, et al. (2005) Effect of homocysteine-lowering nutrients on blood lipids: results from four randomised, placebo-controlled studies in healthy humans. *PLoS Med.* **2**, e135

18. Ueland PM, Holm PI & Hustad S (2005) Betaine: a key modulator of one-carbon metabolism and homocysteine status. *Clin. Chem. Lab. Med.* **43**, 1069–75.

19. Schwahn BC, Hafner D, Hohlfeld T, et al. (2003) Pharmacokinetics of oral betaine in healthy subjects and patients with homocystinuria. *Br. J. Clin. Pharmacol.* **55**, 6–13.

20. Hayes KC, Pronczuk A, Cook MW, et al. (2003) Betaine in sub-acute and sub-chronic rat studies. *Food Chem. Toxicol.* **41**, 1685–700.

21. Knox C, Law V, Jewison T, et al. (2011) DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. Nucleic Acids Res. 39:D1035-41. Link to Glycine Betaine: http://www.drugbank.ca/drugs/DB06756.

22. Liver Disease Research Branch of the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) and the Division of Specialized Information Services of the National Library of Medicine (NLM) (2017) LiverTox: Betaine. https://livertox.nlm.nih.gov//Betaine.htm.

23. Schwab U, Törrönen A, Toppinen L, et al. (2002) Betaine supplementation decreases plasma homocysteine concentrations but does not affect body weight, body composition, or resting energy expenditure in human subjects. *Am. J. Clin. Nutr.* **76**, 961–7.

24. Steenge GR, Verhoef P & Katan MB (2003) Betaine supplementation lowers plasma homocysteine in healthy men and women. *J. Nutr.* **133**, 1291–5.

25. Olthof MR, van Vliet T, Boelsma E, et al. (2003) Low dose betaine supplementation leads to immediate and long term lowering of plasma homocysteine in healthy men and women. *J. Nutr.* **133**, 4135–8.

26. Olthof MR, Bots ML, Katan MB, et al. (2006) Effect of folic acid and betaine supplementation on flow-mediated dilation: a randomized, controlled study in healthy volunteers. *PLoS Clin. Trials* **1**, e10.

27. Schwab U, Alfthan G, Aro A, et al. (2011) Long-term effect of betaine on risk factors associated with the metabolic syndrome in healthy subjects. *Eur. J. Clin. Nutr.* **65**, 70–76.

28. National Institute for Health and Care Excellence (2018) NICE Guidance: Betaine. https://bnf.nice.org.uk/drug/betaine.html.

29. National Center for Biotechnology Information (2017) Compound Summary for CID

493570 (Riboflavin). *PubChem Compd. Database*.

30.   National Center for Biotechnology Information (2017) Compound Summary for CID 5311498 (Vitamin B12). *PubChem Compd. Database*.

31.   National Center for Biotechnology Information (2017) Compound Summary for CID 6037 (Folic Acid). *PubChem Compd. Database*.

32.   Scott JM & Weir DG (1998) Folic acid, homocysteine and one-carbon metabolism: a review of the essential biochemistry. *J. Cardiovasc. Risk* **5**, 223–7.

33.   Finkelstein JL, Layden AJ & Stover PJ (2015) Vitamin B-12 and Perinatal Health. *Adv. Nutr.* **6**, 552–63.

34.   Hustad S, McKinley MC, McNulty H, et al. (2002) Riboflavin, Flavin Mononucleotide, and Flavin Adenine Dinucleotide in Human Plasma and Erythrocytes at Baseline and after Low-Dose Riboflavin Supplementation. *Clin. Chem.* **48**, 1571–1577.

35.   Hustad S, Schneede J & Ueland PM (2013) Madame Curie Bioscience Database [Internet]. In *Madame Curie Bioscience Database [Internet].* Austin (TX): Landes Bioscience.

36.   Mason JB (2003) Biomarkers of Nutrient Exposure and Status in One-Carbon (Methyl) Metabolism. *J. Nutr.* **133**, 941S–947.

37.   Green R & Miller JW (2007) Vitamin B12. In *Handbook of Vitamins*, 4th ed. [ZempleniJ, Rucker R, McCormick D, et al., editors]. Boca Ralton, FL: Taylor & Francis Group.

38.   Institute of Medicine (2006) Dietary Reference Intakes: The Essential Guide to Nutrient Requirements. Washington, DC: The National Academies Press.

39.   Powers HJ (2003) Riboflavin (vitamin B-2) and health. *Am J Clin Nutr* **77**, 1352–1360.

40.   British Nutrition Foundation (2015) UK Dietary Reference Values. https://www.nutrition.org.uk/attachments/article/261/Nutrition Requirements_Revised Oct 2016.pdf.

41.   Institute of Medicine (US) Committee to Review Dietary Reference Intakes for Vitamin D and Calcium (2011) Dietary Reference Intakes for Calcium and Vitamin D. Washington (DC): National Academies Press.

42.   Hininger I, Favier M, Arnaud J, et al. (2004) Effects of a combined micronutrient supplementation on maternal biological status and newborn anthropometrics measurements: a randomized double-blind, placebo-controlled trial in apparently healthy pregnant women. *Eur. J. Clin. Nutr.* **58**, 52–9.

43.   Ramakrishnan U, González-Cossío T, Neufeld LM, et al. (2003) Multiple micronutrient supplementation during pregnancy does not lead to greater infant birth size than does iron-only supplementation: a randomized controlled trial in a semirural community in Mexico. *Am. J. Clin. Nutr.* **77**, 720–5.

44.   Kirke PN, Daly LE & Elwood JH (1992) A randomised trial of low dose folic acid to prevent neural tube defects. The Irish Vitamin Study Group. *Arch. Dis. Child.* **67**, 1442–

6.

45.   Rush D, Stein Z & Susser M (1980) A randomized controlled trial of prenatal nutritional supplementation in New York City. *Pediatrics* **65**, 683–97.

46.   Hall MN, Howe CG, Liu X, et al. (2016) Supplementation with Folic Acid, but Not Creatine, Increases Plasma Betaine, Decreases Plasma Dimethylglycine, and Prevents a Decrease in Plasma Choline in Arsenic-Exposed Bangladeshi Adults. *J. Nutr.* **146**, 1062–7.

47.   Bhutta ZA, Rizvi A, Raza F, et al. (2009) A comparative evaluation of multiple micronutrient and iron-folic acid supplementation during pregnancy in Pakistan: impact on pregnancy outcomes. *Food Nutr. Bull.* **30**, S496-505.

48.   Zagré NM, Desplats G, Adou P, et al. (2007) Prenatal multiple micronutrient supplementation has greater impact on birthweight than supplementation with iron and folic acid: a cluster-randomized, double-blind, controlled programmatic study in rural Niger. *Food Nutr. Bull.* **28**, 317–27.

49.   Zeng L, Dibley MJ, Cheng Y, et al. (2008) Impact of micronutrient supplementation during pregnancy on birth weight, duration of gestation, and perinatal mortality in rural western China: double blind cluster randomised controlled trial. *BMJ* **337**, a2001.

50.   Kumwenda N, Miotti PG, Taha TE, et al. (2002) Antenatal vitamin A supplementation increases birth weight and decreases anemia among infants born to human immunodeficiency virus-infected women in Malawi. *Clin. Infect. Dis.* **35**, 618–24.

51.   Rumiris D, Purwosunu Y, Wibowo N, et al. (2006) Lower rate of preeclampsia after antioxidant supplementation in pregnant women with low antioxidant status. *Hypertens. pregnancy* **25**, 241–53.

52.   Moore SE, Fulford AJ, Darboe MK, et al. (2012) A randomized trial to investigate the effects of pre-natal and infant nutritional supplementation on infant immune development in rural Gambia: the ENID trial: Early Nutrition and Immune Development. *BMC Pregnancy Childbirth* **12**, 107.

53.   Brough L, Rees GA, Crawford MA, et al. (2010) Effect of multiple-micronutrient supplementation on maternal nutrient status, infant birth weight and gestational age at birth in a low-income, multi-ethnic population. *Br. J. Nutr.* **104**, 437–45.

54.   Christian P, West KP, Khatry SK, et al. (2003) Effects of maternal micronutrient supplementation on fetal loss and infant mortality: a cluster-randomized trial in Nepal. *Am. J. Clin. Nutr.* **78**, 1194–202.

55.   Kaestel P, Michaelsen KF, Aaby P, et al. (2005) Effects of prenatal multimicronutrient supplements on birth weight and perinatal mortality: a randomised, controlled trial in Guinea-Bissau. *Eur. J. Clin. Nutr.* **59**, 1081–9.

56.   Liu J, Mei Z, Ye R, et al. (2013) Micronutrient Supplementation and Pregnancy Outcomes. *JAMA Intern. Med.* **173**, 276.

57.   Osrin D, Vaidya A, Shrestha Y, et al. (2005) Effects of antenatal multiple micronutrient

supplementation on birthweight and gestational duration in Nepal: double-blind, randomised controlled trial. *Lancet* **365**, 955–62.

58. Roberfroid D, Huybregts L, Lanou H, et al. (2008) Effects of maternal multiple micronutrient supplementation on fetal growth: a double-blind randomized controlled trial in rural Burkina Faso. *Am. J. Clin. Nutr.* **88**, 1330–40.

59. Supplementation with Multiple Micronutrients Intervention Trial (SUMMIT) Study Group, Shankar AH, Jahari AB, et al. (2008) Effect of maternal multiple micronutrient supplementation on fetal loss and infant death in Indonesia: a double-blind cluster-randomised trial. *Lancet* **371**, 215–27.

60. Tofail F, Persson LA, El Arifeen S, et al. (2008) Effects of prenatal food and micronutrient supplementation on infant development: a randomized trial from the Maternal and Infant Nutrition Interventions, Matlab (MINIMat) study. *Am. J. Clin. Nutr.* **87**, 704–11.

61. Schmidt MK, Muslimatun S, West CE, et al. (2001) Vitamin A and iron supplementation of Indonesian pregnant women benefits vitamin A status of their infants. *Br. J. Nutr.* **86**, 607–15.

62. West KP, Shamim AA, Mehra S, et al. (2014) Effect of maternal multiple micronutrient vs iron-folic acid supplementation on infant mortality and adverse birth outcomes in rural Bangladesh: the JiVitA-3 randomized trial. *JAMA* **312**, 2649–58.

63. Ebbing M, Bønaa KH, Nygård O, et al. (2009) Cancer incidence and mortality after treatment with folic acid and vitamin B12. *JAMA* **302**, 2119–26.

64. Fawzi WW, Msamanga GI, Urassa W, et al. (2007) Vitamins and perinatal outcomes among HIV-negative women in Tanzania. *N. Engl. J. Med.* **356**, 1423–31.

65. Czeizel AE, Dudás I & Métneki J (1994) Pregnancy outcomes in a randomised controlled trial of periconceptional multivitamin supplementation. Final report. *Arch. Gynecol. Obstet.* **255**, 131–9.

66. Melse-Boonstra A, Holm PI, Ueland PM, et al. (2005) Betaine concentration as a determinant of fasting total homocysteine concentrations and the effect of folic acid supplementation on betaine concentrations. *Am. J. Clin. Nutr.* **81**, 1378–82.

67. Dawson SL, Bowe SJ & Crowe TC (2016) A combination of omega-3 fatty acids, folic acid and B-group vitamins is superior at lowering homocysteine than omega-3 alone: A meta-analysis. *Nutr. Res.* **36**, 499–508.

68. Toole JF, Malinow MR, Chambless LE, et al. (2004) Lowering Homocysteine in Patients With Ischemic Stroke to Prevent Recurrent Stroke, Myocardial Infarction, and Death. *JAMA* **291**, 565.

69. MRC Vitamin Study Research Group (1991) Prevention of neural tube defects: results of the Medical Research Council Vitamin Study. *Lancet* **338**, 131–7.

70. Ramakrishnan U, Nguyen PH, Gonzalez-Casanova I, et al. (2016) Neither Preconceptional Weekly Multiple Micronutrient nor Iron-Folic Acid Supplements Affect Birth Size and Gestational Age Compared with a Folic Acid Supplement Alone

in Rural Vietnamese Women: A Randomized Controlled Trial. *J. Nutr.* **146,** 1445S-52S

71. Friis H, Gomo E, Nyazema N, et al. (2004) Effect of multimicronutrient supplementation on gestational length and birth size: a randomized, placebo-controlled, double-blind effectiveness trial in Zimbabwe. *Am. J. Clin. Nutr.* **80**, 178–84.

72. Galan P, Kesse-Guyot E, Czernichow S, et al. (2010) Effects of B vitamins and omega 3 fatty acids on cardiovascular diseases: a randomised placebo controlled trial. *BMJ* **341**, c6273.

73. European Food Safety Authority (2006) Tolerable upper intake levels for vitamins and minerals. EFSA: Parma, Italy.

74. Expert Group on Vitamins and Minerals (EVM) (2003) Safe Upper Levels for Vitamins and Minerals. Food Standards Angency: London, UK.

75. Office on Women's Health (2012) Folic Acid Fact Sheet. US Department of Healht and Human Services. https://www.womenshealth.gov/publications/our-publications/fact-sheet/folic-acid.html

76. U.S. Food and Drug Administration (1972) GRAS Substances (SCOGS) Database. Select Committee on GRAS Substances, Center for Food Safety and Applied Nutrition.

Chapter 8 Comparison of a novel drink powder and existing multiple micronutrient tablet (UNIMMAP) in reducing plasma homocysteine in non-pregnant women in rural Gambia: a randomised controlled trial to identify future candidates for epigenetic trials.

_____

Summary of Chapter

Background: Maternal periconceptional nutritional status is associated with infant DNA methylation profiles. DNA methylation relies on nutritional inputs for one-carbon metabolic pathways, including efficient recycling of homocysteine. This proof-of-concept trial in non-pregnant women tests the effectiveness of a supplement designed to improve one-carbon-related nutrient status and assess its potential future use in pregnancy trials.

Objective: To assess the effectiveness of a novel drink powder on reducing plasma homocysteine in non-pregnant women in rural Gambia.

Methods: We designed a novel drink powder based on determinants of plasma homocysteine in the target population. We conducted a 3-arm, unblinded, randomized, controlled trial. Non-pregnant women of reproductive age were randomized in a 1:1:1 allocation ratio to 12 weeks of daily supplementation of either a) a novel drink powder (4g betaine, 800μg folic acid, 5.2μg vitamin B12 and 2.8mg vitamin B2), b) a multiple micronutrient tablet (UNIMMAP) containing 15 micronutrients at the dietary recommended intake or c) no intervention. Supplementation was observed daily. Fasted venepuncture samples were collected at baseline, midline and endline to measure plasma homocysteine. We used linear regression models to determine the difference in homocysteine between pairs of trial arms at midline (week 5) and endline (week 12), adjusted for baseline homocysteine, age and body mass index.

Results: 298 eligible women were enrolled and randomised. Compliance was >97.8% for both interventions. At endline, compared to the controls, the drink powder and UNIMMAP reduced mean plasma homocysteine by 23.6% and 15.5% respectively (both p<0.001). Compared to UNIMMAP, the drink powder reduced mean homocysteine by 8.8% (p=0.025).

The effects were stronger at midline.

Conclusions: The trial confirms that dietary supplements can influence metabolic pathways that we have shown in previous studies to predict offspring DNA methylation. Both supplements reduced homocysteine effectively and remain potential candidates for future epigenetic trials in pregnancy in rural Gambia.

Notes

The chapter is written in a journal article format, ready for a further round of co-author review before submission. I also plan to submit a condensed version of Chapter 6 as supplementary material to explain how the supplement was designed, along with the CONSORT checklist once the article is finalised.

RESEARCH PAPER COVER SHEET

SECTION A – Student Details

Student	Philip James
Principal Supervisor	Dr Matt Silver
Thesis Title	Maternal nutrition, seasonality and epigenetics: an exploration of one-carbon metabolism and a novel nutritional supplement design in The Gambia.

SECTION B – Prepared for publication, but not yet published

Where is the work intended to be published?	American Journal of Clinical Nutrition
Please list the paper's authors in the intended authorship order:	Philip T James, Andrew M Prentice, Matt J Silver
Stage of publication	Not yet submitted.

Multi-authored Work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper.

I set up and managed the clinical trial in The Gambia. I performed the statistical analyses and wrote the first draft of the paper. At this stage two co-authors (Andrew Prentice and Matt Silver) have reviewed the first draft and their comments have been incorporated in the following version. The manuscript will be circulated to a wider team of co-authors prior to submission to AJCN.

Student Signature:

Date: 18th October 2018

Supervisor Signature:

Date: 18th October 2018

RESEARCH PAPER 3

Comparison of a novel drink powder and existing multiple micronutrient tablet (UNIMMAP) in reducing plasma homocysteine in non-pregnant women in rural Gambia: a randomised controlled trial to identify future candidates for epigenetic trials.

Philip T James, Andrew M Prentice & Matt J Silver

Author affiliations
Medical Research Council Unit The Gambia at the London School of Hygiene and Tropical Medicine, London, UK (PTJ, AMP, MJS)

Corresponding author
Philip T James, Medical Research Council Unit The Gambia, London School of Hygiene and Tropical Medicine, Keppel Street, London, WC1E 7HT, UK. Tel: 07971 263776. Email: Philip.James@lshtm.ac.uk

Short running head
Homocysteine reduction in Gambian women.

Abbreviations
AE, adverse event; ANCOVA, analysis of covariance; CI, confidence interval; CV, coefficient of variation; DRI, dietary reference intake; MRCG, Medical Research Unit The Gambia; SAH, S-adenosyl homocysteine; SAM, S-adenosyl methionine; SD, standard deviation; UNIMMAP; United Nations Multiple Micronutrient Preparation; VA, village assistant.

Clinical Trial Registry: Clinicaltrials.gov Reference NCT03431597.

Key words:
Betaine, B vitamins, epigenetics, homocysteine, methyl donors, nutritional supplements, randomised controlled trial, UNIMMAP.

8.1 Introduction

Epigenetic processes describe changes to the genome that can alter gene expression without changing the underlying DNA sequence[1]. One mechanism is DNA methylation of cytosine bases, and there is evidence that this can be influenced by a variety of environmental factors, including maternal nutritional status[2]. Epigenetic modifications therefore provide a plausible mechanism linking early life environmental insults with increased risk of adverse health outcomes throughout the lifecourse[3]. The *in utero* period, especially the very first few days after conception, is a period of exceptionally rapid cell differentiation and complex epigenetic remodelling[4] and represents a window in which epigenetic modifications and errors could have significant consequences for the health of the child[5]. Of the factors that may affect the infant epigenome *in utero* through maternal exposure, nutritional one-carbon-related metabolites are of particular interest[2,6,7]. One carbon metabolism encompasses the metabolic pathways crucial for the provision of methyl groups required for DNA methylation. These pathways involve folate, methionine, serine, glycine, choline and betaine in methyl group donation, and vitamins B2, B12 and B6 as essential coenzymes.

Our previous research in The Gambia has explored the interplay of seasonality, maternal nutritional status and offspring DNA methylation. In the Gambian rainy season women have increased plasma concentrations of key one-carbon metabolites compared to the dry season, corresponding to an increased methylation potential[8]. Indeed, offspring of women conceiving in the peak rainy season demonstrate increased DNA methylation at genomic loci where methylation is established in the very early embryo[9–11]. Furthermore, concentrations of several methyl donor metabolites in mothers predict the offspring's methylation at these loci[9]. Recent work suggests that decreased methylation potential in the dry season may lead to suboptimal outcomes, as evidenced by methylation patterns suggestive of a loss of regular imprinting at the non-coding RNA gene *VTRNA2-1*[12]. These observations raise the question whether a nutritional intervention can shift the maternal metabolome to optimise regulation of the infant epigenome by providing micronutrients in the ratio and quantity necessary for optimal one carbon metabolism all year round, for example by increasing methyl donor provision in the dry season.

As a prequel to possible future studies of a pre-conceptional intervention we conducted a proof-of-concept randomised controlled trial in non-pregnant women with homocysteine as the primary outcome. Homocysteine is strongly inversely associated with the ratio of S-

adenosyl methionine (SAM): S-adensoyl homocysteine (SAH)[8], a measure of methylation potential[13]. Homocysteine can be hydrolysed to SAH, which acts as an allosteric inhibitor of the SAM to SAH (transmethylation) reaction[14,15]. Therefore by reducing dry season concentrations of homocysteine we hope to improve methylation potential.

We compared the effectiveness of a novel drink powder containing betaine and B vitamins against the United Nations Multiple Micronutrient Preparation (UNIMMAP).

8.2 Subjects and methods

8.2.1 Trial outline

We conducted a 3-arm, parallel assignment, unblinded, randomized, controlled trial with 298 women. Non-pregnant, non-lactating, healthy women of reproductive age from the West Kiang region of The Gambia were randomized in a 1:1:1 allocation ratio to 12 weeks of daily supplementation of either a) a novel betaine and B vitamins drink powder supplement ('drink powder'), b) an existing available micronutrient tablet (UNIMMAP) or c) no intervention (control). All participants were invited for midline and endline visits for anthropometric measurements and to provide a venepuncture sample. The trial was registered at Clinicaltrials.gov (NCT03431597). The supplementation phase of the trial took place from March to June 2018, in the Gambian dry season.

Our primary objective was to assess the effect of the drink powder on plasma homocysteine after 12 weeks of daily supplementation versus the control group. Our secondary objectives were to look at this comparison at midline, and to compare the effectiveness of the drink powder versus the UNIMMAP tablet at midline and endline.

8.2.2 Participants

Study participants came from 21 villages of the West Kiang region. A list of potentially eligible women within West Kiang was generated using the Medical Research Council Gambia (MRCG) Keneba Health and Demographic Surveillance System database[16]. At the screening stage a team of ten field workers visited women at their homes to explain the study and eligibility criteria, and to invite them to join the study by providing informed consent.

Women were eligible at the screening stage if they were pre-menopausal aged 18-45 years, non-pregnant to their knowledge, non-lactating (at least 9 months post-partum), had no plan to conceive in the ensuing 3 months, no plans to travel, and with no current illness or chronic

health problems (cardiovascular disease, renal disease, thyroid disease or cancer). Women were excluded if they were currently taking B vitamin or multivitamin supplements or taking medication for prevention of seizures (e.g. Carbamazepine). Women with high blood pressure were excluded only when there was an additional history of stroke or heart attack.

8.2.3 Sample size

Unpublished year-round homocysteine data from a recent trial in the same region[17] indicated that women have a mean (SD) plasma homocysteine concentration of 8.06 µmol/L (±2.5) in the peak dry season of February to April. We wanted to detect a minimum change in plasma homocysteine of 1µmol/L, which would bring it under mean rainy season concentration (7.31µmol/L). This required 99 women per arm to give 80% power (two sample t-test, alpha = 0.05), which we increased to 125 to allow for 20% attrition.

8.2.4 Study design

Women who provided informed consent at the screening stage were invited to come for the baseline visit at MRCG Keneba field station. They provided a urine sample for the study nurses to conduct a rapid pregnancy test, and were then formally enrolled onto the trial if they had a confirmed negative pregnancy test result.

Randomisation

Study participants were randomized to the drink powder, UNIMMAP tablet or control arm in a 1:1:1 ratio according to a computer generated randomization scheme. A list of study IDs was generated in advance by the Keneba database manager. In advance of the baseline study the MRC statistician pre-assigned one of the three trial arms ('drink powder', 'tablet', 'control') to each participant's study ID in a 1:1:1 ratio using block randomisation. A block size of 15 was used. The sub-investigator printed the study ID and group allocation onto a piece of paper and sealed in an envelope. The envelopes were arranged in consecutive order of participant 1 to 375, and kept for the baseline visit. The sealed envelopes were not an attempt to keep the group allocation blind, since this was an unblinded trial. However, they helped ensure only one person received a study ID and corresponding group allocation at a time upon enrolment, overseen by the field coordinator.

Study visits

At the baseline visit all enrolled women had their weight (Seca scale ±0.1kg), height (Seca

stadiometer ± 0.1cm) and blood pressure (Omron) measured. Each measurement was taken in triplicate and the average used. The study nurses collected a fasted venepuncture sample from the participants into 9mL EDTA monovettes, which were kept on ice. The samples were processed by the laboratory in Keneba within one hour of collection. Samples were spun using an Eppendorf 5810R centrifuge at 1800 rcf for 10 minutes at 4ºC to separate the plasma. Plasma aliquots were then immediately stored at -70°C until analysis.

Participants were invited back for a midline visit at week 5 and an endline visit at week 12. The midline visit was programmed at week 5 rather than week 6 to ensure samples were collected before the start of Ramadan. The exact same procedures were repeated as for baseline. If women were pregnant at midline they were given the choice of whether to continue in the study or not. Although supplements were safe for pregnancy the information was recorded to be able to account for any effects of hemodilution at the analysis stage. There were no changes to the originally planned methods after trial commencement.

Design of the novel betaine and B vitamins drink powder

To design the supplement we analysed three separate datasets containing plasma nutrient concentrations from women in West Kiang. We ran multivariable linear regression models to assess which nutrients independently predicted plasma homocysteine. We selected those which consistently demonstrated an inverse relationship with homocysteine as candidates for the supplement. The chosen supplement ingredients were then carefully assessed against existing dosage safety information, doses given in previous trials and current plasma nutrient levels to establish final supplement doses well below upper safety limits.

Supplement composition and supplementation schedule

Table 8.1 details the composition of the two intervention products. The drink powder contained 4g of anhydrous betaine with two times the Dietary Reference Intake (DRI) of folic acid, vitamin B12 and vitamin B2 (manufactured by Kendy Ltd., Bulgaria, sourcing betaine from Danisco, Finland). The UNIMMAP tablet contained 15 micronutrients at the DRI (manufactured by DSM Nutritional Products, South Africa).

Those participants who were randomised to the intervention arms at baseline started their supplementation the next day in their home villages. The supplements were supplied to participants on a daily basis by Village Assistants (VAs). VAs were employees of MRCG living in the villages they were supervising. The participants visited their VA on a daily basis. The

VAs dissolved the drink powder into 200mL of water, and participants were given the option to add sugar to taste in case it was too bitter. UNIMMAP was provided in capsule form, to be taken with water. The VAs observed consumption of the supplements.

Once a week field assistants delivered weekly packs of supplements to the VAs to administer to the participants in their village. The weekly pack included a tally sheet for the VA to record any adverse events experienced by the participants daily. They also included seven small coloured cards that were given to the participants, who were asked to give in a card every time they visited the VA in order to measure compliance. Each week the field assistants would compile a report on the adverse events from the tally sheets and record compliance, measured by counting the number of cards a participant had returned that week and cross-checking with the number of remaining supplements in the pack. After the 12 weeks of supplementation the participants continued to be monitored for a further three weeks to capture any delayed adverse events.

Table 8.1 Composition of interventional products, per daily dose.

Novel drink powder		UNIMMAP tablet	
Ingredient	Amount in supplement	Ingredient	Amount in supplement
Folate as folic acid	800µg	Folic Acid	400µg
Vitamin B12 as cyanocobalamin	5.2µg	Vitamin B12	2.6µg
Riboflavin (vitamin B2) as Riboflavin-5'-phosphate	2.8mg	Riboflavin	1.4mg
Betaine as anhydrous betaine	4g	-	-
-	-	Vitamin A	800µg RE
-	-	Vitamin D	5µg
-	-	Vitamin E	10mg
-	-	Thiamine	1.4mg
-	-	Niacin	18mg
-	-	Vitamin B6	1.9mg
-	-	Vitamin C	70mg
-	-	Zinc	15mg
-	-	Iron	30mg
-	-	Iodine	150µg
-	-	Selenium	65µg
-	-	Copper	2mg

Abbreviations: RE, retinol equivalent; UNIMMAP, Daily United Nations Multiple Micronutrient Preparation supplement.

8.2.5 Laboratory methods

We analysed the concentration of plasma homocysteine in samples from baseline, midline and endline visits at the MRCG Keneba laboratory using the COBAS INTEGRA® 400 Plus analyzer. All samples were analysed in one batch on completion of endline. All samples experienced one cycle of freeze-thaw. Inter-assay coefficients of variation (CVs) were <4.0% and intra-assay CVs were <5.1%.

8.2.6 Blinding

Due to the different characteristics of the drink powder and the tablet, the participants, field assistants and the VAs were not blinded to group allocation. The laboratory technician analysing the plasma homocysteine was blinded to group allocation.

8.2.7 Ethical considerations

The trial received ethics approval from The Gambia Government / MRC Joint Ethics Committee (reference SCC 1575v1.2). All participants provided informed consent by signature if literate, otherwise by thumb print in the presence of an impartial witness if illiterate. The MRCG Clinical Trials Support Office conducted four monitoring trips to ensure the trial was implemented according to the international standards of Good Clinical Practice. All procedures were in accordance with the Helsinki Declaration of 1975 as revised in 1983.

8.2.8 Statistical Analysis

We summarized compliance as a percentage of the total number of cards received for each participant over the trial duration. We summarized adverse events by calculating the proportion of participants that had self-reported any adverse event at any point over the trial duration. We compared baseline medians of continuous variables between the three trial arms using the Kruskal-Wallis test. We used the Wilcoxon rank-sum test to compare medians of percentage compliance between the two intervention arms, and the chi-squared test for proportions of participants reporting morbidities (with the Fisher's exact test for any numerator <5).

We used analysis of covariance (ANCOVA) regression models to assess the effectiveness of the supplements at lowering homocysteine at midline and endline. We first assessed the homocysteine variables for normality using the Shapiro-Wilk test. Variables were right-skewed, but satisfied normality assumptions after log-transformation. We then performed the ANCOVA analyses, consisting of linear regression models comparing the difference in log

homocysteine between two trial arms at midline or endline, adjusted for baseline homocysteine, age and body mass index (BMI). The β coefficient provided an estimate of the ratio of geometric mean homocysteine concentration, which we also chose to present as percentage change in the geometric mean using the formula $(e^{\beta}-1)*100$. We analysed the data using a complete case analysis, and also performed a sensitivity analysis restricted to non-pregnant, fasted, compliant (>80% compliance) individuals. All analyses were performed in Stata 15.1 (StataCorp 2017, Texas, USA).

8.3   Results

The participant flow diagram is shown in **Figure 8.1.** Of the 511 women initially approached for the trial, 298 eligible women were enrolled and randomised to the drink powder, UNIMMAP and control arms. Although only 10 women officially withdrew over the study timeframe (9 before midline and 1 before endline), on the days of the study visits there were many absentees. 104 did not attend the baseline visit and were not enrolled. At midline there were 16 absentees and 45 at endline, despite those in the supplement arms continuing with the intervention in their villages. There were 298, 273 and 243 plasma samples available for homocysteine analysis at baseline, midline and endline respectively. Of these, one (in the UNIMMAP arm at midline) had a homocysteine reading below the detection limit.

Baseline characteristics for the enrolled participants are provided in **Table 8.2**. There was no difference in age, BMI, homocysteine concentration, blood pressure and pulse of participants between trial arms.

The unadjusted homocysteine concentration results by time point are shown in **Table 8.3** and **Figure 8.2**. In the drink powder arm, mean homocysteine concentration decreased by 2.67µmol/L and 2.32µmol/L compared to baseline at midline and endline respectively. A similar, though attenuated, pattern was observed in the UNIMMAP arm, where midline and endline mean homocysteine concentrations were 1.28µmol/L and 1.22µmol/L lower than baseline respectively.

Table 8.4 provides the results of the pairwise comparisons between trial arms at endline and midline, adjusted for age, BMI and baseline homocysteine. Compared to the controls the drink powder reduced mean plasma homocysteine by 24.1% and UNIMMAP reduced it by 16.4% (both p<0.001) after 12 weeks of daily supplementation. Comparing the two intervention arms, the drink powder reduced mean homocysteine by 9.2% more than UNIMMAP (p=0.025).
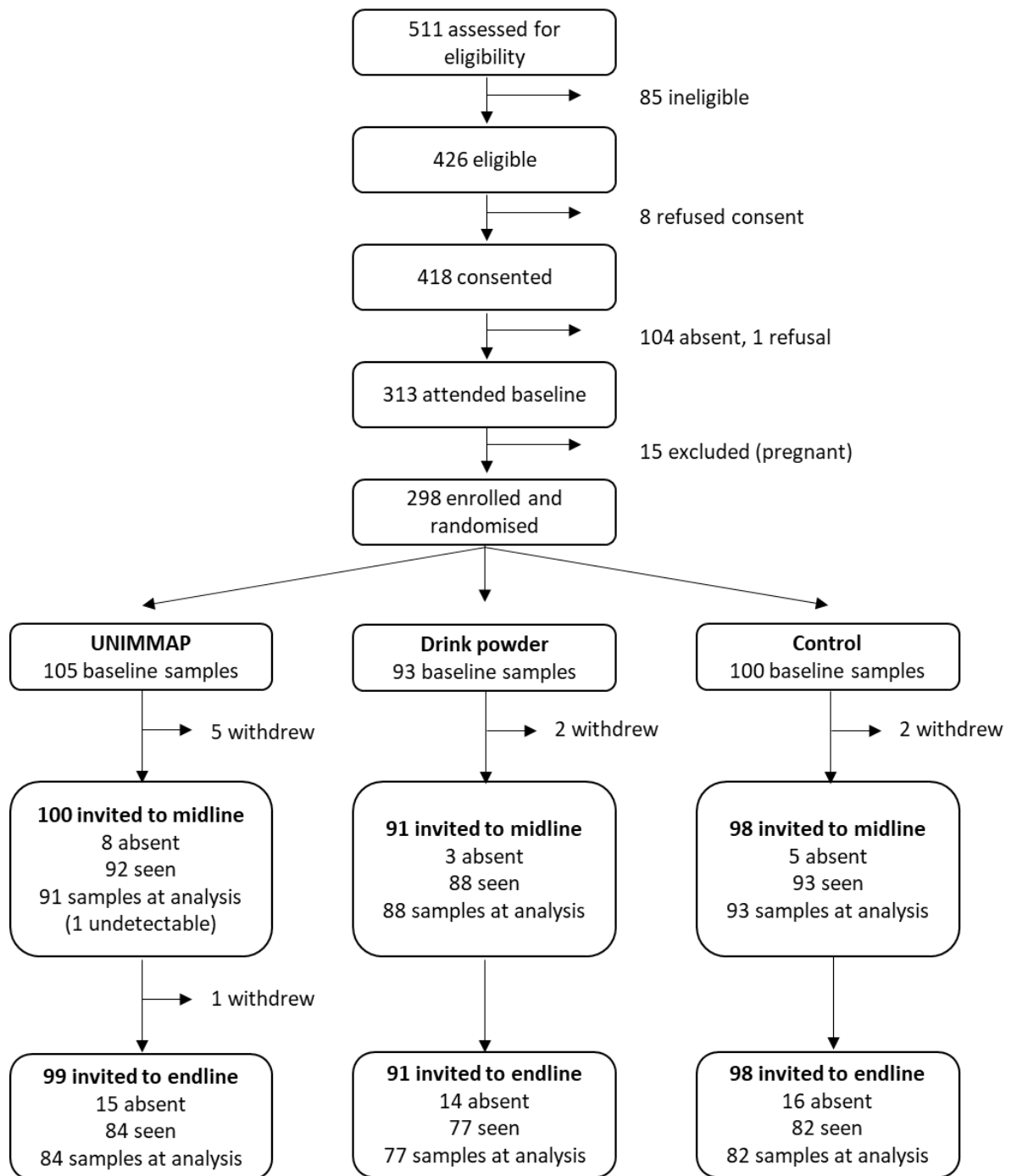
Figure 8.1 Participant flow diagram

Abbreviations: UNIMMAP, Daily United Nations Multiple Micronutrient Preparation supplement.

Table 8.2 Baseline characteristics

Variable	Control (N=100) Median (IQR)	Drink powder (N=93) Median (IQR)	UNIMMAP (N=105) Median (IQR)	P value*
Age (years)	28.5 (22.1, 38.7)	30.7 (21.7, 39.7)	26 (20.2, 38.8)	0.57
BMI (kg/m^2)	21.3 (18.6, 25.3)	21.9 (19.6, 26.3)	21.0 (19.0, 23.5)	0.81
Homocysteine (µmol/L)	9.7 (7.9, 12.3)	10.4 (7.6, 14.1)	10.2 (7.2, 12.7)	0.27
Systolic BP (mmHg)	114 (105, 124)	114 (107, 122)	114 (105, 122)	0.30
Diastolic BP (mmHg)	70 (63, 77)	70 (65, 77)	69 (62, 75)	0.30
Pulse (beats per min)	72 (65, 79)	71 (63, 81)	72 (66, 77)	0.96

* Kruskal-Wallis test for comparing medians

Abbreviations: BMI, body mass index; BP, blood pressure; IQR, interquartile range; UNIMMAP: Daily United Nations Multiple Micronutrient Preparation supplement.

Table 8.3 Unadjusted plasma homocysteine (µmol/L) by time point

		Baseline		Midline		Endline
Study arm	N	Geometric mean (95% CI)	N	Geometric mean (95% CI)	N	Geometric mean (95% CI)
Control	100	10.00 (9.28, 10.78)	93	11.50 (10.62, 12.46)	82	10.18 (9.54, 10.87)
UNIMMAP	105	9.83 (9.10, 10.62)	91	8.55 (8.02, 9.12)	84	8.61 (8.14, 9.12)
Drink powder	93	10.59 (9.67, 11.60)	88	7.90 (7.33, 8.51)	77	8.27 (7.66, 8.92)

Abbreviations: CI, confidence interval; UNIMMAP, Daily United Nations Multiple Micronutrient Preparation supplement.
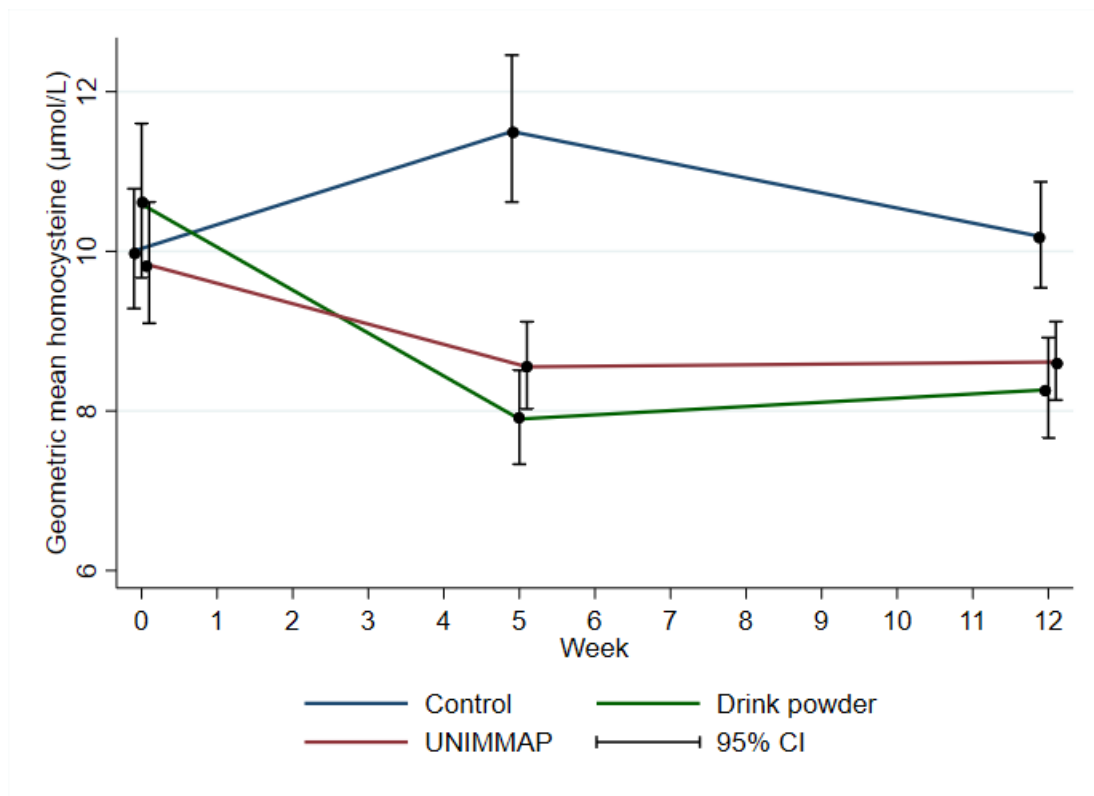
Figure 8.2 Geometric mean homocysteine over time by trial arm

Abbreviations: UNIMMAP, Daily United Nations Multiple Micronutrient Preparation supplement.

Effects at midline were more pronounced, where the drink powder reduced mean plasma homocysteine by 35.8% compared to the controls, and UNIMMAP reduced homocysteine by 22.5%. The drink powder reduced mean homocysteine 15.9% more than UNIMMAP (all p values <0.001).

We performed a sensitivity analysis excluding any participants who became pregnant during the study period, who gave a non-fasted blood sample or had compliance <80%. At midline 8 participants were found to be pregnant and decided to continue with the study (2/93 (2.15%) in control arm, 0 in drink powder arm and 6/92 (6.52%) in UNIMMAP arm). At endline a total of 20 women were pregnant (4/82 (4.88%) in control arm, 7/77 (9.09%) in drink powder arm and 8/84 (9.53%) in UNIMMAP arm). At midline one person reported they had not fasted before venepuncture (in the UNIMMAP arm). At endline 16 participants had not fasted (1 in control arm, 6 in drink powder arm, 9 in UNIMMAP arm). The trial results excluding pregnant, non-fasted and non-compliant individuals are provided in **Supplementary Table 8.1.** Effect sizes and p values are almost identical to those described in the intention-to-treat analysis above.

In this trial supplementation was observed daily and compliance was extremely high in both intervention arms, at 97.8% for UNIMMAP and 98.8% for the drink powder **(Table 8.5).** The compliance was higher than the proportion of women reporting for the endline study visit (Figure 8.1) because some women continued to take their supplements at home even though they did not attend the endline visit. There was no difference between self-reported adverse events (AEs) between the intervention arms. **Table 8.5** describes the proportion of participants who reported an adverse event to their village assistant on at least one day over the 12-week period. The most common AEs described for the two interventions were abdominal pain, nausea and dizziness, followed by urine discoloration and fever. The frequency of these AEs, however, was low. Mean AE duration across the trial was less than half a day per person, and the maximum number of days an individual reported a given AE was 19 days for abdominal pain (additional details in **Supplementary Table 8.2**). There were two serious adverse events reported over the trial duration (one miscarriage in week 9 and one stroke in week 10), both in the drink powder arm. The Trial Safety Monitor and Medical Expert judged both events as highly unlikely to have been caused by the supplement, and both participants resumed participation in the trial after the events.

Table 8.4 Multivariable regression results showing percentage difference in geometric mean plasma homocysteine between trial arms

Trial arms	Time point	N	β (95% CI)*	% difference in geometric means**	P value
Drink powder vs. control	12 weeks	159	-0.27 (-0.35, -0.19)	-23.6 (-29.5, -17.1)	<0.001
UNIMMAP vs. control	12 weeks	166	-0.17 (-0.24, -0.10)	-15.5 (-21.2, -9.4)	<0.001
Drink powder vs. UNIMMAP	12 weeks	161	-0.09 (-0.17, -0.01)	-8.8 (-15.8, -1.2)	0.025
Drink powder vs. control	5 weeks	181	-0.43 (-0.51, -0.36)	-35.1 (-39.7, -30.2)	<0.001
UNIMMAP vs. control	5 weeks	184	-0.26 (-0.33, -0.20)	-23.1 (-28.1, -17.8)	<0.001
Drink powder vs. UNIMMAP	5 weeks	179	-0.15 (-0.22, -0.09)	-14.3 (-20.0, -8.2)	<0.001

* β represents difference in log homocysteine at time point between trial arms, adjusted for baseline homocysteine, age and body mass index at time point.

**Calculated from $(e^\beta-1)*100$

Abbreviations: CI, confidence interval; UNIMMAP, Daily United Nations Multiple Micronutrient Preparation supplement.

Table 8.5 Compliance and self-reported morbidity (condition ever experienced over trial)

	UNIMMAP	Drink Powder	P value*
% Compliance (Median IQR)	97.8 (91.0, 100.0)	98.8 (91.6, 100.0)	0.23
Nausea (N %)	17/104 (16.4)	8/93 (8.6)	0.10
Dizziness (N %)	14/104 (13.5)	12/93 (12.9)	0.91
Urine discolouration (N %)	4/104 (3.9)	8/93 (8.6)	0.23
Abdominal pain (N %)	19/104 (18.3)	13/93 (14.0)	0.41
Fever (N %)	4/104 (3.9)	6/93 (6.5)	0.52

*Wilcoxon rank-sum test of medians for percentage compliance, chi-squared test for morbidity (with Fisher's exact test for any numerator <5)

Abbreviations: UNIMMAP, Daily United Nations Multiple Micronutrient Preparation supplement.

8.4   Discussion

In this trial both supplements were highly effective in reducing homocysteine after twelve weeks of daily supplementation, with the strongest effect sizes apparent after five weeks. There was strong evidence to suggest the drink powder was more effective than UNIMMAP at midline, and that this advantage was maintained, although weaker, at endline. Trial compliance was high and self-reported adverse events were low throughout the trial, and neither varied by intervention arm. The trial confirms an important proof-of-principle that dietary supplements can influence metabolic pathways that we have shown in previous observational studies to predict offspring methylation levels[9,12,18].

Nutritional supplements comprising various combinations of folic acid, B vitamins and methyl donors such as betaine have been successful in lowering plasma homocysteine in multiple trials[19–23]. However, despite a call over ten years ago for more supplementation studies looking at the combined effect of B vitamins with betaine[24], there still remain very few. A meta-analysis of twelve folic acid supplementation trials (with or without additional B vitamins) showed that folic acid supplementation between 1-3 months of supplementation reduced homocysteine by 25% (similar across all doses from 0.5-5mg)[23]. The UNIMMAP multivitamin tablet, which contains one DRI of folic acid, B12, B2 and B6 amongst other ingredients, demonstrated similar effectiveness in our trial.

The drink powder contained double the amount of folic acid, B12 and B2 compared to

UNIMMAP, alongside the addition of betaine, components that may explain its comparative effectiveness. Doubling the B vitamin dosage may not greatly contribute towards increased homocysteine-reduction properties: previous studies suggest that dose of folic acid does not greatly matter[23], that additional B12 also has a disputed effect[23,25], and that B2 supplementation is only effective amongst those with the rare *MTHFR* 677 TT genotype[26,27].

In contrast, in a meta-analysis 4-6g of daily betaine supplementation for 6-12 weeks decreases plasma homocysteine by 1.23 µmol/L, corresponding to a 11.8% reduction of mean baseline values[28]. We can therefore speculate that the betaine in the drink powder has an additive effect compared to the action of folic acid and B vitamins in UNIMMAP. The results also suggest that the vitamin B6 in UNIMMAP is not required to reduce homocysteine in this population, supporting findings from previous trials, including those involving participants with low B6 status[23,29,30]. It appears that tailoring the drink powder design to the nutritional profile of the local population using plasma determinants of homocysteine worked well, and could be a viable approach for future supplement design.

The performance of both interventions surpassed our expectations. Given many of the previous folic acid supplementation trials have been performed in patients with hyperhomocysteinemia, those with increased risk of cardiovascular events and amongst more elderly participants[23], we had envisaged a more modest reduction of around 1µmol/L in our healthy, relatively young population. The heightened effectiveness of the interventions at midline may be a seasonal effect, driven by higher plasma homocysteine in the controls. A previous longitudinal study of one-carbon biomarkers in non-pregnant women in Gambia also showed that maximum plasma homocysteine concentration and minimum folate concentration occurred in late March-early April[8], the timing of our midline. The trial design involving daily observed supplementation, akin to an efficacy rather than effectiveness trial, could also be a factor in the degree of success.

Although the drink powder reduced homocysteine more quickly than UNIMMAP at midline, by endline the differences were less pronounced. We have a particular interest in the potential for a micronutrient supplement to influence the establishment of DNA methylation marks in the very early embryo[9,11]. The sensitive periconceptional period may extend back beyond the three months required for the most active phase of oocyte maturation[6], and there is increasing awareness that interventions should cover longer periods than this to account for unknown timing of conception[31]. Therefore factors such as cost-effectiveness and long-term acceptability become relevant considerations for candidate supplements

alongside speed of action. The high compliance rates in both arms suggests that both interventions were acceptable in the village community settings. The low reported morbidity also reinforces the low-risk nature of these supplements. UNIMMAP has been provided in pre-conception and pregnancy trials in numerous settings[32–40]. Of the trials that reported morbidity the majority found there to be no difference between UNIMMAP and control arms. Given the drink powder arm in our trial reported similar levels of adverse events to the UNIMMAP arm we have assurance that both are well tolerated, further supported by the fact that no participant withdrew from the trial on account of adverse effects.

Lowering homocysteine and providing one-carbon related nutrients are essential for a range of biological processes related to fertility and embryo development, and epigenetic mechanisms are just one of several aspects to consider at periconception[6,41,42]. Future supplement designs may therefore want to consider adding other nutrients essential for optimal fetal development that were beyond the scope of this proof-of-concept trial. Choline, for example, not only acts as a metabolic precursor to betaine, but as an essential component of lipids, lipoproteins and neurotransmitters[43,44]. However, the interrelated nature of metabolic pathways may mean the provision of supplemental folic acid and betaine is sufficient to reduce the demand for choline for transmethylation and therefore spare it for other important functions[45–47]. Further research is required to determine the extent of such compensatory mechanisms and the implication for supplement design.

Our trial had several limitations. We used homocysteine as an integrated indicator of B vitamin status and methylation potential but, had resources allowed, we would have liked to quantify the effects of the two supplements on micronutrient status. The study visits had a number of no-shows due to travel and exam-sitting commitments, particularly at the baseline visit. Unfortunately, we were unable to extend the midline or endline visits due to the need to avoid the month of Ramadan, when venepuncture sample collection was not viable. Whilst this meant we did not achieve our intended sample size of 99 women per arm we were fortunate that the effect sizes of the intervention were much greater than expected. Finally, we did not collect information on genetic polymorphisms that may have affected homocysteine metabolism[48].

In conclusion, both supplements worked to reduce homocysteine concentrations from those seen in the dry season to below those in the rainy season, and therefore met the trial objective. It is premature to speculate whether the improvements on homocysteine reduction offered by the drink powder would be of epigenetic significance in future trials. As

such, they both remain potential candidates for future epigenetic trials in pregnancy in the rural Gambian setting. This is contingent on future research confirming causal links between maternal nutrient exposure, offspring methylation and phenotypic effect, which is an area of continued investigation.

8.5   References

1.   Jaenisch R & Bird A (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.* **33 Suppl**, 245–254.

2.   James P, Sajjadi S, Tomar AS, et al. (2018) Candidate genes linking maternal nutrient exposure to offspring health via DNA methylation: a review of existing evidence in humans with specific focus on one-carbon metabolism. *Int. J. Epidemiol.* dyy153.

3.   Bianco-Miotto T, Craig JM, Gasser YP, et al. (2017) Epigenetics and DOHaD: from basics to birth and beyond. *J. Dev. Orig. Health Dis.* **8**, 513–519.

4.   Seisenberger S, Peat JR, Hore TA, et al. (2013) Reprogramming DNA methylation in the mammalian life cycle: building and breaking epigenetic barriers. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **368**, 20110330.

5.   Fleming TP, Watkins AJ, Velazquez MA, et al. (2018) Origins of lifetime health around the time of conception: causes and consequences. *Lancet* **391**, 1842–1852.

6.   Steegers-Theunissen RPM, Twigt J, Pestinger V, et al. (2013) The periconceptional period, reproduction and long-term health of offspring: the importance of one-carbon metabolism. *Hum. Reprod. Update* **19**, 640–55.

7.   Dominguez-Salas P, Cox SE, Prentice AM, et al. (2012) Maternal nutritional status, C(1) metabolism and offspring DNA methylation: a review of current evidence in human subjects. *Proc. Nutr. Soc.* **71**, 154–65.

8.   Dominguez-Salas P, Moore SE, Cole D, et al. (2013) DNA methylation potential: dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women. *Am. J. Clin. Nutr.* **97**, 1217–27.

9.   Dominguez-Salas P, Moore SE, Baker MS, et al. (2014) Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. *Nat. Commun.* **5**, 3746.

10.   Van Baak T, Coarfa C, Dugué P, et al. (2018) Epigenetic supersimilarity of monozygotic twin pairs. *Genome Biol.* **19,** 2

11.   Kessler NJ, Waterland RA, Prentice AM, et al. (2018) Establishment of environmentally sensitive DNA methylation states in the very early human embryo. *Sci. Adv.* **4**, eaat2624.

12.   Silver MJ, Kessler NJ, Hennig BJ, et al. (2015) Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. *Genome Biol.* **16**, 118.

13.   Mason JB (2003) Biomarkers of Nutrient Exposure and Status in One-Carbon (Methyl) Metabolism. *J. Nutr.* **133**, 941S–947.

14.   Škovierová H, Vidomanová E, Mahmood S, et al. (2016) The Molecular and Cellular Effect of Homocysteine Metabolism Imbalance on Human Health. *Int. J. Mol. Sci.* **17**, 1733

15.   Kerr SJ (1972) Competing methyltransferase systems. *J. Biol. Chem.* **247**, 4248–52.

16.   Hennig BJ, Unger SA, Dondeh BL, et al. (2017) Cohort Profile: The KiangWest Longitudinal Population Study (KWLPS)—a platform for integrated research and health care provision in rural Gambia. *Int. J. Epidemiol.* **46:** e13.

17.   Moore SE, Fulford AJ, Darboe MK, et al. (2012) A randomized trial to investigate the effects of pre-natal and infant nutritional supplementation on infant immune development in rural Gambia: the ENID trial: Early Nutrition and Immune Development. *BMC Pregnancy Childbirth* **12**, 107.

18.   Kühnen P, Handke D, Waterland RA, et al. (2016) Interindividual Variation in DNA Methylation at a Putative POMC Metastable Epiallele Is Associated with Obesity. *Cell Metab.* **24**, 502–9.

19.   Steenge GR, Verhoef P & Katan MB (2003) Betaine supplementation lowers plasma homocysteine in healthy men and women. *J. Nutr.* **133**, 1291–5.

20.   Olthof MR, van Vliet T, Verhoef P, et al. (2005) Effect of homocysteine-lowering nutrients on blood lipids: results from four randomised, placebo-controlled studies in healthy humans. *PLoS Med.* **2**, e135

21.   Dawson SL, Bowe SJ & Crowe TC (2016) A combination of omega-3 fatty acids, folic acid and B-group vitamins is superior at lowering homocysteine than omega-3 alone: A meta-analysis. *Nutr. Res.* **36**, 499–508.

22.   Olthof MR & Verhoef P (2005) Effects of betaine intake on plasma homocysteine concentrations and consequences for health. *Curr. Drug Metab.* **6**, 15–22.

23.   Homocysteine Lowering Trialists' Collaboration. (1998) Lowering blood homocysteine with folic acid based supplements: meta-analysis of randomised trials. *BMJ* **316**, 894–8.

24.   Benevenga NJ (2007) Consideration of betaine and one-carbon sources of N5-methyltetrahydrofolate for use in homocystinuria and neural tube defects. *Am. J. Clin. Nutr.* **85**, 946–9.

25.   Jacques PF, Bostom AG, Wilson PW, et al. (2001) Determinants of plasma total homocysteine concentration in the Framingham Offspring cohort. *Am. J. Clin. Nutr.* **73**, 613–621.

26.   McNulty H, Dowey LRC, Strain JJ, et al. (2006) Riboflavin Lowers Homocysteine in Individuals Homozygous for the *MTHFR* 677C→T Polymorphism. *Circulation* **113**, 74–80.

27.   Hoey L, McNulty H & Strain J (2009) Studies of biomarker responses to intervention with riboflavin: a systematic review. *Am. J. Clin. Nutr.* **89**, 1960S–1980S.

28.   McRae MP (2013) Betaine supplementation decreases plasma homocysteine in healthy adult participants: a meta-analysis. *J. Chiropr. Med.* **12**, 20–25.

29.   Rios-Avila L, Coats B, Ralat M, et al. (2015) Pyridoxine supplementation does not alter in vivo kinetics of one-carbon metabolism but modifies patterns of one-carbon and

tryptophan metabolites in vitamin B-6-insufficient oral contraceptive users. *Am. J. Clin. Nutr.* **102**, 616–25.

30. Lamers Y, Coats B, Ralat M, et al. (2011) Moderate vitamin B-6 restriction does not alter postprandial methionine cycle rates of remethylation, transmethylation, and total transsulfuration but increases the fractional synthesis rate of cystathionine in healthy young men and women. *J. Nutr.* **141**, 835–42.

31. Stephenson J, Heslehurst N, Hall J, et al. (2018) Before the beginning: nutrition and lifestyle in the preconception period and its importance for future health. *Lancet (London, England)* **391**, 1830–1841.

32. Bhutta ZA, Rizvi A, Raza F, et al. (2009) A comparative evaluation of multiple micronutrient and iron-folic acid supplementation during pregnancy in Pakistan: impact on pregnancy outcomes. *Food Nutr. Bull.* **30**, S496-505.

33. Liu J, Mei Z, Ye R, et al. (2013) Micronutrient Supplementation and Pregnancy Outcomes. *JAMA Intern. Med.* **173**, 276.

34. Roberfroid D, Huybregts L, Lanou H, et al. (2008) Effects of maternal multiple micronutrient supplementation on fetal growth: a double-blind randomized controlled trial in rural Burkina Faso. *Am. J. Clin. Nutr.* **88**, 1330–40.

35. Kaestel P, Michaelsen KF, Aaby P, et al. (2005) Effects of prenatal multimicronutrient supplements on birth weight and perinatal mortality: a randomised, controlled trial in Guinea-Bissau. *Eur. J. Clin. Nutr.* **59**, 1081–9.

36. Osrin D, Vaidya A, Shrestha Y, et al. (2005) Effects of antenatal multiple micronutrient supplementation on birthweight and gestational duration in Nepal: double-blind, randomised controlled trial. *Lancet* **365**, 955–62.

37. West KP, Shamim AA, Mehra S, et al. (2014) Effect of maternal multiple micronutrient vs iron-folic acid supplementation on infant mortality and adverse birth outcomes in rural Bangladesh: the JiVitA-3 randomized trial. *JAMA* **312**, 2649–58.

38. Zagré NM, Desplats G, Adou P, et al. (2007) Prenatal multiple micronutrient supplementation has greater impact on birthweight than supplementation with iron and folic acid: a cluster-randomized, double-blind, controlled programmatic study in rural Niger. *Food Nutr. Bull.* **28**, 317–27.

39. Tofail F, Persson LA, El Arifeen S, et al. (2008) Effects of prenatal food and micronutrient supplementation on infant development: a randomized trial from the Maternal and Infant Nutrition Interventions, Matlab (MINIMat) study. *Am. J. Clin. Nutr.* **87**, 704–11.

40. Supplementation with Multiple Micronutrients Intervention Trial (SUMMIT) Study Group, Shankar AH, Jahari AB, et al. (2008) Effect of maternal multiple micronutrient supplementation on fetal loss and infant death in Indonesia: a double-blind cluster-randomised trial. *Lancet* **371**, 215–27.

41. Parisi F, Rousian M, Koning AHJ, et al. (2017) Periconceptional maternal biomarkers of one-carbon metabolism and embryonic growth trajectories: the Rotterdam Periconceptional Cohort (Predict Study). *Fertil. Steril.* **107**, 691–698.

42. Xu J & Sinclair KD (2015) One-carbon metabolism and epigenetic regulation of embryo development. *Reprod. Fertil. Dev.* **27**, 667.

43. Ueland PM (2011) Choline and betaine in health and disease. *J. Inherit. Metab. Dis.* **34**, 3–15.

44. Jiang X, West AA & Caudill MA (2014) Maternal choline supplementation: a nutritional approach for improving offspring health? *Trends Endocrinol. Metab.* **25**, 263–73.

45. Hall MN, Howe CG, Liu X, et al. (2016) Supplementation with Folic Acid, but Not Creatine, Increases Plasma Betaine, Decreases Plasma Dimethylglycine, and Prevents a Decrease in Plasma Choline in Arsenic-Exposed Bangladeshi Adults. *J. Nutr.* **146**, 1062–7.

46. Obeid R (2013) The metabolic burden of methyl donor deficiency with focus on the betaine homocysteine methyltransferase pathway. *Nutrients* **5**, 3481–95.

47. Bertolo RF & McBreairty LE (2013) The nutritional burden of methylation reactions. *Curr. Opin. Clin. Nutr. Metab. Care* **16**, 102–8.

48. Stover PJ (2011) Polymorphisms in 1-carbon metabolism, epigenetics and folate-related pathologies. *J. Nutrigenet. Nutrigenomics* **4**, 293–305.

Supplementary Table 8.1: Multivariable regression results showing percentage difference in geometric mean plasma homocysteine between trial arms, constricted to fasted, non-pregnant, compliant participants

Trial arms	Timepoint	N*	β (95% CI)**	% difference in geometric means Ɨ	P value
Drink powder vs. control	12 weeks	141	-0.28 (-0.36, -0.19)	-24.1 (-30.1, -17.4)	<0.001
UNIMMAP vs. control	12 weeks	143	-0.18 (-0.25, -0.1)	-16.4 (-22.5, -9.9)	<0.001
Drink powder vs. UNIMMAP	12 weeks	130	-0.10 (-0.18, -0.01)	-9.2 (-16.7, -1.1)	0.028
Drink powder vs. control	5 weeks	176	-0.44 (-0.52, -0.37)	-35.8 (-40.4, -30.8)	<0.001
UNIMMAP vs. control	5 weeks	168	-0.26 (-0.33, -0.18)	-22.5 (-27.9, -16.8)	<0.001
Drink powder vs. UNIMMAP	5 weeks	162	-0.17 (-0.25, -0.1)	-15.9 (-21.9, -9.4)	<0.001

* β represents difference in log homocysteine at time point between trial arms, adjusted for baseline homocysteine, age and body mass index at time point.

**Fasted and non-pregnant at time point, and above 80% compliance at time point

Ɨ Calculated from $(e^{\beta}-1)*100$

Supplementary Table 8.2: Additional details on morbidity

Adverse Event	UNIMMAP					Drink Powder				
	N	Mean (days)	Median (days)	Min (days)	Max (days)	N	Mean (days)	Median (days)	Min (days)	Max (days)
Nausea	104	0.26	0	0	5	93	0.21	0	0	7
Dizziness	104	0.37	0	0	10	93	0.30	0	0	6
Urine discolouration	104	0.05	0	0	2	93	0.37	0	0	8
Abdominal pain	104	0.58	0	0	19	93	0.46	0	0	14
Fever	104	0.07	0	0	2	93	0.18	0	0	7

Chapter 9 Discussion

_____

Summary of chapter

In this chapter I summarise the main findings, strengths and weaknesses of the PhD by objective. I then expand on some questions that recurred for me during the PhD period. These include exploring what plasma biomarkers represent and their appropriateness as a proxy measure of the periconceptional nutritional environment. To help with the latter point I draw on research from Assisted Reproductive Technology. I then give an overview of some of the additional exposures we could consider measuring in The Gambia to further our understanding of why seasonality has such an influence on DNA methylation. I finish with suggestions on the way forwards for nutritional epigenetic research in the field of public health nutrition.

9.1 Summary of PhD findings, strengths and weaknesses by objective

9.1.1 *Objective a): A literature review of human studies that investigate associations between maternal periconceptional nutritional exposures, the infant epigenome and later phenotypes.*

There is considerable interest in the extent to which diseases in later life may have been programmed through environmental exposures in very early life, and if so, whether there is the potential for interventions at periconception and in pregnancy to mitigate these effects. Chapter 3 set out the context of nutritional epigenetics within the Developmental Origins of Health and Disease (DOHaD) paradigm. This reviewed what we know to date about how nutritional exposures at periconception and throughout pregnancy influence infant DNA methylation patterns, and in turn whether these epigenetic signatures are further associated with later risk of disease. The literature summarised there painted a complex, heterogeneous picture, showing whilst there is a rapidly building evidence base that maternal nutritional exposures are associated with infant methylation, the patterns are far from consistent. There is also a growing, although often separate, evidence base that DNA methylation marks in infancy, particularly at imprinted genes and metastable epialleles, are associated with adverse phenotypes in the cognitive, growth and cardiometabolic domains. The review highlighted the lack of studies that have been able to link maternal exposure through to offspring phenotype via DNA methylation as a mediating mechanism. Recommendations therefore included more research on the nature of the exposures as well as the confounders, and, where possible, the implementation of randomised trials to help determine causality when linking early exposures to offspring methylation.

Strengths

- The review had a unique focus on identifying research drawing together exposure-methylation-phenotype associations in a single study.
- The genes that featured in the literature searches were summarised with genomic locations, annotations, and coverage on existing commonly used DNA methylation platforms (Illumina 450k / EPIC Arrays). The intention was that this would provide a useful resource for researchers in nutritional and epigenetic epidemiology looking to design a targeted methylation analysis, as well as for those wanting to target specific genes and pathways in future interventions.

- The overview of candidate genes highlighted gene classes of interest to consider alongside imprinted genes, such as metastable epialleles (MEs).

Weaknesses

- Heterogeneity in the type, measurement and timing of nutritional exposures, alongside variation in the genetic loci investigated by study, meant that a systematic literature review and meta-analysis was not possible.

- The review was restricted to DNA methylation and nutrients involved in one-carbon metabolism to allow a feasible scope. Other environmental, including nutritional exposures, were therefore excluded. The nutritional effects on histone modifications and micro RNAs were not investigated, but are important mechanisms to consider in future research.

9.1.2   Objective b): Analyse an expanded set of maternal biomarkers and characterize how the maternal one-carbon metabolome varies across the year

Chapter 4 presented the analyses of the seasonal trends of core one-carbon biomarkers (folate, B2, B6, B12, betaine, choline, homocysteine, dimethylglycine (DMG), methionine, cysteine) and a panel of amino acids in plasma. The dataset contained 350 plasma samples of women in their first trimester of pregnancy, purposively sampled to cover the whole year. Apart from methionine and DMG all the core one-carbon biomarkers and six amino acids demonstrated a seasonal variation in plasma concentrations. The seasonal trends were modelled using Fourier term (FT) linear regression models. The FT models explained between 2.2% and 12.9% of the variation in biomarker concentrations. The coefficient of cyclic variation, a measure of the magnitude of the periodic variation for biomarkers with a seasonal trend, showed that B2 had the greatest seasonal variation (16.2%). This was followed by folate (15.1%) and decreased to approximately 3.5% amongst the amino acids histidine, leucine and serine. Timing of the seasonal peaks and troughs, as well as the amplitude of the seasonal differences, varied greatly between the biomarkers. Co-variation between many of the biomarkers could be explained based on the biochemistry involved in one-carbon metabolism. However, it was not possible to fully explain all the co-variation seen, which could reflect the complexities of interpreting plasma concentrations (explored in section 9.3.1 below).

Strengths

- This dataset contained the first year-round information on one-carbon metabolites in pregnant women in The Gambia. Previous datasets had either been in non-pregnant women, or only captured 6 months of the year.

- The analysis was an ideal opportunity to consolidate prior knowledge from two previous datasets, and expand the focus beyond core one-carbon metabolites. The seasonal trends across the three datasets generally showed a consistent pattern. This confirmed that methylation potential was lower in the dry season, when homocysteine concentrations were higher.

- The dataset was comprised of cross-sectional samples taken from different women across the whole year. This meant the dataset was vulnerable to higher inter-individual variation compared to datasets composed of repeated measures in the same women across the year (e.g. as in the Indicator group dataset). However, the very similar trends with the Indicator group justified the utility of this dataset to investigate seasonal trends.

- The analyses suggested a potential new seasonal window of interest at the end of the rainy season, in September and October, when several biomarkers were at their lowest concentrations.

- The analyses considered a greater number of potential confounders compared to previous studies, being able to adjust the seasonal trends for BMI, inflammation, age and gestational age.

Weaknesses

- The introduction of FTs to model seasonal trends risks oversimplifying seasonal patterns, where not all variation may fit a sinusoidal wave pattern. However, they do enable a standardised way of comparing the amplitude of variation between biomarkers and across datasets.

- The intermediary metabolites S-adenosyl methionine (SAM), S-adenosyl homocysteine (SAH) and formate were planned but not measured due to assay difficulties. These would have provided an additional level of detail on methylation potential using the SAM:SAH ratio, and on an important source of 1-carbons via formate. Future planned studies using SAM:SAH should ensure storage and freeze-thaw cycles are kept to a minimum.

9.1.3 Objective c): Explore whether the profile of maternal nutritional predictors of offspring methylation changes between the rainy and dry season of The Gambia.

Chapter 5 analysed the association of maternal nutrient plasma concentrations in the MDEG cohort, back-extrapolated to the time of conception, with offspring DNA methylation at mean age 3.3 months. In the offspring DNA a new, larger set of genomic loci were investigated. These were 50 CpG loci that were MEs and exhibited differential methylation by season of conception. Unlike previous analyses, in this analysis season was introduced as a potential effect modifier to explore whether the profile of nutritional predictors of methylation differed between the Gambian rainy and dry seasons. There was preliminary evidence of seasonal differences in biomarker-methylation associations for folate, choline and homocysteine. During the dry season plasma biomarkers associated with the folate remethylation pathway predicted DNA methylation. In the rainy season the profile of predictors appeared to switch to a set more aligned to the betaine remethylation pathway. Speculatively this was due to the differences in the underlying nutritional status of women between the seasons. Differences in dietary intake patterns and underlying nutritional status may need to be considered as potential effect modifiers in the relationship between exposures and methylation in all contexts, not just in rural Gambia. The effect of a nutritional intervention, for example, may therefore be dependent on the time of year of implementation and the nutritional status of the population being targeted.

Strengths

- The investigation of season as an effect modifier was a novel approach in this analysis, intended to further our understanding of how seasonality influences DNA methylation through differences in nutrition exposures.
- The use of a larger set of MEs confirmed and strengthened previous findings that offspring conceived in the rainy season have higher levels of ME methylation than offspring conceived in the dry season.
- The analyses used principal components to capture some of the co-variation between biomarkers and their joint influence on methylation. This approach was helpful in considering the role of metabolic pathways rather than considering biomarkers individually.

Weaknesses

- A relatively small sample size, combined with the instability of regression models in variable selection, meant that the findings from this analysis should be viewed as hypothesis-generating, requiring further validation in datasets explicitly powered to test seasonal interactions.

- Linear regression models cannot capture the complexity of metabolism, for example, the non-linear patterns and allosteric regulation of reactions. Models incorporating kinetic data may help interrogate some of the molecular mechanisms underpinning the apparent switch of methylation predictors between the seasons.

- A number of unmeasured seasonal exposures could also influence methylation patterns, and several of these are considered in section 9.4 below.

9.1.4: Objective d): Determine which nutritional components are consistently associated with increased methylation potential in women of rural Gambia

In Chapter 6 three datasets (two cross-sectional and one longitudinal) were used to investigate the nutritional exposures that were consistently inversely associated with plasma homocysteine in the West Kiang population. The purpose was to identify potential ingredients for a supplement using the hypothesis that reducing homocysteine would help improve methylation potential and enable one-carbon metabolic pathways to function unhindered. Three regression approaches were used to obtain a qualitative overview of the patterns of predictors of homocysteine in the different datasets using different variable selection approaches. To assess how generalizable results were from one dataset to another, an important consideration when designing a supplement, the predictive equation for homocysteine in one dataset was applied to the other two datasets to quantify its predictive power.

Folate, B12, B2 and betaine were the most consistent negative predictors of plasma homocysteine. Cysteine, DMG, choline and vitamin B6 were positive predictors of homocysteine. The association of the negative predictors, along with cysteine and DMG, could be explained through an understanding of homocysteine metabolism. The positive associations of B6 and choline with homocysteine were not expected, although similar associations have also been found in other studies. In the two cross-sectional datasets the set of predictors explained over half the variability of plasma homocysteine, and over a third

of the variability in the longitudinal dataset. The predictors performed well in cross-dataset comparisons. Overall the results confirmed that nutritional predictors of homocysteine are generalizable between different cohorts from West Kiang region, and therefore represent robust predictors of homocysteine for supplement design considerations.

Strengths

- The combination of using three datasets, three variable selection approaches and cross-dataset predictive power enabled the identification of robust negative predictors of homocysteine.
- Identifying plasma predictors of homocysteine represented a simple approach to tailor a supplement design to a target population. This approach could be replicated in different contexts.

Weaknesses

- Identifying predictors of plasma homocysteine is useful for supplement ingredient selection but does not provide information on the dose of the component to consider.
- Dietary inputs in the form of a supplement may not necessarily act in the same way as plasma determinants of homocysteine. The complex relationships between dietary intake and plasma concentrations are explored further in section 9.3.1.

9.1.5 Objective e) Use the results from objectives b) to d) to design a novel nutritional supplement to increase methylation potential and to decide which time of year is most suitable for a trial.

Chapter 7 described the design of a novel drink powder supplement using the four negative predictors of plasma homocysteine identified in Chapter 6. The drink powder contained folic acid, B2 and B12 at twice the dietary reference intake (DRI), and betaine at 4g per day. The decision to double the DRI of folic acid and the B vitamins came from the analysis in Chapter 4, which showed that over a quarter of women had a low plasma folate concentration, 39% of women had a low plasma B2 concentration, and 9.4% had a low B12 concentration. The supplement therefore not only had to provide the daily allowances for maintenance of good health, but to correct micronutrient deficiencies as well. Much of Chapter 7 was dedicated to exploring the safety of the drink powder supplement. Previous doses of the supplement ingredients provided in previous trials, particularly pregnancy trials where possible, were

recorded along with summaries of documents assessing safety. Overall it was concluded that the suggested drink powder was a low-risk formulation, extremely unlikely to result in any adverse events when taken at the correct dosage of one sachet per day.

Chapter 5 suggested that the timing of future interventions matters and should be carefully considered in study design. Information from Chapter 4, alongside previous research, helped confirm that the timing of the supplementation trial should be in the dry season. The longitudinal analysis in Chapter 4 suggested that mean plasma homocysteine concentration was higher in the peak dry season (February – April) compared to the peak rainy season (July-September), a pattern also seen in previous studies. A higher homocysteine concentration is associated with a lower methylation potential in the dry season, and indeed, the MRC Gambia research group had previously found methylation patterns suggestive of loss of imprinting at the *VTRNA 2-1* gene in offspring conceived during the dry season[1]. The clinical trial described in Chapter 8 did not intend to look at epigenetic outcomes. However, conducting it in the dry season was important to give the best indication of how the supplement might perform in future epigenetic trials aimed at increasing maternal methylation potential in the dry season and optimising offspring methylation patterns.

Strengths

- The novel drink powder design contains, as far as is known, a unique ingredient and dosage combination designed to lower homocysteine. Betaine, folic acid and B vitamins have been used in previous homocysteine-lowering trials, but have not previously been combined into the same supplement in The Gambia.
- The comprehensive review of previous dosages provided in trials, alongside the review of existing safety information, provided reassurance that the supplement design was safe.

Weaknesses

- Not all previous clinical trials published information on adverse events, therefore it was not possible to completely exclude the possibility of unrecorded side effects.
- There was not as much safety information on betaine compared to folic acid and the B vitamins. Previous betaine trials had been conducted in non-pregnant adults. Therefore manufacturer information for a betaine powder supplement (Cystadane®) was particularly relied upon rather than information from pregnancy trials, which would have been preferable.

9.1.6 Objective f): Conduct a randomized controlled trial to test the effect of the novel supplement on increasing methylation potential via reducing homocysteine.

Chapter 8 described a 3-arm randomised controlled trial conducted amongst 298 non-pregnant, healthy women of rural Gambia. The effectiveness of the novel drink powder in reducing plasma homocysteine was tested in a proof-of-concept trial, against a control arm and an arm receiving the UNIMMAP multiple micronutrient tablet. The rationale was that a supplement reducing homocysteine would increase methylation potential and therefore be a suitable candidate for future epigenetic trials if it was safe and well accepted. Compared to the control arm the drink powder reduced mean plasma homocysteine by 24.1% and UNIMMAP by 16.4% (both $p<0.001$) after 12 weeks of daily supplementation. These effects exceeded expectations. The drink powder was more effective than UNIMMAP, reducing mean homocysteine by a further 9.2% ($p=0.025$), which could be attributed to the addition of betaine.

Both interventions brought dry season homocysteine concentrations to below those found in the rainy season, and therefore met the trial aim. Both supplements therefore remain candidates for future pregnancy trials looking at infant epigenetic outcomes. Only after such trials will it be possible to verify whether the additional effectiveness of the drink powder provides any advantage over UNIMMAP in terms of influence on infant DNA methylation. The drink powder works more quickly and effectively than UNIMMAP to reduce plasma homocysteine. However, UNIMMAP has the advantage of having previously been used in several large pregnancy trials, and may therefore be easier to justify to ethics committees. Also, being a tablet, UNIMMAP may be logistically easier to administer than the drink powder.

Strengths
- This is the first time a nutritional supplement has been designed and tested in The Gambia with the specific aim of acting on metabolic pathways shown in previous observational studies to predict offspring methylation levels.
- The trial results confirmed that using plasma concentration data to tailor the design of a homocysteine-lowering supplement to the target population is an effective strategy.

- Trial compliance was high (>97%) and self-reported adverse events were low throughout the trial, and neither varied by intervention arm. This confirmed that both supplements were safe and well accepted by the participants.

Weaknesses

- The trial would have benefitted from qualitative research investigating participants' attitudes to the two interventions to provide further information on future acceptability and sustainability.

- The Gambian government recently changed regulations so that food supplements were categorised as medicines, and therefore the same regulations as for full drug trials needed to be met. This meant trial set-up took longer than expected and the trial finished at the end of the dry season in June, when plasma homocysteine levels are approaching their lowest concentration (Chapter 4). Running the trial earlier in the dry season, when homocysteine concentrations are higher, may have resulted in an even stronger effect of the interventions.

9.2 Unanswered questions

This final chapter offers space to reflect on some unresolved issues in the hope they might be useful considerations for researchers going forwards. The main recurring questions that surfaced during the thesis included:

- We use plasma nutrient concentrations in the analyses. What exactly do these plasma concentrations represent?

- Is plasma the best proxy we have for the nutritional environment surrounding the oocyte as it matures and the embryo in the first days after fertilization?

- Assisted Reproductive Technologies have been developing pre-implantation embryos *in vitro* for decades. What can we learn from this field about the optimal nutritional environment required at periconception?

- We know that differences in one-carbon metabolites do not fully account for the effect of seasonality on DNA methylation. What are some of the other potential exposures to explore in future research?

- Is there really a need to focus on molecular mechanisms within the field of public health nutrition?

- Should future interventions focus exclusively on epigenetic outcomes? Are we ready for such interventions, and if not, why not?

These questions are addressed in three broad sections below.

9.3 Plasma biomarkers and the search for the optimal nutritional environment for the maturing oocyte and the early embryo

Given many studies, including the ones presented in this thesis, use plasma nutrient concentrations, I first ask how well these represent dietary intake. Practical interventions in the public health nutrition field are likely to involve dietary advice or supplements, and therefore a thorough understanding of the relationship between dietary intake and plasma concentrations is important. I then explore the extent to which plasma nutrient concentrations represent the nutritional environment experienced by the early embryo. I define this in two main stages: firstly, the nutritional content of the follicular fluid in which the oocyte matures before ovulation, and secondly, the nutrients available in the uterine tract environment supporting the embryo in the first few days prior to implantation and gastrulation. Together these topics help consolidate what we know about the periconceptional nutritional environment, the degree to which dietary intake may influence it and what knowledge gaps remain.

9.3.1 The relationship between dietary intake and plasma concentrations

Biomarkers of nutritional status can reflect different time-scales, from the short-term to the long-term. Plasma nutrient concentrations generally represent short-term dietary intake over recent hours[2], or can sometimes reflect intake over previous days if the subject has fasted overnight[3]. Other tissues, such as hair follicles and adipose, can provide an insight about intake in the previous months or years[4,5]. The appropriateness of the tissue sample depends on the extent to which the nutrient is stored there and to what extent any stores reflect nutrient intake. For example, plasma folate is often used as a measure of recent intake but red blood cell folate is more appropriate for an average of the status in the previous 60 days (the average half-life of an erthyrocyte)[2].

For population studies that require samples from many people at multiple time points, accessibility and acceptability become key issues alongside biological appropriateness. Venepunctures and the assessment of nutritional status using either red blood cell or plasma

concentrations therefore represent a relatively feasible option. However, multiple considerations are required before making assertions about how the plasma concentrations reflect dietary intake and nutrient status. Many reviews have addressed some of the issues to consider[3,4,6–8], from which the following points are summarised.

Firstly, the degree to which plasma concentrations and dietary intake can be correlated depends on the accuracy of intake data. There is frequent under-reporting of energy intake in dietary recalls, alongside the difficulties of being able to recall accurate information on meal composition and portion size. Even when directly measuring intake through observed, weighed records, there is the possibility of participants changing their normal eating habits under assessment conditions. Secondly, there are the factors that affect nutrient concentration in food before consumption, such as preparation style, cooking time, light exposure and storage conditions. Thirdly, after ingestion factors affecting nutrient absorption come into play. These include genetic considerations, homeostatic control, nutrient-nutrient interactions, underlying nutritional status, gut transit times, fibre and phytate content, and the influence of microbiota. Furthermore, smoking, alcohol intake, physical activity levels, disease, the interaction of medications and levels of inflammation can also affect nutrient absorption. Finally, plasma concentrations can be affected by laboratory issues such as length of storage, number of freeze-thaw cycles, processing time, potential contamination and differing stability according to the anti-coagulant used. The different laboratory assays chosen for analyses have specific levels of precision, limits of detection and intra-assay, inter-assay and inter-laboratory variation. Considering all the above factors it is unsurprising that using plasma biomarkers to assess nutritional exposures so often generates variable or sometimes counter-intuitive results.

9.3.2 Is plasma a useful proxy for the nutritional environment of the early embryo?

The above section has summarised many of the considerations required in order to relate diet to plasma concentrations. The next steps are to ask where the nutrients come from that sustain methylation processes at periconception, and the extent to which maternal plasma reflects the available nutrient stores. Whilst the vast majority of transmethylation reactions occur in the liver[9,10], we are particularly interested in two phases relevant to periconception: the maturation of the oocyte in the ovarian follicular fluid prior to ovulation (which provides the oocyte nutrient stores) and the uterine tract environment in the first few days post-fertilization prior to implantation and gastrulation. Both the maternal oocyte stores and the nutritional environment of the early embryo are therefore essential to consider for

understanding optimal periconceptional nutrition[11].

Follicular fluid

Follicular fluid obtains its nutrients from plasma[12,13], hence maternal plasma concentration may actually be a very appropriate proxy for the environment in which the oocyte is developing. There is limited information on the exact nature of the follicular fluid. However, homocysteine and folate have been shown to have a similar concentration in both serum and ovarian follicular fluid, whereas methionine concentration is higher in serum than follicular fluid[14]. Lower homocysteine in the follicular fluid is associated with improved oocyte maturity and embryo quality in Assisted Reproductive Therapy[15]. A mouse study tested the effect of an *in vitro* culture low in methyl donors (low methionine, folic acid, B12, choline, B6) on the maturation of ooctytes. The low methyl donor culture led to decreased follicular development and reduced methylation at *MEST* compared to controls, but no difference in methylation at *SNPRN*, *IGF2R* and *H19*[16]. Furthermore, folic acid supplementation in humans does appear to influence follicular fluid homocysteine and folate concentrations[17]. Considering the oocyte goes through a period of intense maturation within the follicular fluid in the 14 weeks prior to ovulation[18], combined with the suggestion that supplies of nutrients in the ooplasm fuel the oocyte post-ovulation and when it is fertilized[18,19], the follicular fluid is of great importance in determining the nutritional environment of the pre-implantation embryo.

Whilst in our Gambian studies we have back-extrapolated the plasma samples of mothers to the time of conception, perhaps this is only capturing part of the story, more appropriate to the environment of the uterine tract at conception. This may not necessarily, however, be reflecting the three months prior to ovulation where the maternal nutritional environment affects the follicular fluid and oocyte cytoplasm composition, influencing the concentration of metabolites required for the initial cell divisions in the pre-implantation embryo[18,19]. The metabolic substrates and enzymes in the early embryo prior to maternal-to-zygotic transition (MZT), being prior to zygote gene activation, have to be provided by the maternal oocyte storage pools[11,20]. In humans the MZT is at the eight-cell stage around 2-3 days post-fertilisation[21]. In future studies it may also therefore be worth assessing the average nutrient concentration of the three months prior to conception to see if this sheds any more light on the relative contribution of the oocyte development phase to DNA methylation patterns in the early embryo.

Uterine tract environment

The extent to which plasma nutrient concentrations reflect uterine tract concentrations is not well defined. The fact that previous Gambian studies using plasma nutrient concentrations back-extrapolated to the time of conception (Chapter 5 and Dominguez-Salas et al. 2014[22]) have been associated with infant DNA patterns suggests that maternal plasma may still be a good enough proxy. Indeed, considering the complexities involved in interpreting plasma concentration data outlined in section 9.3.1, alongside the modelling assumptions involved in back-extrapolation (see Annex 4.2. for more details), we would expect the various imprecisions to tend towards a null result rather than producing biased estimates. Yet the associations between plasma concentrations at conception and infant DNA methylation remain significant and follow expected directions of association. However, formal studies assessing correlations between plasma and uterine environment are very few. For example, one study characterising the amino acid profile in human uterine fluid found that although branched chain amino acids reflected dietary intake, the uterine concentrations did not correlate with serum concentrations[23]. To the best of my knowledge there is no literature investigating the correlations of one-carbon metabolites between human serum and uterine fluid.

Given the dearth of information on this topic it is helpful to turn to findings from Assisted Reproductive Technology, since there have now been decades of research into developing an embryo *in vitro* prior to introducing the pre-implantation embryo into the uterus at the day 3-5 mark.

9.3.3 What can Assisted Reproductive Technologies teach us about the uterine environment?

Over 6 million babies have now been born using practices of Assisted Reproductive Technology (ART)[24]. The culture used to surround the early embryo during *in vitro* fertilization (IVF) can therefore tell us a lot about the nutritional environment required to sustain a viable embryo. From the outset, it should be noted that the IVF cultures do not necessarily represent the optimum environment. Indeed, there is an increasing recognition of the potential aberrant DNA methylation that can occur in ART at the pre-implantation stage. There is accumulating evidence that DNA methylation differences between offspring from ART compared with those naturally conceived affect loci at imprinted genes and MEs[11,19,24,25], with some suggesting this could be due to perturbations in the folate and methionine cycles[19,26]. Recent reviews document how children born using ART have

increased risk of low birthweight, congenital anomalies and conditions related to imprinting disorders compared to those conceived spontaneously[24,27]. There may also be an increased risk of adverse cardio-metabolic phenotypes[28]. However, despite the increased risks, the incidence of these adverse phenotypes is still very low[19,27]. It is debated whether these phenotypes are determined by aberrant DNA methylation dynamics during the *in vitro* window of the first 3-5 days before implantation, or whether they are a reflection of parental characteristics seeking ART (e.g. they may be older, be infertile etc.) or the increased incidence of multiple births associated with ART[19,28]. However, there is preliminary research comparing offspring conceived using ART with those born naturally to sub-fertile women suggesting that that ART does have an increased risk of low birthweight and prematurity, reviewed in Canovas *et al.* (2017)[24].

Given the above, what could be the differences between the ART cultures versus the human uterine tract secretions? We know from recent studies that the type of embryo culture medium is associated with pregnancy and perinatal outcomes[29], although high-quality data is lacking and it is therefore difficult to determine precisely which components are causing the effects[19,30]. IVF cultures were originally modelled on genital tract secretions of various mammals in the early 1970s, reviewed in Ménézo *et al.* 2013[11]. It is difficult to obtain the specific methyl donor content of commercially available culture mediums, though reviews suggest they vary by manufacturer[26,31]. Although some culture components are listed by companies, many are not, and the concentrations are often unspecified due to commercial sensitivity[11].

Human embryonic stem cells (hESC) have been used as models to generate insights into pre-implantation embryo metabolism. A study by Steele et al. (2005) examined which one-carbon metabolic enzymes are expressed by hESC in an attempt to better understand the dynamics of one-carbon pathways in the early embryo[26]. All key one-carbon enzymes were expressed, albeit it with a low expression of betaine homocysteine methyltransferase-2 (BHMT2). When the experiments introduced methotrexate (a folate inhibitor) to the cell culture to deplete folate concentrations the intracellular (cytosolic) concentration of homocysteine increased by 63% and methionine decreased, further emphasising the necessity of folate in the culture. The authors speculated that since the one-carbon metabolic enzymes were expressed in hESC there was a rationale for believing that maternal diet or culture medium could influence methylation potential through one-carbon pathways in the preimplantation embryo.

Since the research described above in the hESC cell lines, more has been discovered about the human pre-implantation embryo directly, building on evidence from bovine oocyte and preimplantation studies[32,33]. Whilst most one-carbon metabolic enzymes do seem to be expressed[34], the transsulfuration pathway (catalysed by cystathionine beta synthase) is absent in the very early human embryo until the maternal to zygotic transition[11,20]. This is why folate and B vitamins are particularly essential for the developing embryo, to avoid build-up of homocysteine in the absence of the transsulfuration pathway (which can decrease methylation potential) by recycling it to methionine[11].

Further research is needed to better define the nutritional environment of the uterine tract and how well plasma concentrations mirror this. It is also not yet clearly defined what the influence of maternal nutrition is during the oocyte development phase (affecting the oocyte intracellular stores) relative to the environment of the uterine tract (*i.e.* the extracellular nutritional resources for the pre-implantation embryo). However, research from population-based intergenerational nutritional epigenetic studies and from the ART field both point to the possibility that perturbations in one-carbon metabolites during the periconceptional period have the potential to influence methylation potential and methylation patterns in the developing embryo. Improved synergy between these fields of research may help to further understanding of what is happening at the molecular level and how these findings might eventually influence population–level interventions.

9.4   Towards a rural Gambian 'exposome'

In our Gambian studies plasma differences in one-carbon metabolites do not fully account for the effect of season on offspring methylation. There remain a number of other unmeasured seasonal exposures that may act as confounders or effect modifiers of the associations we have seen to date. The totality of all the different environmental exposures that can be experienced across the lifecourse has been termed the 'exposome'[35,36]. Studying the exposome in greater detail could help define what other components contribute to the seasonal influence on DNA methylation. One example of ongoing research utilising such an approach is the Human Early-Life Exposome (HELIX) project, a consortium of six European birth cohorts, which seeks to characterize early-life non-genetic exposures from a range of environmental factors and therefore provide a wealth of information on the 'early-life exposome' from conception onwards[37]. The project is ongoing and may prove useful in highlighting candidate exposures to consider in The Gambia.

9.4.1   The role of metabolomics

Within the exposome nutrition exposures are often considered using a metabolomics approach, which investigates the entire set of metabolites present in a given tissue[38]. This approach may help identify metabolic pathways associated with DNA methylation that lie more distal to the core one-carbon biomarkers we have explored to date. We will need to adopt sophisticated ways to analyse metabolomics datasets with their multiple, correlated exposures and possible interactions. Statistical approaches will need to take account of multiple testing yet still have the flexibility to be useful for hypothesis generation. A review of some different statistical techniques is provided by Barrera-Gomez *et al.* (2017)[39]. An important limitation to be aware of in metabolomics is the accuracy and reproducibility of the laboratory assays used. Whilst techniques are constantly improving, there is a possibility for misclassification of some metabolites[40]. In non-targeted approaches measuring hundreds of metabolites it is not possible to adopt the same use of internal and external standards as in a smaller, targeted approach. This is important to consider for Gambian studies where the differences we find in certain nutrient concentrations between the seasons are small and therefore require assays with a high degree of precision.

As previously discussed, it can be challenging to correlate dietary intake data with plasma nutrient concentrations. However, in the long-term we would want to consider nutritional interventions that are food-based in the interests of acceptability and sustainability. There is a developing field of metabolomics research that may help in this regard, using 'multi-metabolite biomarker models', reviewed in Garcia-Aloy *et al.* (2017)[6]. These map profiles of metabolites onto certain foods, and may be able to help identify dietary patterns in a more objective way than traditional questionnaires and recalls. Accurate panels of biomarkers have been developed that mirror the intake of certain food and drink, such as for cocoa, bread, wine and coffee[6]. Metabolomic profiling in 24h urine has been used to differentiate dietary patterns in a controlled feeding setting (distinguishing between four different diets)[41]. There are also preliminary investigations into combining dietary intake data from questionnaires with the biomarker panels to predict certain phenotypes of interest with greater sensitivity[42].

Such approaches are in the early days of development but could offer an interesting strategy for helping to determine whether there are certain foods responsible for the differing one-carbon biomarker concentrations we find in the different Gambian seasons. The development of biomarker profiles for the Gambian diet would be an intensive and

expensive undertaking, so one approach may be to select a shortlist of the foods most likely to differ between the seasons based on the current dietary intake information we have from the Gambia. Dominguez-Salas *et al.* (2013)[43] considered the correlation between intake of individual nutrients and their plasma concentrations but did not extend this to a food group analysis. Some preliminary hypotheses to test include whether the mango season towards the end of the dry season contributes to the peak in plasma folate, and whether betaine concentrations may be associated with timing of Ramadan due to the potentially increased intake of bread when breaking the fast.

9.4.2   Additional exposures to consider

Annex 9.1 provides a summary of the literature on exposures that may be promising candidates for future research in The Gambia. These include the microbiome, additional nutritional exposures such as polyunsaturated fatty acids and vitamin C, oxidative stress, inflammation and infection, toxin exposure and the consideration of paternal influences.

9.5   The future of nutritional epigenetic work in public health nutrition

9.5.1   The search for epigenetically-mediated phenotypic effects

There is a concerted effort to investigate whether the alterations to DNA methylation associated with periconceptional nutrition and other exposures result in tangible phenotypic outcomes. The review in Chapter 3 highlighted some of the challenges in assessing causality in methylation-phenotype associations and the research still required. As an overall motivation to continue the Gambian epigenetic research, despite the difficulty in defining concrete phenotypes at present, it can be helpful to re-visit one of the original inspirations to investigate molecular mechanisms. In rural Gambia there remains a striking, and as yet still unexplained, observation that children born in the rainy season are six times more likely to die between age 15-65 years than those born in the dry season[44,45]. Given the mortality curves diverge in puberty, and that there is no difference in the clinical diagnosis of premature death (the majority are infection-related), it appears these different mortality risks may be due to a programmed effect. Epigenetic mechanisms remain a leading candidate to explore.

Although the current evidence base for exposure-methylation-phenotype studies is sparse, this is changing. There are several ongoing preconception nutrition interventions, reviewed in Barker *et al.* (2018)[46], that will be generating data within the next few years on offspring

phenotypes, with the ability to investigate epigenetic mechanisms. One particular consortium of interest, for example, is the Pregnancy and Childhood Epigenetics (PACE) group. This compromises 39 studies with 29,000 samples to date[47] with a promising breadth of research into phenotypes to come.

Within The Gambia we have ongoing studies that will also contribute to this research gap. The EMPHASIS study, described in Fall *et al.* (2017)[48], is following up children aged 5-9 years who were conceived during preconception nutritional intervention trials in The Gambia and India. The Gambia trial used UNIMMAP whilst the Indian intervention was a food-based supplement. The phenotypes of interest are cognitive, cardiometabolic, growth and body composition outcomes. Another current study is a cohort investigating the observations arising from previous literature on how the *POMC* gene displays differential methylation by season of conception in Gambian infants, and how in separate German case-control studies these methylation differences are associated with obesity at age 11 and in adulthood[49,50]. The cohort study will examine three different life stages: 0-2 years, mid-childhood and adulthood to assess the effects of *POMC* methylation on weight regulation and energy balance over a wide range of ages in a single study. Some of the children now being followed up in mid-childhood (age 5-9 years) were born to mothers included in the MDEG-2 dataset (Chapter 4), hence the data on maternal one-carbon biomarkers and amino acids will be utilized to explore nutritional predictors of *POMC* methylation.

In The Gambia given season of conception affects methylation at a variety of genetic loci, notably MEs, it may be necessary to jointly consider a wide range of potential phenotypes that could be affected (both beneficial and adverse). This is different to the conventional epigenome-wide association study (EWAS) approach where each study tends to focus on a single, well-defined phenotype.

9.5.2   The way forwards: are nutritional epigenetic trials premature?

Intervention trials that target changes in DNA methylation as the primary outcome in intergenerational settings require that certain criteria be met. Firstly, that there is sufficient evidence the maternal nutritional intervention has the potential to affect offspring methylation. Secondly, that these nutrition-associated methylation changes are causally associated with beneficial phenotypic outcomes. Whilst the scientific community is not yet in a position to be able to fully meet these criteria, as discussed in the previous section the

next few years will generate an unprecedented level of evidence to evaluate the options. Existing knowledge gaps include exploring whether increased methylation potential is always beneficial. Increased methylation at one genetic locus may be beneficial whilst at another it could be detrimental. At present there is no blueprint of the methylation patterns that constitute the 'optimum' methylome. As part of this, further research is also required on how DNA methylation influences gene expression, either directly or via interaction with histone modifications and micro RNAs.

Not all public health nutrition practitioners are accustomed to, nor convinced of the need of, studying the molecular as part of informing population-level interventions. Indeed, it is important to step back and reflect on the extent to which we need to know the mechanisms at work to be able to improve health outcomes. There is very strong evidence, for example, that neural tube defects are preventable through preconceptional folic acid, which is now recommended worldwide and has led to mandatory food fortification in many countries[51]. The mechanism for why folic acid is protective is still undefined, although some studies suggest epigenetics could play a role[52,53]. Using this example it could be argued that the knowledge of providing folic acid at preconception is enough, and understanding the underpinning mechanism is now of secondary importance. Unfortunately, similar solutions have not been discovered for all adverse phenotypes of interest. For example, meta-analyses of the effect of multiple micronutrient supplementation during pregnancy (often missing the periconceptional period) on birth outcomes have shown generally disappointing results, with only modest improvements in birthweight and very limited impact on other outcomes such as pre-term births, miscarriages, still births and perinatal mortality[54–57]. Continued investigations into mechanisms that might link maternal nutrition to offspring phenotype are therefore of critical importance to further guide the timing and composition of interventions, and to ensure research funding is directed towards interventions that work.

Whilst ongoing research is undoubtedly required to build the evidence base to safely and effectively target epigenetic outcomes through nutritional interventions, there is still enough evidence to suggest that epigenetic mechanisms be considered as secondary outcome measures for planned preconception trials. This would apply not only to food fortification and supplementation trials, but also to the cash transfer and behavioural change preconception interventions that are being advocated for in a recent conceptual model laid out in *The Lancet* to improve maternal nutrition and infant outcomes[46]. Collecting information on offspring DNA methylation would enable these mechanisms to be further

explored, even if the primary outcomes are targeting other pathways. This recommendation has time and resource implications, and therefore would need to be considered at the project planning stage.

9.6   Conclusion

The research in this thesis has contributed three key components to our existing knowledge of nutritional epigenetics in The Gambia. Firstly, it has further characterised some of the seasonal differences in plasma nutritional biomarkers, validating the findings from previous studies and extending the consideration to biomarkers outside the core one-carbon metabolism subset. Secondly, it has highlighted the importance of considering seasonality and nutrient status as potential effect modifiers for maternal nutritional predictors of infant methylation. Thirdly, it has identified a novel nutritional supplement and the existing UNIMMAP tablet as promising candidates for future epigenetic trials acting on metabolic pathways to increase methylation potential. Whilst the explicit targeting of DNA methylation profiles by nutritional interventions remains premature, accumulating evidence continues to support the focus on epigenetic mechanisms as important considerations in understanding the developmental origins of health and disease.

9.7 References

1. Silver MJ, Kessler NJ, Hennig BJ, et al. (2015) Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. *Genome Biol.* **16**, 118.

2. Gibson R (2005) *Principles of Nutritional Assessment*. 2nd ed. Oxford University Press.

3. Potischman N (2003) Biologic and Methodologic Issues for Nutritional Biomarkers. *J. Nutr.* **133**, 875S–880.

4. Hedrick VE, Dietrich AM, Estabrooks PA, et al. (2012) Dietary biomarkers: advances, limitations and future directions. *Nutr. J.* **11**, 109.

5. Marshall JR (2003) Methodologic and Statistical Considerations Regarding Use of Biomarkers of Nutritional Exposure in Epidemiology. *J. Nutr.* **133**, 881S–887S.

6. Garcia-Aloy M, Rabassa M, Casas-Agustench P, et al. (2017) Novel strategies for improving dietary exposure assessment: Multiple-data fusion is a more accurate measure than the traditional single-biomarker approach. *Trends Food Sci. Technol.* **69**, 220–229.

7. Jenab M, Slimani N, Bictash M, et al. (2009) Biomarkers in nutritional epidemiology: applications, needs and new horizons. *Hum. Genet.* **125**, 507–25.

8. Blanck HM, Bowman BA, Cooper GR, et al. (2003) Laboratory Issues: Use of Nutritional Biomarkers. *J. Nutr.* **133**, 888S–894S.

9. Scotti M, Stella L, Shearer EJ, et al. (2013) Modeling cellular compartmentation in one-carbon metabolism. *Wiley Interdisci. Rev. Syst. Biol. Med.* **5**, 343–365.

10. Mudd SH & Poole JR (1975) Labile methyl balances for normal humans on various dietary regimens. *Metabolism.* **24**, 721–35.

11. Ménézo Y, Lichtblau I & Elder K (2013) New insights into human pre-implantation metabolism in vivo and in vitro. *J. Assist. Reprod. Genet.* **30**, 293–303.

12. Lew R (2018) Natural history of ovarian function including assessment of ovarian reserve and premature ovarian failure. *Best Pract. Res. Clin. Obstet. Gynaecol.* YBEOG1835.

13. Edwards RG (1974) Follicular fluid. *J. Reprod. Fertil.* **37**, 189–219.

14. Steegers-Theunissen RPM, Steegers EAP, Thomas CMG, et al. (1993) Study on the presence of homocysteine in ovarian follicular fluid. *Fertil. Steril.* **60**, 1006–1010.

15. Ocal P, Ersoylu B, Cepni I, et al. (2012) The association between homocysteine in the follicular fluid with embryo quality and pregnancy rate in assisted reproductive techniques. *J. Assist. Reprod. Genet.* **29**, 299–304.

16. Anckaert E, Romero S, Adriaenssens T, et al. (2010) Effects of Low Methyl Donor Levels in Culture Medium During Mouse Follicle Culture on Oocyte Imprinting Establishment1. *Biol. Reprod.* **83**, 377–386.

17. Boxmeer JC, Brouns RM, Lindemans J, et al. (2008) Preconception folic acid treatment affects the microenvironment of the maturing oocyte in humans. *Fertil. Steril.* **89**, 1766–1770.

18. Steegers-Theunissen RPM, Twigt J, Pestinger V, et al. (2013) The periconceptional period, reproduction and long-term health of offspring: the importance of one-carbon metabolism. *Hum. Reprod. Update* **19**, 640–55.

19. Hoeijmakers L, Kempe H & Verschure PJ (2016) Epigenetic imprinting during assisted reproductive technologies: The effect of temporal and cumulative fluctuations in methionine cycling on the DNA methylation state. *Mol. Reprod. Dev.* **83**, 94–107.

20. Menezo YJR, Silvestris E, Dale B, et al. (2016) Oxidative stress and alterations in DNA methylation: two sides of the same coin in reproduction. *Reprod. Biomed. Online* **33**, 668–683.

21. Eckersley-Maslin MA, Alda-Catalinas C & Reik W (2018) Dynamics of the epigenetic landscape during the maternal-to-zygotic transition. *Nat. Rev. Mol. Cell Biol.* **19**, 436–450.

22. Dominguez-Salas P, Moore SE, Baker MS, et al. (2014) Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. *Nat. Commun.* **5**, 3746.

23. Kermack AJ, Finn-Sell S, Cheong YC, et al. (2015) Amino acid composition of human uterine fluid: association with age, lifestyle and gynaecological pathology. *Hum. Reprod.* **30**, 917–924.

24. Canovas S, Ross PJ, Kelsey G, et al. (2017) DNA Methylation in Embryo Development: Epigenetic Impact of ART (Assisted Reproductive Technologies). *BioEssays* **39**, 1700106.

25. Lazaraviciute G, Kauser M, Bhattacharya S, et al. (2014) A systematic review and meta-analysis of DNA methylation levels and imprinting disorders in children conceived by IVF/ICSI compared with children conceived spontaneously. *Hum. Reprod. Update* **20**, 840–852.

26. Steele W, Allegrucci C, Singh R, et al. (2005) Human embryonic stem cell methyl cycle enzyme expression: modelling epigenetic programming in assisted reproduction? *Reprod. Biomed. Online* **10**, 755–766.

27. el Hajj N & Haaf T (2013) Epigenetic disturbances in in vitro cultured gametes and embryos: implications for human assisted reproduction. *Fertil. Steril.* **99**, 632–641.

28. Fleming TP, Watkins AJ, Velazquez MA, et al. (2018) Origins of lifetime health around the time of conception: causes and consequences. *Lancet* **391**, 1842–1852.

29. Kleijkers SHM, Mantikou E, Slappendel E, et al. (2016) Influence of embryo culture medium (G5 and HTF) on pregnancy and perinatal outcome after IVF: a multicenter RCT. *Hum. Reprod.* **31**, 2219–2230.

30. Youssef MM, Mantikou E, van Wely M, et al. (2015) Culture media for human pre-implantation embryos in assisted reproductive technology cycles. *Cochrane Database*

Syst. Rev. **11**, CD007876.

31. Chronopoulou E & Harper JC (2015) IVF culture media: past, present and future. *Hum. Reprod. Update* **21**, 39–55.

32. Kwong WY, Adamiak SJ, Gwynn A, et al. (2010) Endogenous folates and single-carbon metabolism in the ovarian follicle, oocyte and pre-implantation embryo. *Reproduction* **139**, 705–15.

33. Ikeda S, Koyama H, Sugimoto M, et al. (2012) Roles of one-carbon metabolism in preimplantation period--effects on short-term development and long-term programming. *J. Reprod. Dev.* **58**, 38–43.

34. Benkhalifa M, Montjean D, Cohen-Bacrie P, et al. (2010) Imprinting: RNA expression for homocysteine recycling in the human oocyte. *Fertil. Steril.* **93**, 1585–1590.

35. Wild CP (2005) Complementing the genome with an 'exposome': the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol. Biomarkers Prev.* **14**, 1847–50.

36. Robinson O & Vrijheid M (2015) The Pregnancy Exposome. *Curr. Environ. Heal. Reports* **2**, 204–213.

37. Maitre L, de Bont J, Casas M, et al. (2018) Human Early Life Exposome (HELIX) study: a European population-based exposome cohort. *BMJ Open* **8**, e021311.

38. Brennan L & Hu FB (2018) Metabolomics-Based Dietary Biomarkers in Nutritional Epidemiology-Current Status and Future Opportunities. *Mol. Nutr. Food Res.*, 1701064.

39. Barrera-Gómez J, Agier L, Portengen L, et al. (2017) A systematic comparison of statistical methods to detect interactions in exposome-health associations. *Environ. Heal.* **16**, 74.

40. Slama R & Vrijheid M (2015) Some challenges of studies aiming to relate the Exposome to human health. *Occup. Environ. Med.* **72**, 383–4.

41. Garcia-Perez I, Posma JM, Gibson R, et al. (2017) Objective assessment of dietary patterns by use of metabolic phenotyping: a randomised, controlled, crossover trial. *Lancet Diabetes Endocrinol.* **5**, 184–195.

42. Freedman LS, Kipnis V, Schatzkin A, et al. (2010) Can we use biomarkers in combination with self-reports to strengthen the analysis of nutritional epidemiologic studies? *Epidemiol. Perspect. Innov.* **7**, 2.

43. Dominguez-Salas P, Moore SE, Cole D, et al. (2013) DNA methylation potential: dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women. *Am. J. Clin. Nutr.* **97**, 1217–27.

44. Moore SE, Cole TJ, Poskitt EME, et al. (1997) Season of birth predicts mortality in rural Gambia. *Nature* **388**, 434.

45. Moore SE, Cole TJ, Collinson AC, et al. (1999) Prenatal or early postnatal events

predict infectious deaths in young adulthood in rural Africa. *Int. J. Epidemiol.* **28**, 1088–95.

46.   Barker M, Dombrowski SU, Colbourn T, et al. (2018) Intervention strategies to improve nutrition and health behaviours before conception. *Lancet.* **391**, 1853–1864.

47.   Felix JF, Joubert BR, Baccarelli AA, et al. (2018) Cohort Profile: Pregnancy And Childhood Epigenetics (PACE) Consortium. *Int. J. Epidemiol.* **47**, 22–23u.

48.   Chandak GR, Silver MJ, Saffari A, et al. (2017) Protocol for the EMPHASIS study; epigenetic mechanisms linking maternal pre-conceptional nutrition and children's health in India and Sub-Saharan Africa. *BMC Nutr.* **3**, 81.

49.   Kühnen P, Handke D, Waterland RA, et al. (2016) Interindividual Variation in DNA Methylation at a Putative POMC Metastable Epiallele Is Associated with Obesity. *Cell Metab.* **24**, 502–9.

50.   Kuehnen P, Mischke M, Wiegand S, et al. (2012) An Alu element-associated hypermethylation variant of the POMC gene is associated with childhood obesity. *PLoS Genet.* **8**, e1002543.

51.   Czeizel A, Dudás I, Vereczkey A, et al. (2013) Folate Deficiency and Folic Acid Supplementation: The Prevention of Neural-Tube Defects and Congenital Heart Defects. *Nutrients* **5**, 4760–4775.

52.   Steegers-Theunissen RP, Obermann-Borst SA, Kremer D, et al. (2009) Periconceptional maternal folic acid use of 400 microg per day is related to increased methylation of the IGF2 gene in the very young child. *PLoS One* **4**, e7845.

53.   Gonseth S, Roy R, Houseman EA, et al. (2015) Periconceptional folate consumption is associated with neonatal DNA methylation modifications in neural crest regulatory and cancer development genes. *Epigenetics* **10**, 1166–76.

54.   Haider BA & Bhutta ZA (2012) Multiple-micronutrient supplementation for women during pregnancy. *Cochrane Database Syst. Rev.* **11**, CD004905.

55.   Haider BA & Bhutta ZA (2017) Multiple-micronutrient supplementation for women during pregnancy. *Cochrane Database Syst. Rev.* **4,** CD004905

56.   Fall CHD, Fisher DJ, Osmond C, et al. (2009) Multiple Micronutrient Supplementation during Pregnancy in Low-Income Countries: A Meta-Analysis of Effects on Birth Size and Length of Gestation. *Food Nutr. Bull.* **30**, S533–S546.

57.   Bhutta ZA, Das JK, Rizvi A, et al. (2013) Evidence-based interventions for improvement of maternal and child nutrition. *Lancet.* **382**, 452–477.

ANNEXES

ANNEX 1.1 Collaborators and Field team details and contributions

Name	Position & Institution	Role in thesis
Dr Matt Silver	Senior Investigator Scientist, MRC Unit The Gambia at LSHTM	PhD supervisor, provided bioinformatics data for Chapter 5, co-author for papers in Chapters 3, 5 & 8, commented on drafts of all chapters.
Prof Andrew Prentice	Nutrition Theme Leader, MRC Unit The Gambia at LSHTM	Associate supervisor, line manager, co-author for papers in Chapters 3, 5 & 8, reviewed planned supplement composition for trial.
Prof Patrick Stover	Director of the Division of Nutritional Sciences, College of Human Ecology, Cornell University, Ithaca, USA.	Advisory Committee Member, reviewed planned supplement composition for trial.
Dr Sophie Moore	Reader in Global Women's and Children's Health, King's College London, London, UK	Advisory Committee Member, reviewed planned supplement composition for trial, co-author for paper in Chapter 5.
Dr Martha Field	Senior Research Associate, Division of Nutritional Sciences, College of Human Ecology, Cornell University, Ithaca, USA.	Reviewed planned supplement composition for trial.
Roger Dyer	Research Laboratory Manager, Child & Family Research Institute, University of British Columbia, Vancouver, Canada.	Lead collaborator for biomarker analyses and primary technician for HPLC-MS platform optimisation in Chapter 4.
Prof Rajavel Elango	Assistant Professor, Department of Pediatrics, University of British Columbia, Vancouver, Canada.	Amino acid analyses for Chapter 4.
Dr Klaus Kraemer	Managing Director, Sight and Life, Kaiseraugst, Switzerland	Inspiration for Sight and Light article and book chapter (Annex 3), introduction to Kendy Ltd. and facilitator for in-kind contribution of UNIMMAP tablets for trial.
Antonia Ignatova	International Department, Kendy Ltd, Bankya, Bulgaria	Management of supplement development for trial.
Kabiru Ceesay	Field Coordinator, Keneba field station, MRC Unit The	Trial staff recruitment and supervision oversight for trial.

Name	Position & Institution	Role in thesis
	Gambia at LSHTM	
Ousubie Jawla	Field Supervisor, Keneba field station, MRC Unit The Gambia at LSHTM	Direct supervision of all field workers and logistics oversight for trial.
Ansumana Kombo	Clinical Trials Assistant, Keneba field station, MRC Unit The Gambia at LSHTM	Assistance with maintenance of clinical trial paperwork.
Lamin Jarju, Isatou Jammeh, Fakebba Camara, Lamin F Ceesay, Musa Camara, Alhagie Mbenga, Lamin Camara, Fatou Gitteh, Sutay Njai,	Field assistants, Keneba field station, MRC Unit The Gambia at LSHTM	Overseeing field activities in West Kiang through recruitment of participants and supervisory visits to Village Assistants, trial logistics, and help with anthropometry at trial visits.
Hawa Darboe Momodou Jallow Yerro Sowe	Research nurses, Keneba field station, MRC Unit The Gambia at LSHTM	Pregnancy test and venepuncture oversight at trial visits.
Dr Fatai Akemokwe	Head of Clinical Services, Keneba field station, MRC Unit The Gambia at LSHTM	Medical Expert for trial, helping assess serious adverse events and medical database support
Dr Karen Forrest	Head of Clinical Services, MRC Unit The Gambia at LSHTM	Safety Monitor for trial, helping assess serious adverse events
Ebrima Sise	Analytical Project Manager, Keneba field station, MRC Unit The Gambia at LSHTM	Cobas Integra technician for AGP in Chapter 4 and homocysteine in Chapter 8.
Tolu Osunnuyi, Ebrima Danso, Patrick Okot	Laboratory managers, Keneba field station, MRC Unit The Gambia at LSHTM	Laboratory management for Keneba analyses.
Alasana Saidykhan, Ebrima Danso	Laboratory technicians, Keneba field station, MRC Unit The Gambia at LSHTM	Plasma separation and aliquot preparation for trial samples.
Amulai Touray	Nutrition Theme Projects Coordinator, MRC Unit The Gambia at LSHTM	Support with trial recruitment and budget management.
Bakary Sonko	Database Manager, Keneba field station, MRC Unit The	Database design and maintenance for trial.

Name	Position & Institution	Role in thesis
	Gambia at LSHTM	
Dr Nuredin Ibrahim Mohammed	Statistician, MRC Unit The Gambia at LSHTM	Performed participant randomisation code for trial, approved trial statistical analysis plan.
Edrisa Sinjanka	Research Nurse Coordinator, Keneba field station, MRC Unit The Gambia at LSHTM	Medical training and nurse supervision for trial.
Musa Jarjou	Data entry manager	Coordination of all data entry activities for trial.
Natoma Jarra , Alimatou Jatta, Amie Demba, Lamin Jorbateh	Data entry clerks, Keneba field station, MRC Unit The Gambia at LSHTM	Data entry for trial.
Abdoulie Faal	Database technician, Keneba field station, MRC Unit The Gambia at LSHTM	Help with maintenance of clinical database for trial.
Abubaccar Kandeh	Store Manager, Keneba field station, MRC Unit The Gambia at LSHTM	Oversight of interventional products and equipment for trial.
Ebrima Jallow, Bakary Jatta, Baba Bajo	Drivers, Keneba field station, MRC Unit The Gambia at LSHTM	Transportation of participants for trial visits and medical referrals.
Vivat Thomas	Coordinator of the Clinical Trials Support Office, MRC Unit The Gambia at LSHTM	Guidance for Trial registration, liaising with Medicines Control Agency, advice with compliance with Good Clinical Practice.
Gibbi Sey, Samba Baldeh	Trial monitors, Clinical Trials Support Office, MRC Unit The Gambia at LSHTM	Monitoring of trial to ensure Good Clinical Practice.
Dr Branwen Hennig	Senior Portfolio Lead, Population Health, Science Division, Wellcome Trust, London, UK	Co-author on Chapter 5, supervisor for original MDEG studies.
Dr Paula Dominguez-Salas	Department of Production and Population Health, Royal Veterinary College, London, UK	Co-author on Chapter 5, lead for original MDEG studies.
Sara Sajjadi, Ashutosh Singh Tomar, Ayden Saffari, Caroline H D	Various – all part of EMPHASIS collaboration between MRC Unit The Gambia; Genomic Research on Complex Diseases (GRC	Co-authors on Chapter 3.

Name	Position & Institution	Role in thesis
Fall, Smeeta Shrestha, Prachand Issarapu, Dilip Kumar Yadav, Lovejeet Kaur, Karen Lillycrop, Giriraj R Chandak	Group), CSIR-Centre for Cellular and Molecular Biology, Hyderabad, India; MRC Life course Epidemiology Unit, University of Southampton, Southampton General Hospital, Southampton, UK.	
Dr Ines Thiele	LCSB Associate Professor, Luxembourg Centre for Systems Biomedicine, Luxembourg University, Luxembourg.	Lead collaborator for *in silico* modelling, Annex 4.1.
Dr Maike Aurich	LCSB Research Associate, Luxembourg Centre for Systems Biomedicine, Luxembourg University, Luxembourg.	Lead collaborator for *in silico* modelling, Annex 4.1.

ANNEX 1.2 PhD Timeline

	Oct-15	Jan-16	Apr-16	Jul-16	Oct-16	Jan-17	Apr-17	Jul-17	Oct-17	Jan-18	Apr-18	Jul-18	Oct-18
Academic Year		2015-16				2016-17				2017-18			2018
1 Upgrading													
Upgrading report writing			█										
Upgrading presentation				█									
2 Literature reviews													
Book chapter	█	█											
Narrative review						█	█						
3 Phase 1: Plasma biomarkers & modelling													
UBC visit		█											
UBC analysis			█	█	█								
Fourier series modelling						█							
Predictors of homocysteine modelling					█	█							
Predictors of methylation data analysis					█	█							
Predictors of methylation paper write-up									█				
4 Phase 2: Trial intervention													
Identify supplier for supplement					█								
Design supplements							█	█					
Register trial, SCC ethics								█	█	█			
Database design, form pilot testing								█					
Team recruitment & training									█	█			
Recruitment of subjects										█			
Intervention										█	█		
Lab analysis of plasma homocysteine												█	
Write up trial results												█	
5 Write Up													
Thesis writing													█

ANNEX 3.1 Supplementary Material 1:   Additional information on the role of nutritional factors within one-carbon metabolism

An overview of one-carbon metabolism

S-adenosyl methionine (SAM) is the methyl donor for hundreds of transmethylation reactions[1], and its importance in epigenetics stems from its ability to methylate cytosine bases and histones. The loss of a methyl group from SAM forms S-adenosyl homocysteine (SAH), a build-up of which can reduce methylation rates through allosteric inhibition[2]. The SAM:SAH ratio can be used as a proxy indicator of methylation potential[3]. SAH is further hydrolysed to homocysteine (Hcy)[4] and, since this reaction is reversible, Hcy needs to be removed from the system to maintain favourable methylation conditions. Hcy can be methylated to form methionine or be irreversibly degraded in the transsulfuration pathway. In the case of methionine formation, Hcy can either receive the methyl group from N^5-methyl tetrahydrofolate ('methyl-THF'), produced through the reduction of dietary folates and folic acid, or from betaine in the liver and kidneys, a product formed through the oxidation of choline. Methionine can then be condensed with ATP to form SAM and complete the cycle. Genetic variants in one-carbon enzymes can also affect the flow of metabolites through these pathways[5].

The role of individual nutrients

Folate: Dietary folates and folic acid are reduced to form tetrahydrofolate (THF), which is a scaffold upon which one-carbon units can be attached and activated. THF in turn is reduced to methylene-THF, then to 5-methyl-THF, which donates its methyl group to homocysteine using vitamin B12[6].

B vitamins: Several B vitamins act as cofactors in one-carbon pathways (B12, B2, B6). Vitamin B12 is a cofactor in the methylation of homocysteine by 5-methyl-THF. Vitamin B6 (in the active form of pyridoxal-5'-phosphate; PLP) is required to reduce THF to methylene-THF, and is also a cofactor in the transsulfuration pathway converting homocysteine to cysteine[7]. B2 is required as a precursor to FAD, which is a cofactor for MTHFR to reduce methylene-THF to methyl-THF[3]. The different forms of THF are interconvertible, except the methyl-THF form. In the absence of sufficient B12 to utilise the methyl group from methyl-THF in homocysteine methylation (a process which re-generates THF), methyl-THF can accumulate at the expense of other THF forms. This is termed the 'folate trap' or 'methyl trap' and means pathways dependent on other forms of THF (e.g. purine synthesis and the thymidylate pathway) can

become impaired[8].

Choline can be synthesised endogenously or obtained from the diet (good sources include red meat, poultry, milk, eggs and fish[9]). A two-step oxidation reaction converts it to **betaine**[10], which can donate its methyl group to homocysteine via betaine-homocysteine methyl transferase (BHMT)[10]. Betaine can also be directly sourced from the diet (e.g. wheat bran, wheat germ, spinach, beets[9]).

Polyunsaturated fatty acids (PUFAs) influence the 1-carbon pathway through two mechanisms. Firstly, ω-3 PUFAs upregulate enzymes responsible for the methylation of homocysteine to methionine[11,12]. Secondly, the availability of PUFAs can influence methyl balance via phosphatidylcholine (PC). PC is used in the production of very low density lipoproteins (VLDLPs), bile and surfactants. It is formed via two main pathways: the cytidine diphosphocholine (CDP)–choline ('Kennedy') pathway and the PEMT pathway, which converts phosphatidylethanolamine (PE) to PC (Fig. 1). PC differs in composition according to the pathway of its formation; the former incorporates more saturated fatty acids and the latter more PUFAs, such as arachidonic acid (AA) and docosahexaenoic acid (DHA)[13]. In the PEMT pathway the conversion of a molecule of PE to PC requires three methyl groups from SAM. In rodent studies imbalances in methyl donor supply and the coenzymes involved in 1-carbon metabolism affect PEMT activity, manifested in disturbances to tissue levels of PUFAs[12,14], since the PEMT pathway is the major way phospholipids are transported from the liver to the plasma and other tissues[15,16]. These effects have been most pronounced in pregnant dams consuming a diets designed to resemble a rural Indian intake, with excess folic acid and restricted vitamin B12 lowering offspring placental and brain DHA, as well as lowering placental global methylation[17,18]. In these experiments, supplementation with ω-3 PUFAs reversed the effect of the micronutrient imbalances[12]. The PEMT pathway represents a major demand for methyl groups[19,20]. Hence any factor that reduces the demand for methyl groups in the PEMT pathway (e.g. low PUFA intake) could in theory make more methyl groups available for DNA methylation, with potential consequences for epigenetic outcomes[21].
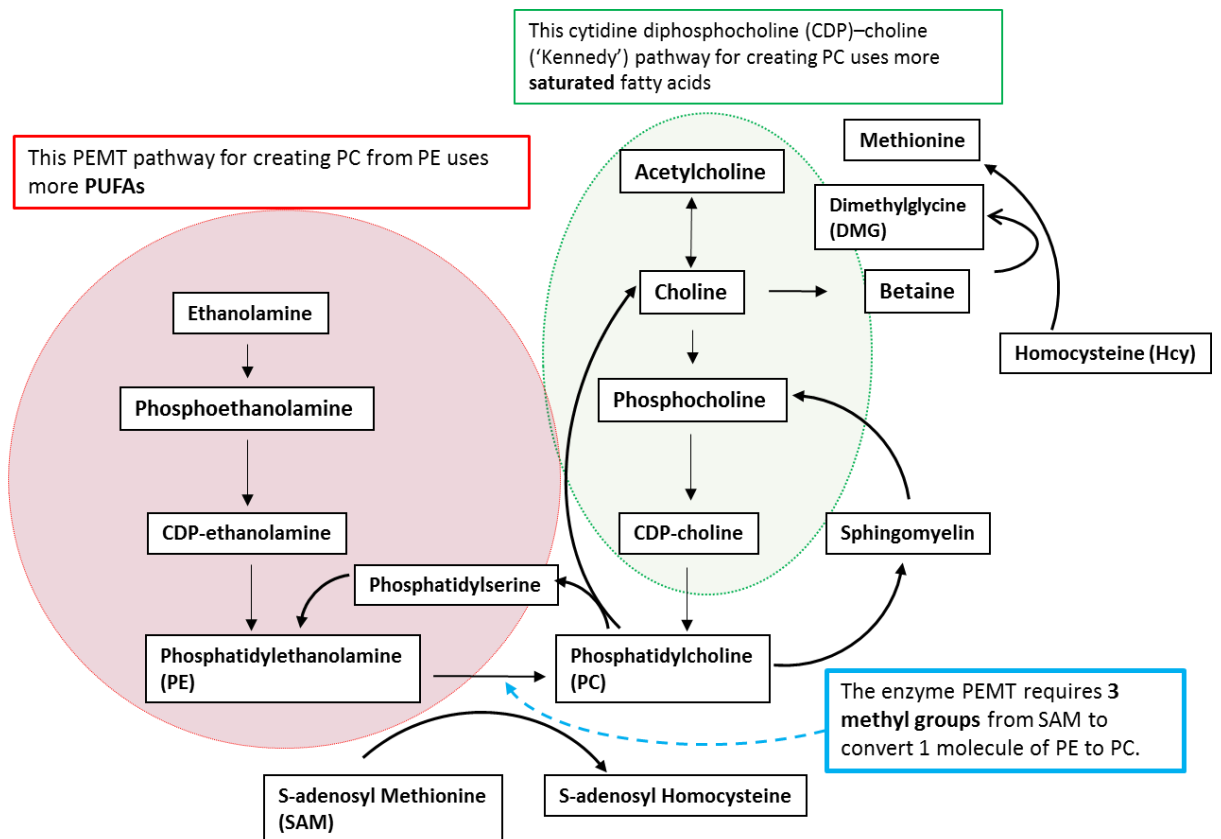
Fig 1: Formation of phosphatidylcholine by two metabolic pathways

Abbreviations: CDP, cytidine diphosphocholine; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PEMT, phosphatidylethanolamine N-methyltransferase; PUFA, polyunsaturated fatty acids; SAM, S-adenosyl methionine.

Amino acids: Several amino acids are thought to contribute 1-carbon units[22], and in turn deficiencies in 1-carbon-related analytes may affect amino acid metabolism[23–25]. The most documented amino acids in this context are serine and glycine. Serine donates a 1-carbon unit at the stage of converting THF to methylene-THF through the action of Serine Hydroxymethyltransferase (SHMT) and PLP[26]. In the process serine is converted to glycine. Serine is also involved in the transsulfuration pathway at the steps of converting homocysteine to cystathionine and then to cysteine. Glycine is not simply a by-product of reactions involving serine, however. For example, it accepts methyl groups from SAM via GNMT, especially when SAM concentrations are high[27]. Its catabolism by the glycine cleavage system in the mitochondria generates methylene-THF with the involvement of THF, which is then used as a carbon donor in the 1C pathways[28]. The catabolism of amino acids tryptophan and histidine is of also of particular interest within 1-carbon metabolism as this produces formate[22,29]. Formate is produced in the mitochondria and then released into the

cytosol, forming formyl-THF by condensation with THF. Formyl-THF can then either be used in purine synthesis or be interconverted into other THF oxidation states as part of folate metabolism. Taken together, formate, serine and glycine help link the mitochondrial and cytosolic folate pools[30].

Methionine: Methionine is obtained from the diet or formed by the methylation of homocysteine. Its adenylation forms SAM, making it a key metabolite in the 1-carbon pathways. Methionine also upregulates glycine N-methyltransferase (GNMT) which can use excess SAM to methylate glycine to sarcosine.

Homocysteine: Homocysteine is an intermediary metabolite that is hydrolysed from SAH. Since a build-up of SAH inhibits transmethylation reactions[2] homocysteine is often inversely associated with methylation. For example, van Mil *et al.* (2014) found an increase in maternal plasma homocysteine was associated with 0.04% lower methylation at *NR3C1* in the infant epigenome (p=0.03)[31]. Although it is not a nutritional exposure *per se*, nutritional inputs can influence the concentration of homocysteine. Trials that have aimed to decrease homocysteine as part of lowering blood pressure could therefore provide additional information as to exposures with the potential to influence the epigenome. For example, several studies indicate that increased consumption of ω-3 PUFAs is associated with reduced homocysteine levels[32], and that when combined with folic acid and other B vitamins the association is even stronger[33]. Betaine supplementation has also been shown to reduce homocysteine levels[27]

References

1. Lu SC & Mato JM (2012) S-adenosylmethionine in liver health, injury, and cancer. *Physiol. Rev.* **92**, 1515–42.

2. Scotti M, Stella L, Shearer EJ, et al. (2013) Modeling cellular compartmentation in one-carbon metabolism. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **5**, 343–365.

3. Mason JB (2003) Biomarkers of Nutrient Exposure and Status in One-Carbon (Methyl) Metabolism. *J. Nutr.* **133**, 941S–947.

4. Pajares MA & Pérez-Sala D (2006) Betaine homocysteine S-methyltransferase: just a regulator of homocysteine metabolism? *Cell. Mol. Life Sci.* **63**, 2792–803.

5. Fredriksen A, Meyer K, Ueland PM, et al. (2007) Large-scale population-based metabolic phenotyping of thirteen genetic polymorphisms related to one-carbon metabolism. *Hum. Mutat.* **28**, 856–65.

6. Scott JM & Weir DG (1998) Folic acid, homocysteine and one-carbon metabolism: a review of the essential biochemistry. *J. Cardiovasc. Risk* **5**, 223–7.

7. Selhub J (1999) Homocysteine metabolism. *Annu. Rev. Nutr.* **19**, 217–46.

8. Finkelstein JL, Layden AJ & Stover PJ (2015) Vitamin B-12 and Perinatal Health. *Adv. Nutr.* **6**, 552–63.

9. Cho E, Zeisel SH, Jacques P, et al. (2006) Dietary choline and betaine assessed by food-frequency questionnaire in relation to plasma total homocysteine concentration in the Framingham Offspring Study. *Am. J. Clin. Nutr.* **83**, 905–11.

10. Ueland PM (2011) Choline and betaine in health and disease. *J. Inherit. Metab. Dis.* **34**, 3–15.

11. Huang T, Hu X, Khan N, et al. (2013) Effect of Polyunsaturated Fatty Acids on Homocysteine Metabolism through Regulating the Gene Expressions Involved in Methionine Metabolism. *Sci. World J.* **2013**, 1–8.

12. Khot V, Kale A, Joshi A, et al. (2014) Expression of genes encoding enzymes involved in the one carbon cycle in rat placenta is determined by maternal micronutrients (folic acid, vitamin B12) and omega-3 fatty acids. *Biomed Res. Int.* **2014**, 613078.

13. DeLong CJ, Shen YJ, Thomas MJ, et al. (1999) Molecular distinction of phosphatidylcholine synthesis between the CDP-choline pathway and phosphatidylethanolamine methylation pathway. *J. Biol. Chem.* **274**, 29683–8.

14. van Wijk N, Watkins CJ, Hageman RJJ, et al. (2012) Combined dietary folate, vitamin B-12, and vitamin B-6 intake influences plasma docosahexaenoic acid concentration in rats. *Nutr. Metab.* **9**, 49.

15. Watkins SM, Zhu X & Zeisel SH (2003) Phosphatidylethanolamine-N-methyltransferase activity and dietary choline regulate liver-plasma lipid flux and essential fatty acid metabolism in mice. *J. Nutr.* **133**, 3386–91.

16. Selley ML (2007) A metabolic link between S-adenosylhomocysteine and polyunsaturated fatty acid metabolism in Alzheimer's disease. *Neurobiol. Aging* **28**, 1834–1839.

17. Kulkarni A, Dangat K, Kale A, et al. (2011) Effects of altered maternal folic acid, vitamin B12 and docosahexaenoic acid on placental global DNA methylation patterns in

Wistar rats. *PLoS One* **6**, e17706.

18. Roy S, Kale A, Dangat K, et al. (2012) Maternal micronutrients (folic acid and vitamin B(12)) and omega 3 fatty acids: implications for neurodevelopmental risk in the rat offspring. *Brain Dev.* **34**, 64–71.

19. Stead LM, Brosnan JT, Brosnan ME, et al. (2006) Is it time to reevaluate methyl balance in humans? *Am. J. Clin. Nutr.* **83**, 5–10.

20. Bertolo RF & McBreairty LE (2013) The nutritional burden of methylation reactions. *Curr. Opin. Clin. Nutr. Metab. Care* **16**, 102–8.

21. Khot V, Chavan-Gautam P & Joshi S (2015) Proposing interactions between maternal phospholipids and the one carbon cycle: A novel mechanism influencing the risk for cardiovascular diseases in the offspring in later life. *Life Sci.* **129**, 16–21.

22. Brosnan ME, MacMillan L, Stevens JR, et al. (2015) Division of labour: how does folate metabolism partition between one-carbon metabolism and amino acid oxidation? *Biochem. J.* **472**, 135–46.

23. Rios-Avila L, Coats B, Chi Y-Y, et al. (2015) Metabolite profile analysis reveals association of vitamin B-6 with metabolites related to one-carbon metabolism and tryptophan catabolism but not with biomarkers of inflammation in oral contraceptive users and reveals the effects of oral contraceptives o. *J. Nutr.* **145**, 87–95.

24. Rios-Avila L, Coats B, Ralat M, et al. (2015) Pyridoxine supplementation does not alter in vivo kinetics of one-carbon metabolism but modifies patterns of one-carbon and tryptophan metabolites in vitamin B-6-insufficient oral contraceptive users. *Am. J. Clin. Nutr.* **102**, 616–25.

25. da Silva VR, Rios-Avila L, Lamers Y, et al. (2013) Metabolite profile analysis reveals functional effects of 28-day vitamin B-6 restriction on one-carbon metabolism and tryptophan catabolic pathways in healthy men and women. *J. Nutr.* **143**, 1719–27.

26. Davis SR, Stacpoole PW, Williamson J, et al. (2004) Tracer-derived total and folate-dependent homocysteine remethylation and synthesis rates in humans indicate that serine is the main one-carbon donor. *Am. J. Physiol. Endocrinol. Metab.* **286**, E272-9.

27. Olthof MR & Verhoef P (2005) Effects of betaine intake on plasma homocysteine concentrations and consequences for health. *Curr. Drug Metab.* **6**, 15–22.

28. Nijhout HF, Reed MC, Lam S-L, et al. (2006) In silico experimentation with a model of hepatic mitochondrial folate metabolism. *Theor. Biol. Med. Model.* **3**, 40.

29. Morrow GP, MacMillan L, Lamarre SG, et al. (2015) In vivo kinetics of formate metabolism in folate-deficient and folate-replete rats. *J. Biol. Chem.* **290**, 2244–50.

30. Lamarre SG, Morrow G, Macmillan L, et al. (2013) Formate: an essential metabolite, a biomarker, or more? *Clin. Chem. Lab. Med.* **51**, 571–8.

31. van Mil NH, Bouwland-Both MI, Stolk L, et al. (2014) Determinants of maternal pregnancy one-carbon metabolism and newborn human DNA methylation profiles. *Reproduction* **148**, 581–92.

32. Huang T, Zheng J, Chen Y, et al. (2011) High consumption of Ω-3 polyunsaturated fatty acids decrease plasma homocysteine: a meta-analysis of randomized, placebo-controlled trials. *Nutrition* **27**, 863–7.

33. Dawson SL, Bowe SJ & Crowe TC (2016) A combination of omega-3 fatty acids, folic acid and B-group vitamins is superior at lowering homocysteine than omega-3 alone: A meta-analysis. *Nutr. Res.* **36**, 499–508.

ANNEX 3.2 Supplementary Material 2: Search terms used for narrative review

Inclusion criteria for the review:
1) Type of exposure: Nutrition-related, prioritising nutrients related to one-carbon metabolism
2) Timing of study: Preconception and pregnancy
3) Type of study: Human
4) Type of study: Intergenerational
5) Mechanism: DNA methylation
6) Outcome: growth-related, cardio-metabolic, cognitive

Databases searched: PubMed, Google Scholar. Search terms were applied to titles and abstracts.

Example of PubMed search terms:
1) **Nutritional Exposure:** (nutrition* OR micronutrient OR vitamin OR multivitamin OR folate OR "folic acid" OR B12 OR B6 OR B2 OR riboflavin OR "pyridoxal 5-phosphate" OR cobalamin OR betaine OR choline OR PUFA* OR "polyunsaturated fatty acids" OR "UNIMMAP" OR "one carbon" OR "1-carbon" OR "methyl donor" OR famine OR carbohydrate OR protein OR fat OR "energy restriction" OR "energy intake" OR diet OR supplement*)

2) **Exposure timing:** (pregnan* OR periconception* OR peri-conception* OR trimester OR "in utero")

3) **Human study:** (women OR woman OR maternal OR human) NOT (mouse OR mice OR rodent* OR rat OR rats OR pig OR pigs OR sheep)

4) **Intergenerational study:** (offspring OR infant OR child OR children OR intergeneration*)

5) **Mechanism of DNA methylation:** (epigenetic* OR epigenom* OR "DNA methylation" OR methylation)

6) **Outcome:** (outcome OR phenotype OR cardio-metabolic OR "lipid profile" OR cholesterol OR "metabolic syndrome" OR "type-2 diabetes" OR "type-II diabetes" OR "body composition" OR "bone composition" OR adipos* OR obesity OR obese OR "body mass index" OR BMI OR growth OR anthropometr* OR "blood pressure" OR "mental health" OR cognit* OR ADHD OR "attention-deficit/hyperactivity disorder" OR "intelligence quotient")

Search Strategy:

- Search 1: For the review of nutritional exposures in pregnancy on infant DNA methylation search terms from 1) to 5) were linked together with 'AND'.

- Search 2: For the review of offspring methylation patterns and phenotypic outcomes there were two approaches. Firstly, the search terms from 4) to 6) were linked together with 'AND'. Secondly, a nutrition-sensitive locus found in the first review (e.g. '*LEP*') was entered and linked with search terms 5) and 6) using 'AND'. Only genes that had featured in the first search were included in the second search.

- Search 3: To check for studies linking maternal nutritional exposure to offspring phenotype, mediated by DNA methylation, we ran the search terms from 1) to 6) linked by 'AND'.

ANNEX 3.3   Supplementary Table 1: Studies investigating associations between maternal nutritional exposures and offspring DNA methylation (Search 1)

KEY: Intervention studies   Studies using broader nutritional exposures   No colour: observational studies involving one-carbon metabolites

Reference	Exposure	Exposure timing	Study design	Genes Investigated	Summary of results
Amarasekera M, et al. *FASEB J.* 2014;**28**: 4068–76[1]	Maternal serum folate. High and low folate groups.	3rd trimester	Nested cohort study. 23 mother-infant pairs, Australia. Tissue: Cord blood CD4+ antigen presenting cells. Platform: Sequenom Epityper	*ZFP57* (main focus) Other DMRS: *ACADM, C21orf56, FZD7, LASP1, LY6E,WNT9A*	High folate group: Hypomethylation at *ZFP57* (mean differential methylation 19%), *LY6E* (8%) and *C21orf56* (11%). Hypermethylation at *LASP1* (8%), *ACADM* (13%), *WNT9A* (10%), *FZD7* (8%)
Amarasekera M, et al. *Epigenetics*. 2014;**9**:1570–1576[2]	Maternal PUFA supplementation	2nd & 3rd trimesters	Subset of randomised controlled trial study, (70 mother-infant pairs, 36 intervention, 34 control). Trial was 3.7 g of fish oil (56.0% as DHA and 27.7% as EPA) from 20 weeks gestation to delivery vs. control. Tissue: Isolated CD4+ cells from cord blood Platform: Illumina Infinium HumanMethylation450 BeadChip	*Genome-wide*	Maternal docosahexaenoic acid (DHA), eicosapentaenoic acid (EPA), or total n-3 PUFAs showed dose-response effects on methylation at certain loci but none reached genome-wide significance.
Azzi S, et al. *Epigenetics*. 2014;**9**: 338–45[3]	Maternal vitamins pre-pregnancy BMI, vitamins B2, B3, B6, B9 (folate), B12	3 months before conception and last trimester	Nested cohort study in France (254 mother-infant pairs). Healthy infants from the 'EDEN' cohort. Maternal FFQs at 15 weeks gestation and post-delivery. Pre-pregnancy weight was by recall. Tissue: Infant cord blood measured Platform: Allele-specific methylated multiplex real-time quantitative PCR.	*PLAGL1 (ZAC1)*	Maternal B2 was positively correlated with *ZAC1* DMR methylation index at pre-pregnancy (r = 0.14, p = 0.04) and in the last trimester (r = 0.11, p = 0.09). Pre-pregnancy BMI was positively associated with methylation at *ZAC1* DMR.
Ba Y, et al. *Eur J Clin Nutr.* 2011;**65**: 480–5[4]	Maternal serum folate and vitamin	Delivery	Hospital-based cross-sectional study. 99 mother-infant pairs, China.	*IGF2*	Higher B12 associated with lower methylation at *IGF*2P3 (0.2% per SD). No association of folate with methylation.

Reference	Exposure	Exposure timing	Study design	Genes Investigated	Summary of results
	B12		Tissue: Cord blood Platform: Power SYBR Green PCR Master Mix and real-time methylation specific PCR.		
Cooper WN, et al. *FASEB J.* 2012;**26**: 1782–90[5]	UNIMMAP supplementation	Periconception	Sub-sample from an RCT, The Gambia (58 mother-infant pairs; 36 intervention, 22 control). Intervention was one tablet of UNIMMAP vs. control from pre-pregnancy to mean 9.5 weeks gestation. UNIMMAP includes 15 micronutrients (vitamins A, D, E, B1, B2, B6, Bi 2, C, Niacin, Folic Acid, Fe, Zn, Cu, I, Se) at the Recommended Daily Allowance level, except for folic acid, which is included at a level of 400 μg Tissue: Infant cord blood and venous blood at 9 months Platform: mass spectrometry (SpectroCHIPs, Sequenom)	*GNASAS, IGF2, IGF2R, MEG3 (GTL2), MEST (PEG1), PEG3, PLAGL1 (ZAC1)*	There were no overall effects of UNIMMAP supplementation. Cord blood results: In girls of supplemented mothers there was 8.6% lower methylation of *IGF2R* DMR compared to controls (p=0.023). In boys of supplemented mothers there was 6.5% lower methylation of *GTL2* DMR2 compared to controls (p=0.044). 9 month venous blood results: In girls of supplemented mothers there was 7.7% lower methylation of *PEG1* compared to controls (p=0.030), and 5.5% lower methylation of *GNASAS* (p=0.047).
Dominguez-Salas P, et al. *Nat Commun*. 2014;**5**: 3746[6]	1-carbon metabolites / season	Periconception	Cohort study, The Gambia. Followed 84 women conceiving in peak of rainy season and 83 in peak of dry season and their infants. Maternal measurement: plasma folate, B6, B12, active B12, choline, betaine, methionine, homocysteine, SAM, SAH, dimethyl glycine and erythrocyte B2 back-extrapolated to conception. Tissue: Infant venous blood taken at mean 3.6 months of age. Platform: PSQTM HS 96 pyrosequencer (Biotage)	*RBM46, BOLA3, FLJ20433 (EXD3), LOC654433 (PAX8-AS1), ZFYVE28, ZNF678*	Methylation potential in maternal plasma was higher during the rainy season compared to the dry season. Offspring of rainy season conceptions had higher levels of CpG methylation at the six metastable epialleles in peripheral blood monocytes when considered jointly. Individually only *RBM46* demonstrated a significant difference.
Drake AJ, et al., *Clin Endocrinol*. 2012;**77**:808-15[7].	Maternal diet (food groups)	'Early' (<20 weeks) and 'late' (>20 weeks)	Retrospective cohort study ('Motherwell'). 34 offspring at 40 years of age and mother's dietary records looked up.	*HSD2, NR3C1, IGF2, H19 ICR, GR* exon 1F	Higher methylation at *GR* exon 1F was observed in mothers having higher meat/fish/vegetables and lower bread/potato intake in late pregnancy. Higher

Reference	Exposure	Exposure timing	Study design	Genes Investigated	Summary of results
		pregnancy	Tissue: Whole blood Platform: Pyrosequencing using PSQTM HS-96A (Qiagen)		methylation at *HSD2* with increased meat and fish intake.
Finer S, et al. *BMJ Open.* 2016;**6**:e011768[(8)].	Famine	All trimesters	Case-control study, Bangladesh. Cases: Offspring exposed to famine *in utero* for at least 7 months of gestation (n=13). Controls: unexposed (n=18). Tissue: Whole blood age 28-31 years Platform: Illumina Infinium HumanMethylation450 BeadChip	*VTRNA2-1, PAX8, PRDM9, HLA-DQB2, PLD6, near ZFP57, AKAP12, ATP5B, LRRC14B, SPG20, near BOLA, RBM46, ZFYVE28, EXD3, PARD6G, ZNF678* and *ZFYVE28*	Comparing cases with unexposed controls in Bonferroni- corrected post-hoc pairwise comparisons there was higher methylation at *PAX8* (mean beta 0.771 vs. 0.730, p<0.005), lower methylation at *PRDM9* (0.656 vs 0.720, p<0.001) and lower methylation near *ZFP57* (0.681 vs. 0.683, p<0.001).
Godfrey KM, et al. *Diabetes*. 2011;**60:** 1528–1534[(9)]	Maternal carbohydrate intake	2nd trimester	Cohort, UK (78 mother-child pairs from Princess Anne Hospital and 239 mother-infant pairs in replication cohort - Southampton Women's Study). Nutrition exposure measured by FFQ. Tissue: Infant cord blood Platform EpiTyper software v1.0 (Sequenom)	*RXRA, NOS3, SOD1, IL8, PI3KCD*	*RXRA* methylation was inversely associated with maternal carbohydrate intake in early pregnancy.
Gonseth S, et al. *Epigenetics* 2015;**10**:1166–76[(10)]	Maternal folate intake	Periconception	Used healthy controls from an existing case-control study. N=343, USA. Maternal folate at conception assessed by retrospective FFQs, median 4 years later. Top hits were validated in an independent dataset. Tissue: Cord blood dried blood spots Platform: Illumina Infinium HumanMethylation450 BeadChip	*STX11, OTX2, TFAP2A, CYS1*	All CpG methylation in the four top hits showed an inverse relationship with folate intake. Lower methylation at these sites were associated with increased gene expression. When stratified by folate intake category, women with low intake (<200 µg/day) showed more positive associations between folate and methylation than inverse ones using epigenome-wide results. For women with high folate intakes (>600 µg/day) this trend reversed and more inverse associations were seen.

Reference	Exposure	Exposure timing	Study design	Genes Investigated	Summary of results
Haggarty P, et al. *Am J Clin Nutr* 2013; **97**:94–9[11]	Maternal folate and folic acid supplementation	Throughout pregnancy	Cohort study. 913 mother-infant pairs, UK Maternal folic acid supplementation assessed by FFQ covering pre-conception, <12 weeks gestation and >12 weeks gestation. Maternal folate was measured using red blood cell folate at 19 weeks gestation. Tissue: Infant cord blood Platform: Pyrosequencing (PyroMark, Qiagen).	*PEG3, IGF2 DMR, SNRPN*	Maternal folic acid supplementation started after 12 weeks gestation was associated with increased methylation in *IGF2* (0.7%, p=0.044) and decreased methylation in *PEG3* (-0.5%, p=0.018) compared to no supplementation. There was no effect of folic acid taken preconception or <12 weeks gestation. There was no association when the same analysis was done using maternal red blood cell folate.
Heijmans BT, et al. *Proc Natl Acad Sci USA*. 2008;**105:** 17046–9[12]	Famine	Throughout pregnancy	Dutch hunger winter. Retrospective case-control study six decades after famine (cases n=60 exposed to famine periconceptionally, n=62 exposed in late gestation, each case had a same-sex, unexposed sibling control). Tissue: Adult whole blood Platform: a mass spectrometry–based method (Epi- typer, Sequenom).	*IGF2*	In those offspring exposed to famine periconceptionally there was lower methylation at *IGF2* DMR (-5.2%, p= 5.9 x 10^{-5}) compared to controls. There was no association with those exposed during late gestation.
Hoyo C, et al. *Epigenetics*. 2011;**6:** 928–936[13]	Maternal folic acid intake ('none', 'moderate' and 'high' groups).	Preconception & Prenatal (all trimesters)	Cohort study ('NEST' cohort). 438 mother-infant pairs, USA. Tissue: Cord blood leukocytes Platform: Pyromark Q96 MD Pyrosequencer (Qiagen).	*IGF2, H19*	No association of pre-conception folic acid intake with *IGF2* DMR methylation. 'Moderate' intake associated with 2% lower methylation at *H19* compared with 'none' group. 'High' intake associated with 2.8% lower methylation. Associations stronger in boys. Similar results for pre-natal folate intakes.
Hoyo C, et al. *Epigenetics*. 2014;**9:** 1120–30[14]	Maternal erythrocyte folate	1st trimester (median 12 weeks gestation)	Cohort study ('NEST' cohort). 496 mother-infant pairs, USA. Tissue: Cord blood leukocytes. Platform: Pyromark Q96 MD Pyrosequencer (Qiagen).	*IGF2, H19, PEG1/MEST, PEG3, PLAGL1, MEG3-IG, PEG10/SGCE, NNAT, DLK1/MEG3*	Higher folate levels associated with decreased methylation at *MEG3, PLAGL1* and *PEG3* and increased methylation at *IGF2*.

Reference	Exposure	Exposure timing	Study design	Genes Investigated	Summary of results
Jiang X, et al. *FASEB J.* 2012;**26**: 3563–74[15]	Maternal serum choline	3rd trimester	Randomised controlled trial, USA (24 mother-infant pairs, 12 per arm). 480 vs. 930 mg choline / day given from 26-29 weeks gestation for 12 weeks. Tissue: Placenta & cord blood leukocytes Platform: EpiTyper (Sequenom).	*CRH, GNASAS, IGF2, IL10, LEP, NR3C1*	*CRH*: Higher maternal choline intake associated with higher methylation of placental promoter (~4%, p=0.05) but lower methylation (~2.5%, p=0.04) of cord blood (effect size estimated from figure). *GNASAS*, *IGF2*, *IL10*, *LEP*: no effect *NR3C1*: Higher maternal choline intake associated with higher methylation of placental promoter (0.7%, p=0.002) but lower methylation (-0.6%, p=0.04) of cord blood.
Joubert BR, et al. *Nat Commun.* 2016;**7**: 10577[16]	Maternal plasma folate	2nd trimester (median 18 and 13 weeks gestation for the two cohorts).	Meta-analysis of two cohorts: Norwegian Mother and Child Cohort Study ('MoBa', N=1275); Generation R (Netherlands, N=713). Tissue: Infant cord blood Platform: Illumina Infinium HumanMethylation450 BeadChip	*APC2, GRM8, KLK4, LHX1, OPCML, PRPH, PRSS21, SLC16A12*	443 CpGs demonstrated differential methylation by folate levels using false discovery rate (FDR) of 5%. 94% showed an inverse relationship between folate and methylation. 48 CpGs demonstrated differential methylation by folate levels after Bonferroni correction (P<1.17 x 10^{-7}). The genes listed here represent the ones with the smallest p values.
Kühnen P, et al. *Cell Metab*. 2016;**24**:502–509[17]	1-carbon metabolites / season	Periconception	Cohort study, The Gambia. 144 mother-child pairs from 'MDEG' cohort. Tissue: Infant venous blood taken at mean 3.6 months of age. Platform: Illumina Infinium HumanMethylation450 BeadChip used for initial screen.	*POMC*	Lower methylation at the *POMC* VMR in children conceived in the dry season compared to those conceived in the rainy season (coefficient -0.152, p=0.034).
Lee H-S, et al. *Physiol Genomics*. 2014;**46**: 851–7[18]	Maternal PUFA supplementation	2nd & 3rd trimester	Randomised controlled trial. 261 mother-child pairs (131 supplemented, 130 control) in Mexico. 400mg docosahexaenoic acid (DHA) vs. placebo given daily from 18-22wk gestation to birth. Tissue: Infant cord blood Platform: PyroMark ID system (Qiagen)	*H19, IGF2*	No overall differences. The DHA supplemented group infants had lower *H19* methylation (-1.66%, p=0.01) among mothers of BMI < 25 kg/m². Methylation of *IGF2* DMR was higher (1.35%, p=0.03) in the supplemented group for infants born to overweight mothers. DNA methylation levels in *IGF2* P3 were higher (1.88%, p=0.04) in the DHA group than the control group in preterm infants.

Reference	Exposure	Exposure timing	Study design	Genes Investigated	Summary of results
Lin, X, et al. *BMC Med.* 2017;15:50[19].	Maternal BMI, glucose, plasma fatty acids, plasma vitamin D, serum B12, B6, folate, iron, zinc, magnesium	3rd trimester (26-28 weeks)	Cohort study ('GUSTO' cohort), 987 mother-infant pairs, Singapore. Tissue: Infant cord blood Platform: Illumina Infinium HumanMethylation450 BeadChip	EWAS	Methylation of cg25685359 (*MIRLET7BHG*) was inversely associated with maternal n-6 PUFA levels ($P = 4.2 \times 10^{-4}$). Methylation of cg23671997 (IGDCC4) was positively associated with maternal fasting glucose levels ($P = 2.7 \times 10^{-4}$).
Marchlewicz EH, et al. *Sci Rep.* 2016;6:34857[20].	Maternal metabolomic profile, including PUFAs	1st trimester	Cohort study, USA (40 mother-infant pairs) Tissue: Infant cord blood Platform: PyroMark Q96 MD pyrosequencing (Qiagen).	*IGF2, H19, ESR1, PPARα*	Several derivatives of linoleic acid and α-linolenic acid were associated with candidate gene methylation, but direction of associations differed. Very Long Chain Fatty Acids were positively correlated with methylation at *ESR1* and *PPARα.*
McCullough LE, et al. *Clin Epigenetics.* 2016;8: 8[21]	Maternal plasma B12, B6, Hcy	1st trimester (mean 12 weeks gestation)	USA (496 mother-infant pairs), 'NEST' cohort Tissue: Infant cord blood Platform: Pyrosequencing (Pyromark Q96 MD, Qiagen).	*H19, MEG3 (GTL2), PEG10 / SGCE, PLAGL1 (ZAC1)*	No association between maternal micronutrient levels and *H19, PEG10 / SGCE, PLAGL1 (ZAC1)* Vitamin B6 (PLP) in quartiles positively associated with methylation at *MEG3* DMR (βQuartile 3 = 2.01 and βQuartile 4 = 3.24, p <0.05).
McKay JA, et al. *PLoS One.* 2012;7: e33290[22]	Maternal red blood cell folate, serum vitamin B12	1st trimester (mean 10.6 weeks gestation)	Nested cohort study, UK (121 mother-infant pairs). Also looked at genetic variants. Tissue: Infant cord blood Platform: Pyrosequencing (Pyromark MD, Qiagen).	*IGFBP3, IGF2*	Methylation of the *IGFBP3* locus inversely correlated with infant cord vitamin B12 concentration (standardised β -0.345, p<0.001), but not maternal nutrient concentrations. Maternal B12 concentration inversely correlated with infant global DNA methylation. Maternal MTHFR C677T genotype associated with *IGF2* methylation (but no association with micronutrients).

Reference	Exposure	Exposure timing	Study design	Genes Investigated	Summary of results
Pauwels S, et al. *Clin Epigenetics* 2017;**9**:16[(23)]	Maternal choline, betaine, folate, methionine, folic acid	Throughout pregnancy, including periconception	Cohort study ('MANOE' study). 114 mother-infant pairs, Belgium. Maternal intakes assessed by food frequency questionnaires (FFQs) at different stages of pregnancy (periconception, 1st, 2nd, 3rd trimester) for different mothers. Tissue: Buccal epithelial cell in 6 months old infants Platform: Pyrosequencing (PyroMark Q24, Qiagen).	*IGF2 DMR, DNMT1, LEP, RXRA*	Periconceptional maternal intake (n=21): At *RXRA* (all CpGs) each increase in 100 μg folate was associated with 0.685% increase in methylation (p=0.027), and a 0.875% increase in methylation per 100 mg increase in betaine. At *LEP* (all CpGs) each increase in 100 μg folate was associated with a decrease in 1.233% methylation (p=0.030). At *IGF2* DMR (all CpGs) an increase in 100 μg folic acid was associated with a decrease of 0.706% methylation (p=0.010) 2nd trimester intake (n=85): At *DNMT1* (all CpGs) an increase in 100 μg folic acid was associated with a decrease of 0.027% methylation (p=0.020) 3rd trimester intake (n= 82): At *DNMT1* CpG1 an increase in 100mg choline was associated with an increase of 0.156% methylation (p=0.017). At *DNMT1* CpG3 an increase of 100 μg folate was associated with an increase of 0.131% methylation (p=0.026).
Pauwels S, et al. *Epigenetics* 2017;**12**:1–10[(24)]	Maternal choline, betaine, folate, methionine, folic acid	Throughout pregnancy, including periconception	Cohort study ('MANOE' study). 115 mother-infant pairs, Belgium. Maternal intakes assessed by FFQ at different stages of pregnancy (periconception, 1st, 2nd, 3rd trimester) for different mothers. Tissue: Cord blood Platform: Pyrosequencing (PyroMark Q24 instrument, Qiagen).	*IGF2 DMR, DNMT1, LEP, RXRA*	Periconceptional maternal intake (n=24): At *DNMT1* CpG4 each increase in 100 mg betaine was associated with 0.675% increase in methylation (p=0.039). At *LEP* CpG each increase in 100 mg methionine was associated with 0.427% increase in methylation (p=0.048). 2nd trimester (n=89): At *LEP* CpG2 each increase in 100 mg betaine was associated with 0.575% decrease in methylation (p=0.05). Each increase in 100 μg folate was associated with 0.507% decrease in methylation (p=0.009). At *DNMT1* CpG4 each increase in 100 mg choline was associated with 0.301% decrease in methylation (p=0.031). Each increase in 100 μg folate was associated with 0.226% decrease in methylation (p=0.045).

Reference	Exposure	Exposure timing	Study design	Genes Investigated	Summary of results
					3rd trimester (n=83): At *DNMT1* CpG2 each increase in 100 mg choline was associated with 0.291% increase in methylation (p=0.022). At *RXRA* CpG2 each increase in 100 μg folate was associated with 1.0% decrease in methylation (p=0.041). There was 2.3% higher methylation at *LEP* (mean of CpGs) when women had taken folic acid for 6 months prior to conception compared to those not taking any supplementation. Mean *RXRA* methylation was 1.2% higher in mothers who took folic acid supplementation throughout pregnancy compared to those who stopped in the 2nd trimester (p=0.01).
Qian YY, et al., J *Hum Nutr Diet*. 2016;**29**:643-51[25].	Folic Acid	Pregnancy – no details	Case-control study. 39 small-for-gestational age (SGA) infants and 49 appropriate-for-gestational-age (AGA) controls, China. Information on folic acid supplementation taken by mothers obtained through questionnaire responses. Tissue: Umbilical cord blood Platform: EpiTYPER, Sequenom.	*MEST,H19*	Six sites in the *H19* DMR were found to have significantly higher methylation in SGA when compared to AGA. The associations were stronger in males in the folic acid supplemented group.
Rijlaarsdam J, et al., *J Child Psychol Psychiatry*. 2017;**58**:19-27[26].	High-fat and -sugar diet	3rd trimester (32 weeks)	Nested cohort study ('ALSPAC' cohort). 164 youth with low conduct problems (CP) and early onset persistent (EOP) ADHD. FFQs administered to mothers. Tissue: Cord blood, peripheral blood at age 7 Platform: Illumina Infinium HumanMethylation450 BeadChip	*IGF2*	High-fat and -sugar diet taken prenatally was associated with higher *IGF2* methylation.
Silver MJ, et al. *Genome Biol*. 2015;**16**: 118[27]	1-carbon metabolites / season	Periconception	Cohort study, The Gambia. 110 children conceived in peak of dry season, 105 conceived in peak of rainy season. Used an epigenome-wide screen to create a list of potential MEs susceptible to	*VTRNA2-1*	The epigenome-wide screen top hit for differential methylation by season of conception was *VTRNA2-1*. Offspring conceived in the dry season had increased levels of hypomethylation (<40%) of *VTRNA2-1* DMR (p=0.004).

Reference	Exposure	Exposure timing	Study design	Genes Investigated	Summary of results
			periconceptional environment. Tissue: Infant venous blood taken at mean 3.6 months of age. Platform: 450k (Illumina) used for initial screen.		
Steegers-Theunissen RP, et al. *PLoS One*. 2009;4: e784[28]	Maternal folic acid	Periconception	Cross sectional study amongst controls for another study ('HAVEN'). 120 mother-infant pairs, The Netherlands. Folic acid use ascertained from questionnaires at 17 months post-delivery to determine two groups: 400 ug periconceptional supplementation (n=86) and unsupplemented (n=34). Tissue: Infant blood taken at 17 months. Platform: Mass spectrometry based platform (Epityper, Sequenom)	*IGF2*	Relative methylation of *IGF2* DMR (from 5 CpGs) 4.5% higher in children born to supplemented mothers (p=0.014).
Tobi EW, et al. Hum *Mol Genet*. 2009;**18**: 4046–53[29]	Famine	Periconception, late gestation	Dutch hunger winter. Retrospective case-control study six decades after famine (cases n=60 exposed to famine periconceptionally, n=60 same-sex, unexposed sibling controls). Also cases exposed to famine in late gestation (n=62) and controls (n=62). Tissue: Adult whole blood Platform: a mass spectrometry–based method (Epi- typer, Sequenom).	*ABCA1, CRH, GNASAS, IGF2R, IL10, INSIGF, LEP, MEG3 (GTL2)*	Famine exposure was significantly associated with methylation at *INSIGF* (↓), *GNASAS* (↑), *MEG3* (↑), *IL10* (↑), *LEP* (↑) and *ABCA1* (↑). Effect sizes ranged from -1.6 % to + 2.4 % (all p values <0.017). In addition, *INSIGF*, *GNASAS* and *LEP* demonstrated differences in methylation by sex. There was lower methylation at *GNASAS* for those exposed to famine in late gestation.
Tobi EW, et al. *PLoS One*. 2012;**7**:e37933[30]	Famine	Periconception	Dutch hunger winter. Retrospective case-control study six decades after famine (cases n=60 exposed to famine periconceptionally, n=60 same-sex, unexposed sibling controls) Tissue: Adult whole blood	*INS, INSIGF, IGF2* and *H19*	Methylation at *INSIGF* was 1.5% lower in the cases compared to controls. Directions of associations with methylation at *IGF2* varied according to the specific DMR.

Reference	Exposure	Exposure timing	Study design	Genes Investigated	Summary of results
			Platform: a mass spectrometry–based method (Epi- typer, Sequenom).		
Tobi EW, et al. *Int J Epidemiol.* 2015 May 5;44:1211–1223[31]	Famine	Gestational weeks 1–10, 11–20, 21–30, or 31 to delivery	Dutch hunger winter. 384 subjects exposed in the four timing windows. Tissue: Adult whole blood Platform: Illumina Infinium HumanMethylation450 BeadChip	*FAM150B, SLC38A2, PPAP2C, OSBPL5/MRGPRG, TMEM105, TACC1* and *ZNF385A*	Famine exposure in gestational weeks 1-10 had increased methylation at *FAM150B, SLC38A2, PPAP2C*, and lower methylation at *OSBPL5/MRGPRG* compared to later windows of exposure. Methylation of *TACC1* and *ZNF385A* was increased after exposure during any time in gestation. Methylation at *TMEM105* was lower when exposure was at periconception.
Van Dijk SJ, et al. *Clin Epigenetics*. 2016;8:114[32]	Maternal PUFA supplementation	2nd & 3rd trimesters	Subset of RCT trial, Australia (n= 369 at birth; 190 intervention, 179 control). Trial was 800 mg/day DHA vs. placebo from 20 weeks gestation to delivery. Tissue: Infant blood spots at delivery, and venepuncture at 5 years Platform: Illumina Infinium HumanMethylation450 BeadChip	*ESYT3, SLC12A6, CCK, MAD1L1, GTF2A1L, STON1, RPS6KA2, FAM110B, PWWP2B, TRAK1, RAET1L*	The listed genes contained CpGs showing differential methylation according to intervention arm ($p<5\times10^{-5}$). Directions of associations changed, but mostly DHA was inversely associated with methylation. Maximum effect size was a difference of 4.5% methylation. These trends were still seen in the samples at 5 years, although they no longer reach genome-wide significance.
van Mil NH, et al. *Reproduction*. 2014;148: 581–92[33]	Maternal plasma folate, plasma homocysteine (Hcy) and folic acid intake	1st trimester (mean 13 weeks)	Nested cohort study ('Generation R'). 463 mother-infant pairs, The Netherlands. Questionnaires assessed periconceptional folic acid intake. Tissue: Infant cord blood Platform: Mass spectrometry based platform (Epityper, Sequenom)	*H19, NR3C1, DRD4, 5-HTT, IGF2 DMR, KCNQ1OT1, and MTHFR*	No association between maternal plasma folate and folic acid intake with infant DNA methylation when exposures treated as continuous variables. A 1 µmol/l increase in plasma homocysteine was associated with 0.04% lower methylation at *NR3C1* (p=0.03). In folate deficient mothers (<7 nmol/l) there was lower methylation at *NR3C1* compared to folate replete mothers (≥7 nmol/l) (-0.60%, p<0.001). Mothers with elevated Hcy (>11 µmol/l) showed 0.92% higher methylation at *H19* compared to those

250

Reference	Exposure	Exposure timing	Study design	Genes Investigated	Summary of results
					with low Hcy (≤11 μmol/l) (p=0.05). There were no significant associations at any other gene regions.
Waterland RA, et al. *PLoS Genet*. 2010;**6**:e1001252[34]	Season	Periconception	Cross-sectional study taking samples from 30 children conceived in the rainy season and 30 conceived in the dry season, rural Gambia. Tissue: Venous blood taken at mean 8.9 years. Platform: PSQTM HS 96 pyrosequencer (Biotage)	*BOLA3, FLJ20433, PAX8 ZFYVE28, SLITRK1*	Methylation at all 5 metastable epialleles was higher in children conceived in the rainy season compared to the dry season (all p<0.03).

Abbreviations: DMR, differentially methylated region; EWAS, epigenome-wide association study; ME, metastable epiallele; RCT, randomised controlled trial; PUFA, polyunsaturated fatty acid; VMR, variably methylated region.

References

1. Amarasekera M, Martino D, Ashley S, et al. (2014) Genome-wide DNA methylation profiling identifies a folate-sensitive region of differential methylation upstream of ZFP57-imprinting regulator in humans. *FASEB J.* **28**, 4068–76.

2. Amarasekera M, Noakes P, Strickland D, et al. (2014) Epigenome-wide analysis of neonatal CD4 + T-cell DNA methylation sites potentially affected by maternal fish oil supplementation. *Epigenetics* **9**, 1570–1576.

3. Azzi S, Sas TCJ, Koudou Y, et al. (2014) Degree of methylation of ZAC1 (PLAGL1) is associated with prenatal and post-natal growth in healthy infants of the EDEN mother child cohort. *Epigenetics* **9**, 338–45.

4. Ba Y, Yu H, Liu F, et al. (2011) Relationship of folate, vitamin B12 and methylation of insulin-like growth factor-II in maternal and cord blood. *Eur. J. Clin. Nutr.* **65**, 480–5.

5. Cooper WN, Khulan B, Owens S, et al. (2012) DNA methylation profiling at imprinted loci after periconceptional micronutrient supplementation in humans: results of a pilot randomized controlled trial. *FASEB J.* **26**, 1782–90.

6. Dominguez-Salas P, Moore SE, Baker MS, et al. (2014) Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. *Nat. Commun.* **5**, 3746.

7. Drake AJ, McPherson RC, Godfrey KM, et al. (2012) An unbalanced maternal diet in pregnancy associates with offspring epigenetic changes in genes controlling glucocorticoid action and foetal growth. *Clin. Endocrinol.* **77**, 808–15.

8. Finer S, Iqbal MS, Lowe R, et al. (2016) Is famine exposure during developmental life in rural Bangladesh associated with a metabolic and epigenetic signature in young adulthood? A historical cohort study. *BMJ Open* **6**, e011768.

9. Godfrey KM, Sheppard A, Gluckman PD, et al. (2011) Epigenetic Gene Promoter Methylation at Birth Is Associated With Child's Later Adiposity. *Diabetes* **60**, 1528–1534.

10. Gonseth S, Roy R, Houseman EA, et al. (2015) Periconceptional folate consumption is associated with neonatal DNA methylation modifications in neural crest regulatory and cancer development genes. *Epigenetics* **10**, 1166–76.

11. Haggarty P, Hoad G, Campbell DM, et al. (2013) Folate in pregnancy and imprinted gene and repeat element methylation in the offspring. *Am. J. Clin. Nutr.* **97**, 94–9.

12. Heijmans BT, Tobi EW, Stein AD, et al. (2008) Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 17046–9.

13. Hoyo C, Murtha AP, Schildkraut JM, et al. (2011) Methylation variation at IGF2 differentially methylated regions and maternal folic acid use before and during pregnancy. *Epigenetics* **6**, 928–936.

14. Hoyo C, Daltveit AK, Iversen E, et al. (2014) Erythrocyte folate concentrations, CpG methylation at genomically imprinted domains, and birth weight in a multiethnic newborn cohort. *Epigenetics* **9**, 1120–30.

15. Jiang X, Yan J, West AA, et al. (2012) Maternal choline intake alters the epigenetic state of fetal cortisol-regulating genes in humans. *FASEB J.* **26**, 3563–74.

16. Joubert BR, den Dekker HT, Felix JF, et al. (2016) Maternal plasma folate impacts differential DNA methylation in an epigenome-wide meta-analysis of newborns. *Nat. Commun.* **7**, 10577.

17. Kühnen P, Handke D, Waterland RA, et al. (2016) Interindividual Variation in DNA Methylation at a Putative POMC Metastable Epiallele Is Associated with Obesity. *Cell Metab.* **24**, 502–9.

18. Lee H-S, Barraza-Villarreal A, Biessy C, et al. (2014) Dietary supplementation with polyunsaturated fatty acid during pregnancy modulates DNA methylation at IGF2/H19 imprinted genes and growth of infants. *Physiol. Genomics* **46**, 851–7.

19. Lin X, Lim IY, Wu Y, et al. (2017) Developmental pathways to adiposity begin before birth and are influenced by genotype, prenatal environment and epigenome. *BMC Med.* **15**, 50.

20. Marchlewicz EH, Dolinoy DC, Tang L, et al. (2016) Lipid metabolism is associated with developmental epigenetic programming. *Sci. Rep.* **6**, 34857.

21. McCullough LE, Miller EE, Mendez MA, et al. (2016) Maternal B vitamins: effects on offspring weight and DNA methylation at genomically imprinted domains. *Clin. Epigenetics* **8**, 8.

22. McKay JA, Groom A, Potter C, et al. (2012) Genetic and non-genetic influences during pregnancy on infant global and site specific DNA methylation: role for folate gene variants and vitamin B12. *PLoS One* **7**, e33290.

23. Pauwels S, Ghosh M, Duca RC, et al. (2017) Maternal intake of methyl-group donors affects DNA methylation of metabolic genes in infants. *Clin. Epigenetics* **9**, 16.

24. Pauwels S, Ghosh M, Duca RC, et al. (2017) Dietary and supplemental maternal methyl-group donor intake and cord blood DNA methylation. *Epigenetics* **12**, 1–10.

25. Qian Y-Y, Huang X-L, Liang H, et al. (2016) Effects of maternal folic acid supplementation on gene methylation and being small for gestational age. *J. Hum. Nutr. Diet.* **29**, 643–651.

26. Rijlaarsdam J, Cecil CAM, Walton E, et al. (2017) Prenatal unhealthy diet, insulin-like growth factor 2 gene (IGF2) methylation, and attention deficit hyperactivity disorder symptoms in youth with early-onset conduct problems. *J. Child Psychol. Psychiatry.* **58**, 19–27.

27. Silver MJ, Kessler NJ, Hennig BJ, et al. (2015) Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. *Genome Biol.* **16**, 118.

28. Steegers-Theunissen RP, Obermann-Borst SA, Kremer D, et al. (2009) Periconceptional maternal folic acid use of 400 microg per day is related to increased methylation of the IGF2 gene in the very young child. *PLoS One* **4**, e7845.

29. Tobi EW, Lumey LH, Talens RP, et al. (2009) DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Hum. Mol. Genet.* **18**, 4046–53.

30. Tobi EW, Slagboom PE, van Dongen J, et al. (2012) Prenatal famine and genetic

variation are independently and additively associated with DNA methylation at regulatory loci within IGF2/H19. *PLoS One* **7**, e37933.

31. Tobi EW, Slieker RC, Stein AD, et al. (2015) Early gestation as the critical time-window for changes in the prenatal environment to affect the adult human blood methylome. *Int. J. Epidemiol.* **44**, 1211–1223.

32. van Dijk SJ, Zhou J, Peters TJ, et al. (2016) Effect of prenatal DHA supplementation on the infant epigenome: results from a randomized controlled trial. *Clin. Epigenetics* **8**, 114.

33. van Mil NH, Bouwland-Both MI, Stolk L, et al. (2014) Determinants of maternal pregnancy one-carbon metabolism and newborn human DNA methylation profiles. *Reproduction* **148**, 581–92.

34. Waterland RA, Kellermayer R, Laritsky E, et al. (2010) Season of conception in rural gambia affects DNA methylation at putative human metastable epialleles. *PLoS Genet.* **6**, e1001252.

ANNEX 3.4   Supplementary Table 2: Studies investigating associations between nutrition-sensitive candidate genes and phenotypes (Search 2).

Reference	Study design	Genes Implicated	Phenotype investigated	Summary of results
Azzi S, et al. *Epigenetics*. 2014;**9**: 338–45[1]	Nested cohort study in France. 254 mother-infant pairs. Healthy infants from the 'EDEN' cohort. Tissue: Cord blood Platform: Allele-specific methylated multiplex real-time quantitative PCR.	*PLAGL1 (ZAC1)*	Pre- and post-natal growth	ZAC1 methylation index was positively correlated with estimated fetal weight at 32 weeks gestation (r=0.15, p=0.01). It was positively correlated with weight (r = 0.14, p=0.03) and BMI z-scores (r = 0.15, p=0.01) at age 1 year.
Bens S, et al. *Eur J Hum Genet*. 2013;**21**:838-43[2]	Case-control study. 98 SGA infants and 50 AGA controls from centres of the BMBF consortium, Germany (range 0-18 years). Tissue: Whole blood Platform: Pyrosequencer ID	*PLAGL1, IGF2R, GRB10, H19, IGF2, MEG3, NDN, SNRPN, NESP, NESPAS*	Size at birth	Cases showed hypomethylation at *GRB10* (n=1) and *H19* 2CTCF-binding site (n=1), and hypermethylation at *NDN* (n=1) and *IGF2* (n=1). Note case study approach therefore not included in main narrative Table 2.
Bouwland-Both MI, et al. *PLoS One*. 2013;**8**:e81731[3]	Nested cohort study ('Generation R' cohort subset). 69 small-for-gestational age (SGA) vs. 471 controls (appropriate-for-gestational-age; AGA). Tissue: Cord blood Platform: EpiTYPER, Sequenom	*IGF2, H19, MTHFR*	Birthweight and anthropometry. Weight gain at 3 months	Methylation at the *MTHFR* locus did not vary significantly between SGA and AGA infants. An inverse association was found between SGA and *IGF2* DMR0 methylation (β =-1.07, p=0.015). SGA was not significantly associated with *H19* promoter DMR methylation. *IGF2* DMR0 methylation was inversely associated with birth-three months weight gain (β=-0.46, p=0.022).
Burris HH, et al., *Epigenomics*. 2013; **5**: 271–281[4]	Cohort study. 219 infants, Mexico. Tissue: Cord blood. Platform: Pyrosequencing	*IGF2/H19, KCNQ1OT1, GCR, NR3C1, LINE-1, Alu*	Birthweight	No significant methylation-birthweight associations were found for the studied loci.
Córdova-Palomera A, et al. *PLoS One*. 2014;**9**:e103639[5]	Nested cohort study. 34 monozygotic twin pairs of European descent, Spain. Age 22-56 years. Tissue: Whole blood Platform: Illumina Infinium HumanMethylation450 (450K) Bead-Chip.	*IGF2, IGF2BP1, IGF2BP2, IGF2BP3*	Birthweight, working memory	Mean methylation at 2 CpG sites in *IGF2BP1* was associated with birthweight (β=83.3 x 10^{-3}, p=0.033) and working memory (β= -4.4 x 10^{-3}, p=0.009). No variation in methylation was observed in other loci studied.

Reference	Study design	Genes Implicated	Phenotype investigated	Summary of results
Deodati A, et al. *Horm Res Paediatr*. 2013;**79**:361-7[6]	Cross-sectional study. 85 children, Italy. Mean age 11.6 years. Tissue: Whole blood Platform: Methyl-Profiler DNA Methylation qPCR Assay	*IGF2*	Lipid profile	Children with intermediate methylation at *IGF2* had significantly higher levels of triglycerides (107.6 ± 41.99 vs. 76.6 ± 30.18 mg/dl, p < 0.005) and higher triglyceride:high-density lipoprotein-cholesterol ratio (2.23 ± 0.98 vs. 1.79 ± 0.98, p < 0.02) in comparison with children showing hypomethylation at *IGF2*.
Drake AJ, et al. *Clin Endocrinol*. 2012;**77**:808-15[7]	Retrospective cohort study ('Motherwell'). 34 offspring at 40 years of age. Tissue: Whole blood Platform: Pyrosequencing using PSQTM HS-96A (Qiagen)	*IGF2, H19 ICR, HSD2, NR3C1*	Birthweight, current height, weight, waist circumference, blood pressure	There was an inverse association between methylation at specific CpGs in *HSD2* and neonatal ponderal index (Region 2- CpG14 r=-0.38, CpG19 r=-0.34, p<0.05) and *H19* ICR methylation and birth length (r=-0.36, p<0.05). *HSD2* methylation was positively associated with birthweight (r=0.49, p<0.05), adiposity measures and blood pressure in adulthood. *H19* ICR methylation was not significantly associated with birthweight but was positively associated with weight in adulthood (β=0.37,p=0.03). Both *H19* and *NR3C1* Exon methylation were associated with waist circumference, BMI and blood pressure at age 40.
Dunstan J, et al., *Clin Epigenetics*. 2017; **29**:29[8]	431 adolescents, USA. Age 10-15 years. Tissue: Saliva Platform: Pyromark Q24, Qiagen	*LEP, ICAM-1, CRH, LINE-1*	BMI, waist circumference, percent body fat	In obese boys, *LEP* methylation was inversely associated with the obesity. No significant associations were found for *ICAM-1*, *CRH* and *LINE-1*.
García-Cardona MC, et al. *Int J Obes* 2014;**38**:1457-65[9]	Cross-sectional study, n=106 divided into lean and obese adolescents, Mexico. Age 10-16 years. Tissue: Whole blood Platform: Methylation-specific PCR	*LEP, ADIPOQ*	BMI, insulin resistance, glucose, cholesterol and triglycerides levels	No significant variation in overall methylation of *LEP, ADIPOQ*. Lower methylation at both genes in obese subjects with insulin resistance.
Godfrey KM, et al. *Diabetes*. 2011;**60**: 1528–1534[10]	Cohort study, UK (78 mother-child pairs from Princess Anne Hospital (PAH) and 239 mother-infant pairs in replication cohort - Southampton Women's Study(SWS)). Childhood adiposity measurements were made using dual energy X-ray absorptiometry DEXA (PAH study, age 9 years; SWS, age 6 years).	*RXRA, eNOS*	Adiposity at age 9 years	*RXRA* methylation was associated with sex-adjusted childhood fat mass (exponentiated regression coefficient [β] 17% per SD change in methylation, p=0.009) and % fat mass (β=10%, p=0.023) at age 9 in the PAH cohort and also had similar association in the SWS cohort (Fat mass- β=6%, p=0.002; %fat mass- β=4%, p=0.002).

Reference	Study design	Genes Implicated	Phenotype investigated	Summary of results
	Tissue: Cord blood Platform EpiTyper software v1.0 (Sequenom)			Methylation at *eNOS* was associated with fat mass (β=20%, p<0.001) and %fat mass (β=12%, p=0.002) in the PAH cohort only.
Harvey, NC et al. *J Bone Miner Res.* 2014; **29.3:** 600–607[11]	Cohort, UK (230 Southampton Women's Study and 64 Princess Anne Hospital mother-infant pairs) Tissue: Cord blood Platform: Sequenom MassARRAY Compact System	*RXRA*	Bone mineral content at age 4 years	Methylation for four of six *RXRA* CpG sites inversely correlated with % bone mineral content. Note that maternal free 25(OH)-vitamin D index inversely associated with methylation at one *RXRA* CpG site (β=−3.3 SD/unit, p=0.03).
Houde AA, et al. *BMC Med Genet.* 2015;**16:**29[12]	Cross-sectional study. 73 severely obese adults, Canada. Mean age 34.7 years. Tissue: Blood, subcutaneous (SAT) and visceral adipose tissues (VAT). Platform: Pyrosequencing	*LEP, ADIPOQ*	BMI, anthropometry, blood pressure, lipid profile	Higher *ADIPOQ* methylation levels in SAT were associated with higher BMI and waist circumference. Lower *LEP* methylation in blood was associated with higher BMI. A positive correlation was found between fasting LDL-C levels and *LEP* in blood and SAT, and with *ADIPOQ* in *SAT* and VAT.
Hoyo C, et al. *Cancer Causes Control*. 2012;**23:**635-45[13]	Cohort study (NEST cohort), USA. 300 mother-infant pairs. Tissue: Cord blood Platform: Pyromark Q96 MD (Qiagen)	*IGF2, H19*	Birthweight	*IGF2* DMR0 methylation was inversely associated with *IFG2* protein concentrations in cord blood (β = -9.87, p<0.01), having stronger associations in infants of obese women (β =-20.21, p<0.0001). Higher concentrations of *IGF2* were related to higher birth weight (p<0.001). No associations found with *H19* DMR.
Hoyo C, et al. *Epigenetics*. 2014;**9:** 1120–30[14]	Cohort study ('NEST' cohort). 438 mother-infant pairs, USA. Tissue: Cord blood leukocytes Platform: Pyromark Q96 MD Pyrosequencer (Qiagen).	*H19, PEG10/SGCE, PLAGL1, MEG3*	Birth weight	Higher methylation at *H19*, *PEG10/SGCE* and *PLAGL1* DMRs, and lower *MEG3* methylation, associated with higher birth weight.
Huang RC, et al. *Clin Epigenetics*. 2012;**4:**21[15]	Cohort study ('Raine' Study). 315 children, Australia. Anthropometric parameters measured at birth and follow ups taken at 8 time points after birth. Tissue: Whole blood at age 17 Platform: EpiTyper, Sequenom.	*IGF2/H19*	BMI, anthropometry, birthweight	No association was noted between *IGF2/H19* ICR1 and anthropometric measures at birth. A 3.4% increase in methylation at 2 specific CpGs in the *IGF2/H19* ICR was associated with 18 mm decrease in head circumference at age 17 (p= 0.006).
Kappil MA, et al. *Epigenetics*. 2015;**10:**842-9[16]	Cohort study (Rhode Island Child Health Study (RICHS), USA). 677 mother-infant pairs, subset of n=211 for methylation analysis.	108 imprinted genes	Size at birth	Increased methylation of *MEST* observed in SGA infants. The paper focuses on gene expression data.

Reference	Study design	Genes Implicated	Phenotype investigated	Summary of results
	Tissue: Placenta Platform: Infinium HumanMethylation450 BeadChip.			
Kuehnen, P. et al. *PLoS genetics*. 2012; **8:** p.e1002543[17]	Case-control study, Germany. Normal weight controls (mean age 17.9 years, N=90), obese cases (mean age 11 years, N=171). Tissue: Peripheral blood Platform: Direct sequencing of bisulfite-converted DNA	*POMC*	Obesity at age 11 years	Increased methylation score at *POMC* in obese children compared to controls. Overall CpG methylation score from position -4 to +6 was 40% in cases and 25% in controls (p<0.001).
Kühnen P, et al. *Cell Metab*. 2016;**24**:502–509[18]	Case-control study in German adults: 103 normal-weight and 125 obese, mean age 48.2 years. Tissue: Peripheral blood Platform: Pyromark Q24 (Qiagen)	*POMC*	BMI	Positive correlation of average *POMC* VMR methylation with BMI (r = 0.18, p = 0.006).
Lesseur C, et al. *Mol Cell Endocrinol*. 2013;**381**:160-7[19]	Nested cohort study, USA. Rhode Island Child Health Study 'RICHS' cohort, 81 mother-infant pairs, categorised by birth size. Tissue: Cord blood, placenta Platform: Pyromark MD (Qiagen)	*LEP*	Size at birth	Higher *LEP* methylation levels were observed in SGA children (p=4.6×10⁻³). Infants born to pre-pregnancy obese mothers had lower cord blood methylation levels (p=0.03). There was an interaction between placental *LEP* methylation and sex (p=0.05).
Lesseur C, et al. *Psychoneuroendocrinology*. 2014;**40**:1-9[20]	Nested cohort study, USA. Rhode Island Child Health Study 'RICHS' cohort. N=444. NICU Network Neurobehavioral Scales (NNNS) used for neurobehavior testing. Tissue: Placenta Platform: Pyromark MD (Qiagen)	*LEP*	Neurobehavior	In male infants increased *LEP* methylation a was associated with membership in the profile of increased lethargy and hypotonicity (OR = 1.9; 95% CI: 1.07–3.4), and reduced risk of membership in the profile of decreased lethargy and hypotonicity (OR = 0.54; 95% CI: 0.3–0.94).
Lester BM, et al. *Epigenomics*. 2015;**7**:1123-36[21]	Cross-sectional study of 67 preterm infants, USA. Tissue: Buccal cells Platform: Pyrosequencing	*HSD2* and *NR3C1*	Neurobehaviour	Infants with a high-risk neurobehavioral profile showed increased *NR3C1* methylation and decreased *HSD2* methylation compared to infants with a low-risk neurobehavioral profile.
Lin, X, et al. *BMC Med*. 2017 Mar 7;**15**:50[22]	Cohort study ('GUSTO' cohort), 987 mother-infant pairs, Singapore. Tissue: Cord blood. Platform: Illumina Infinium HumanMethylation450 BeadChip	*ANK3, CDKN2B, IGDCC4, P4HA3, MIRLET7BHG, CACNA1G* and *ZNF423*	Birthweight, size and adiposity at 4 years	Methylation at *ANK3, CDKN2B, IGDCC4, P4HA3, MIRLET7BHG, CACNA1G* and *ZNF423* were associated with birth weight. *MIRLET7BHG* was the only 'nutrition-sensitive' loci considered in this review, however.

Reference	Study design	Genes Implicated	Phenotype investigated	Summary of results
Murphy R, et al. *BMC Med Genet*. 2014;**15**:67[23]	Cohort study (Auckland Birthweight Collaborative ('ABC') study) following 153 children out of which 80 were SGA. Age 11 years Tissue: whole blood Platform: Methylation-specific multiplex-ligation-dependent probe amplification assay and pyrosequencing.	*IGF2, H19, KCNQ10T1*	Size at birth	*IGF2* DMR0 methylation was 2.7% lower in SGA children. Methylation did not vary at *H19* and *KCNQ10T1* ICRs.
Paquette AG, et al. *Epigenomics*. 2015;**7**:767-79[24]	Rhode Island Child Health Study 'RICHS' cohort . N=547 infants. NICU Network Neurobehavioral Scales (NNNS) used for neurobehavior testing. Tissue: Placenta Platform: Pyromark MD (Qiagen)	*NR3C1, HSD11B2, FKBP5, ADCYAP1R1*	Neurobehaviour	Using maximum likelihood factor analysis, 3 factors were identified as explaining the maximum variability in DNA methylation. *NR3C1* loaded strongly onto a factor that was associated with decreased quality of movement and self-regulation, increased arousal and excitability, and increased non-optimal refluxes and stress abstinence scores, also referred to as a 'poorly regulated' profile. *HSD11B2* methylation loaded onto factor associated with reducing risk of being in the poorly regulated profile.
Qian YY, et al. *J Hum Nutr Diet*. 2016;**29**:643-51[25]	Case-control study. 39 small-for-gestational age (SGA) infants and 49 appropriate-for-gestational-age (AGA) controls, China. Tissue: Umbilical cord blood Platform: EpiTYPER, Sequenom.	*MEST,H19*	Size at birth	Methylation in SGA children was higher (p<0.05) than controls at 3 sites in the *H19* DMR. Six sites in the *H19* DMR had higher methylation in SGA compared to AGA, but only in only males born to mothers who supplemented in pregnancy with folic acid.
Rijlaarsdam J, et al. *J Child Psychol Psychiatry*. 2017;**58**:19-27[26]	Nested cohort study ('ALSPAC' cohort). 164 youth with low conduct problems (CP) and early onset persistent (EOP) ADHD. Tissue: Cord blood, peripheral blood at age 7 Platform: Illumina Infinium HumanMethylation450 BeadChip	*IGF2*	ADHD	In the case of EOP individuals, *IGF2* methylation was positively associated with ADHD symptoms.
Souren NY, et al., *Obesity*. 2011;**19**:1519-22[27]	Monozygotic (MZ) twins study. 16 BMI discordant MZ twin pairs recruited from the East Flanders Prospective Twin Survey (EFPTS), Belgium. Mean age 30.9 years. Tissue: Saliva Platform: PCR amplification followed by a single-nucleotide primer extension reaction	*KvDMR1, H19, IGF2, GRB10, MEST, SNRPN, GNAS*	BMI	Methylation differences of the studied imprinted genes in salivary DNA did not account for the discordance in BMI between twins in this cohort.

Reference	Study design	Genes Implicated	Phenotype investigated	Summary of results
St-Pierre J, et al. *Epigenetics*. 2012;**7**:1125-32[28]	Cohort study. 50 mother-infant pairs, French-Canadian origin, Canada. Tissue: Placenta Platform: Pyromark Q24 (Qiagen)	*IGF2, H19*	Birthweight	*IGF2* DMR2 mean methylation on the fetal placental side was positively correlated with birthweight (Spearman rank correlation=0.44, p< 0.01), height (r=0.40, p< 0.01), head (r=0.32, p< 0.05) and thorax (r=0.32,p < 0.05) circumference. 31% variance in birth weight could be accounted for by *IGF2/H19* genotype and epigenotype jointly.
Steegers-Theunissen RP, et al. *PLoS One*. 2009;4: e784[29]	Cross sectional study amongst controls for another study ('HAVEN'). 120 mother-infant pairs, The Netherlands. Tissue: Infant blood taken at 17 months. Platform: Epityper, Sequenom	*IGF2*	Birth weight	*IGF2* methylation was inversely associated with birthweight (-1.7% methylation per SD birthweight; p=0.034).
Tobi EW, et al. *Epigenetics*. 2011;**6**:171-6[30]	Cohort study (The Dutch Project on Preterm and Small for Gestational Age Infants ('POPS')). 38 small for gestational age (SGA) and 75 appropriate for gestational age (AGA) individuals studied. Tissue: Whole blood at age 19. Methylation analysis by Platform: Epityper, Sequenom.	*IGF2, GNASAS, INSIGF, LEP*	Size at birth	Methylation levels of *IGF2* DMR0, *GNASAS*, *INSIGF* and *LEP* did not significantly vary between SGA and AGA individuals at 19 years of age.
Wijnands KP, et al. *Nutr Metab Cardiovasc Dis.* 2015;**25**:608-14[31]	Cohort study. 120 healthy children at 17 months of age, Netherlands. Tissue: Whole blood Platform: EpiTyper, Sequenom.	*TNFα, LEP*	Lipid profile	High-density lipoprotein-cholesterol levels in the children were inversely associated with *TNFα* methylation (-6.1%, p =0.058) and *LEP* (-3.4%, p=0.021)

Abbreviations: ADHD, attention-deficit/hyperactivity disorder; BMI, body mass index; CI, confidence interval; DMR, differentially methylated region; ICR, imprinting control region; LDL-c, low-density lipoprotein cholesterol; OR, odds ratio; SD, standard deviation; VMR, variably methylated region.

References

1.   Azzi S, Sas TCJ, Koudou Y, et al. (2014) Degree of methylation of ZAC1 (PLAGL1) is associated with prenatal and post-natal growth in healthy infants of the EDEN mother child cohort. *Epigenetics* **9**, 338–45.

2.   Bens S, Haake A, Richter J, et al. (2013) Frequency and characterization of DNA methylation defects in children born SGA. *Eur. J. Hum. Genet.* **21**, 838–43.

3.   Bouwland-Both MI, van Mil NH, Stolk L, et al. (2013) DNA methylation of IGF2DMR and H19 is associated with fetal and infant growth: the generation R study. *PLoS One* **8**, e81731.

4.   Burris HH, Braun JM, Byun H, et al. (2013) Association between birth weight and DNA methylation of IGF2, glucocorticoid receptor and repetitive elements LINE-1 and Alu. *Epigenomics* **5**, 271–81.

5.   Córdova-Palomera A, Alemany S, Fatjó-Vilas M, et al. (2014) Birth weight, working memory and epigenetic signatures in IGF2 and related genes: a MZ twin study. *PLoS One* **9**, e103639.

6.   Deodati A, Inzaghi E, Liguori A, et al. (2013) IGF2 methylation is associated with lipid profile in obese children. *Horm. Res. Paediatr.* **79**, 361–7.

7.   Drake AJ, McPherson RC, Godfrey KM, et al. (2012) An unbalanced maternal diet in pregnancy associates with offspring epigenetic changes in genes controlling glucocorticoid action and foetal growth. *Clin. Endocrinol.* **77**, 808–15.

8.   Dunstan J, Bressler JP, Moran TH, et al. (2017) Associations of LEP, CRH, ICAM-1, and LINE-1 methylation, measured in saliva, with waist circumference, body mass index, and percent body fat in mid-childhood. *Clin. Epigenetics* **9**, 29..

9.   García-Cardona MC, Huang F, García-Vivas JM, et al. (2014) DNA methylation of leptin and adiponectin promoters in children is reduced by the combined presence of obesity and insulin resistance. *Int. J. Obes.* **38**, 1457–65.

10.   Godfrey KM, Sheppard A, Gluckman PD, et al. (2011) Epigenetic Gene Promoter Methylation at Birth Is Associated With Child's Later Adiposity. *Diabetes* **60**, 1528–1534.

11.   Harvey NC, Sheppard A, Godfrey KM, et al. (2014) Childhood bone mineral content is associated with methylation status of the RXRA promoter at birth. *J. Bone Miner. Res.* **29**, 600–7.

12.   Houde A-A, Légaré C, Biron S, et al. (2015) Leptin and adiponectin DNA methylation levels in adipose tissues and blood cells are associated with BMI, waist girth and LDL-cholesterol levels in severely obese men and women. *BMC Med. Genet.* **16**, 29.

13. Hoyo C, Fortner K, Murtha AP, et al. (2012) Association of cord blood methylation fractions at imprinted insulin-like growth factor 2 (IGF2), plasma IGF2, and birth weight. *Cancer Causes Control* **23**, 635–645.

14. Hoyo C, Daltveit AK, Iversen E, et al. (2014) Erythrocyte folate concentrations, CpG methylation at genomically imprinted domains, and birth weight in a multiethnic newborn cohort. *Epigenetics* **9**, 1120–30.

15. Huang R-C, Galati JC, Burrows S, et al. (2012) DNA methylation of the IGF2/H19 imprinting control region and adiposity distribution in young adults. *Clin. Epigenetics* **4**, 21.

16. Kappil MA, Green BB, Armstrong DA, et al. (2015) Placental expression profile of imprinted genes impacts birth weight. *Epigenetics* **10**, 842–9.

17. Kuehnen P, Mischke M, Wiegand S, et al. (2012) An Alu element-associated hypermethylation variant of the POMC gene is associated with childhood obesity. *PLoS Genet.* **8**, e1002543

18. Kühnen P, Handke D, Waterland RA, et al. (2016) Interindividual Variation in DNA Methylation at a Putative POMC Metastable Epiallele Is Associated with Obesity. *Cell Metab.* **24**, 502–9.

19. Lesseur C, Armstrong DA, Paquette AG, et al. (2013) Tissue-specific Leptin promoter DNA methylation is associated with maternal and infant perinatal factors. *Mol. Cell. Endocrinol.* **381**, 160–7.

20. Lesseur C, Armstrong DA, Murphy MA, et al. (2014) Sex-specific associations between placental leptin promoter DNA methylation and infant neurobehavior. *Psychoneuroendocrinology* **40**, 1–9.

21. Lester BM, Marsit CJ, Giarraputo J, et al. (2015) Neurobehavior related to epigenetic differences in preterm infants. *Epigenomics* **7**, 1123–36.

22. Lin X, Lim IY, Wu Y, et al. (2017) Developmental pathways to adiposity begin before birth and are influenced by genotype, prenatal environment and epigenome. *BMC Med.* **15**, 50.

23. Murphy R, Thompson JM, Tost J, et al. (2014) No evidence for copy number and methylation variation in H19 and KCNQ10T1 imprinting control regions in children born small for gestational age. *BMC Med. Genet.* **15**, 67.

24. Paquette AG, Lester BM, Lesseur C, et al. (2015) Placental epigenetic patterning of glucocorticoid response genes is associated with infant neurodevelopment. *Epigenomics* **7**, 767–779.

25. Qian Y-Y, Huang X-L, Liang H, et al. (2016) Effects of maternal folic acid supplementation on gene methylation and being small for gestational age. *J. Hum. Nutr. Diet.* **29**, 643–651.

26. Rijlaarsdam J, Cecil CAM, Walton E, et al. (2017) Prenatal unhealthy diet, insulin-like growth factor 2 gene (IGF2) methylation, and attention deficit hyperactivity disorder symptoms in youth with early-onset conduct problems. *J. Child Psychol. Psychiatry.*

58, 19–27.

27.   Souren NYP, Tierling S, Fryns J-P, et al. (2011) DNA methylation variability at growth-related imprints does not contribute to overweight in monozygotic twins discordant for BMI. *Obesity.* **19**, 1519–22.

28.   St-Pierre J, Hivert M-F, Perron P, et al. (2012) IGF2 DNA methylation is a modulator of newborn's fetal growth and development. *Epigenetics* **7**, 1125–32.

29.   Steegers-Theunissen RP, Obermann-Borst SA, Kremer D, et al. (2009) Periconceptional maternal folic acid use of 400 microg per day is related to increased methylation of the IGF2 gene in the very young child. *PLoS One* **4**, e7845.

30.   Tobi EW, Heijmans BT, Kremer D, et al. (2011) DNA methylation of IGF2, GNASAS, INSIGF and LEP and being born small for gestational age. *Epigenetics* **6**, 171–6.

31.   Wijnands KPJ, Obermann-Borst SA & Steegers-Theunissen RPM (2015) Early life lipid profile and metabolic programming in very young children. *Nutr. Metab. Cardiovasc. Dis.* **25**, 608–14.

ANNEX 3.5 Textbook Chapter: Epigenetics, Nutrition and Infant Health

RESEARCH PAPER COVER SHEET

SECTION A – Student Details

Student	Philip James
Principal Supervisor	Dr Matt Silver
Thesis Title	Maternal nutrition, seasonality and epigenetics: an exploration of one-carbon metabolism and a novel nutritional supplement design in The Gambia.

SECTION B – Paper already published

Where was the work published?	Text book: The Biology of the First 1000 Days		
When was the work published?	20th September, 2017		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N/A		
Have you retained the copyright for the work?*	No	Was the work subject to academic peer review?	No, editorial only

If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.
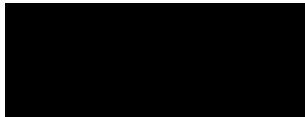
Confirmation Number: 11756503
Order Date: 10/15/2018
Book Title: The Biology of the First 1,000 Days
Order detail ID: 71609779
ISBN: 9781498756792
Publication Type: Book
Publisher: Taylor and Francis
Author/Editor: Kraemer, Klaus; Green, Tim J.; Whitfield, Kyly C.
Permission Status: Granted
Permission type: Republish or display content
Type of use: Thesis/Dissertation
Order License Id: 4450231223812
Requestor type: Author of requested content
Format: Print, Electronic; Portion: chapter/article
Number of pages: 19

Multi-authored Work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper.

Andrew Prentice and I were originally approached by Sight and Life to write a textbook chapter on epigenetics and growth. I replied to the editors with an outline of this chapter, which they agreed would work well for a chapter. I drafted the chapter, and incorporated comments from Matt Silver and Andrew Prentice for the final version.

Student Signature:

Date: 16th October 2018

Supervisor Signature:

Date: 16th October 2018

22 Epigenetics, Nutrition, and Infant Health

Philip T. James, Matt J. Silver,
and Andrew M. Prentice

CONTENTS

INTRODUCTION

The field of epigenetics is currently garnering a great deal of interest, exploring how our very molecular makeup in the form of modifications to the genome can be altered by factors as diverse as aging, disease, nutrition, stress, alcohol, and exposure to pollutants. Epigenetic changes have previously been implicated in the etiology of a variety of diseases [1], notably in the development of certain cancers [2], and inherited growth disorder syndromes [3], but the exploration of epigenetics' role in fetal programming is still in its infancy. This chapter focuses on how nutritional exposures during pregnancy may affect the infant epigenome, and the impact that such modifications may have on the long-term health of the child. We start by describing some keys concepts in epigenetics and discuss windows of epigenetic plasticity in the context of the *developmental origins of health and disease* (DOHaD) hypothesis. We then review some of the key mechanisms by which nutrition can affect the epigenome, with a particular focus on the role of one-carbon metabolism. We finish by outlining some of the child health outcomes that have been linked to epigenetic

dysregulation, and discuss possible next steps that need to be realized if insights into the basic science of epigenetics are to be translated into tangible public health benefits.

A BRIEF PRIMER ON EPIGENETICS

Epigenetic processes describe changes to the genome that can alter gene expression without changing the underlying DNA sequence [4]. These changes are mitotically heritable, and involve the interplay of DNA methylation, histone modifications, and RNA-based mechanisms (Figure 22.1). DNA methylation most commonly occurs at loci where a cytosine is found next to a guanine on a DNA strand along its linear sequence, hence termed *cytosine-phosphate-guanine* or *CpG sites*. It involves the covalent bonding of a methyl (CH_3) group to the cytosine at the 5′ carbon position to form 5-methylcytosine. CpGs found in high densities are termed *CpG islands*. Roughly two-thirds of human genes contain CpG islands in their promoter regions, although repetitive elements in the genome can also contain many CpG sites [5]. CpGs are generally methylated in nonpromoter regions and unmethylated at promoter regions. Methylation at CpG sites in promoters is usually associated with transcriptional silencing, although not consistently [6]. Although methylation is the most studied chemical alteration to date, others modifications (e.g., hydroxymethylation) can occur at cytosine bases [7].

DNA methylation is catalyzed by DNA methyltransferases (DNMTs). Mammals have three types of DNMT. DNMT1 recognizes hemimethylated DNA; therefore, after DNA replication and cell division, it methylates the newly synthesized strand to
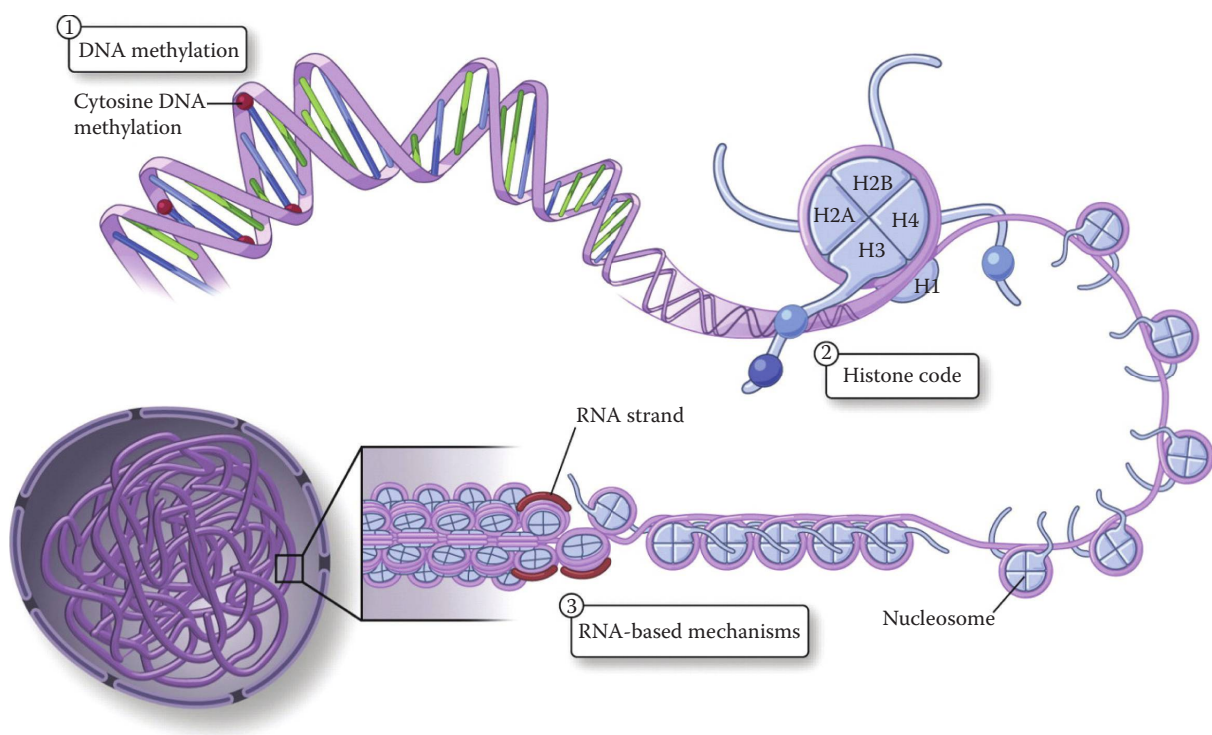
FIGURE 22.1 An overview of epigenetic mechanisms. (From Yan MS-C, Matouk CC, Marsden PA, *J Appl Physiol* 2010, 109:916–26.)

maintain methylation patterns of the original template strand (maintenance methylation). DNMT3a and 3b appear to primarily methylate fully unmethylated CpG sites (*de novo* methylation) [7]. DNA methylation is important for a host of biological processes, including transcriptional silencing, X-chromosome inactivation, genomic imprinting, and the maintenance of cellular identity by enabling tissue-specific gene expression [8].

DNA is tightly woven around histone proteins, forming compact complexes of DNA and protein called nucleosomes. Nucleosomes, in turn, are packed together to form chromatin. Within the nucleosomes, the histone proteins are arranged in an eight-part formation, comprising two copies each of histones H2A, H2B, H3, and H4. Histone modifications involve various posttranslational chemical alterations to the amino acids of the histone tails, including acetylation of lysine, methylation of lysine and arginine, phosphorylation of serine and threonine, and the ubiquitination of lysine [9]. There are several mechanisms by which chemical modifications of CpG sites and histones are thought to influence gene expression. The methyl group from 5-methylcytosine may block transcription factors either directly or through the recruitment of a methyl-binding protein. Alternatively, the DNMT enzymes acting on CpG sites may be physically linked to other enzymes, which bring about histone methylation and deacetylation [10]. Although chromatin remodeling is intricately controlled, a simplified summary is that the hyperacetylation of histones and hypomethylation of histones and CpGs is associated with a euchromatin (open) configuration, generally associated with facilitation of transcriptional activity. Conversely, hypoacetylation of histones and hypermethylation of histones and CpGs is associated with a heterochromatin (closed) structure and transcriptional repression [11].

Although noncoding RNAs do not code for proteins, many are functional and may affect gene expression [12]. Of those that influence gene expression, microRNAs (miRNAs) have been the most studied to date. These are short pieces (~22 nucleotides) of RNA that affect the epigenome through binding to target mRNAs controlling the expression of key regulators such as DMNTs and histone deacetylases [13]. In turn, CpG methylation and histone modifications can influence the transcription of certain miRNA classes [14]. MicroRNAs may also affect gene expression directly by binding to messenger RNAs, repressing their translation [14].

TIME POINTS OF PLASTICITY IN THE EPIGENOME

Times of increased cell turnover, such as during fetal development and infancy, may be particularly susceptible both to epigenetic errors and to environmental influences [15]. In this chapter, we focus on DNA methylation and the *in utero* period, to include periconception, since this falls within the first 1,000 days window, and is also a period of exceptionally rapid cell differentiation and complex epigenetic remodeling (Figure 22.2). In the first 48 hours after fertilization, there is rapid (active) demethylation of the paternal genome and a slower (passive) demethylation of the maternal genome [16]. Erasing the epigenetic marks in the zygote prior to the blastocyst stage is important to enable pluripotency of the developing cells [8]. Imprinted genes and some retrotransposons (defined later) are known to resist demethylation at this stage [17]. Remethylation then occurs in tissue-specific patterns after implantation,
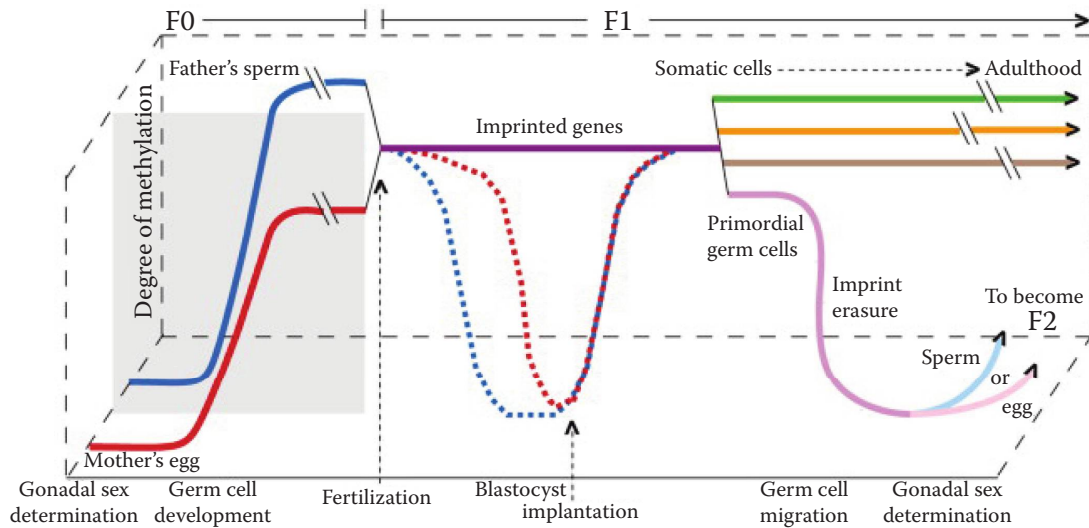
FIGURE 22.2 Epigenetic remodeling in embryogenesis. (From Perera F, Herbstman J, *Reprod Toxicol* 2011, 31:363–73.)

during the process of gastrulation and differentiation of the somatic cells throughout pregnancy. A second wave of demethylation occurs during the epigenetic reprogramming of primordial germ cells (PGCs) in the developing embryo at the point of their migration to the genital ridge [18]. Parental imprints are erased at this stage in preparation for the laying down of sex-specific imprints in the PGCs. Remethylation of sperm cells occurs before the birth of the child, and in oocytes over the duration of their maturation [18]. The periconceptional period is therefore one of huge dynamism in the methylome, representing a window in which epigenetic errors could have significant consequences for the health of the child.

WHICH PARTS OF THE EPIGENOME ARE MOST SUSCEPTIBLE TO ENVIRONMENTAL INFLUENCES?

There are some regions of the genome that demonstrate increased interindividual epigenetic variation, and may be particularly vulnerable to the impact of environmental influences [19]. These include imprinted genes, metastable epialleles, and transposable elements. *Imprinted genes* show monoallelic expression, whereby only the maternally or paternally inherited allele is expressed [20]. If a gene is "paternally expressed" it means the expressed allele comes from the father and the maternal allele is imprinted (silenced). In regard to growth, paternally expressed genes tend to promote *in utero* growth, whereas maternally expressed genes restrict growth, and this forms the basis of the "parental conflict theory" [21]. *Metastable epialleles* (MEs) are genomic loci whose methylation state varies between individuals, but where variation is correlated across tissues originating from all germ layers in a single individual [22]. This indicates that the marks have been laid down in the first few days after conception before cell types start to specialize. MEs therefore provide a useful device to study the influence of the periconceptional environment, including maternal nutrition, on the offspring epigenome [23,24]. *Transposable elements* (TEs) are small pieces of DNA (usually of viral origin earlier in human history) that are

mobile and can insert into new chromosomal locations throughout the genome. They are thought to make up more than half of the human genome [25]. TEs arise either through the use of RNA as an intermediate for transposition (DNA is transcribed to RNA, reverse transcribed to DNA, and then inserted by reverse transcriptase to form a retrotransposon), or through the complete DNA sequence being cut and pasted directly (forming DNA transposons, which are less common in the human genome). TEs are potentially functionally disruptive, for example, if transposed into a functional gene or when increasing copy number, and this may be one reason why most are silenced epigenetically [26]. Some TEs are thought to be vulnerable to the influence of nutrition at key time points [27], and their variable methylation patterns have been shown to affect neighboring gene expression, most notably in the Agouti mouse experiments detailed later in this chapter.

EPIGENETICS AND THE DEVELOPMENTAL ORIGINS OF HEALTH AND DISEASE

The DOHaD hypothesis describes the idea that environmental insults experienced early in life can increase the risk of adverse health outcomes throughout the life course. DOHaD grew out of David Barker's seminal work following a cohort born in Hertfordshire between 1911 and 1930. His early studies found an inverse relationship between birth weight and blood pressure at age 10, with a stronger association at age 36 [28]. His findings soon expanded to identifying associations between low birth weight and adult-onset chronic disease [29], patterns that were also seen in different cohorts such as the Nurses' Health Study I & II [30] and in Helsinki [31].

Although low birth weight provides a useful proxy for an adverse intrauterine environment, subsequent studies have attempted to pinpoint time points of vulnerability more precisely. Data from famine studies have been particularly useful in this respect. The Dutch Hunger Winter occurred toward the end of World War II, when the Western Netherlands was under German control from November 1944 to May 1945. Nazi blockades cut off food and fuel, and the siege was coupled with a harsh and early winter. Calorie intake varied from 500 to 1,000 kcal per day, depending on the area and time period. An estimated 4.5 million people were affected, 20,000 of whom died [32]. Exposure to famine during pregnancy has been associated with a wide range of offspring phenotypes, from lower birth weight [33] to increased adult blood pressure [34], obesity [35], and risk of schizophrenia [36]. Furthermore the specific timing of famine exposure *in utero* appears to have different programming effects. For example, exposure in midgestation is associated with a doubling of the prevalence of obesity for men aged 18 to 19, yet exposure in the third trimester is associated with lower obesity [35]. Many of these associations are also found from records spanning the Chinese Great Leap Forward, where famine was particularly severe from 1959 to 1961. For example, famine exposure *in utero* is associated with a doubled risk of schizophrenia in later life [37], as well as increased hyperglycemia at age 41 to 42, a trend that is exacerbated if an affluent diet was consumed later in life [38].

Although these two famine studies provide rich epidemiological evidence that nutritional exposures in early life are associated with later disease risk, they are unable to determine precise causal factors, for example, if the effects are driven by

depleted maternal energy, deficient levels of certain micronutrients, or a combination of both. The Pune Maternal Nutrition Study (PMNS) in the Maharashtra State of West-Central India sheds some light on more specific nutritional exposures that may be relevant for DOHaD mechanisms. PMNS is a prospective cohort of 797 women, followed pre- and throughout pregnancy, and their offspring. Higher maternal intake of green leafy vegetables, milk, and fruit, and higher erythrocyte folate concentrations, were associated with larger size babies [39]. Higher maternal folate status at 28 weeks gestation predicted higher offspring adiposity and insulin resistance at the age of 6, low maternal vitamin B_{12} status was associated with offspring insulin resistance, and a combination of high maternal folate and low B_{12} produced the strongest associations [40].

Attention is now shifting to an investigation of the mechanisms that may mediate the observations above. In this respect, epigenetic modifications to the genome are emerging as a leading candidate that may, at least partially, explain some of the observations described in the DOHaD literature. For example, Heijmans et al. selected 60 individuals conceived during the Dutch Hunger Winter, and compared them with nonexposed siblings. Five CpG sites within *IGF2* (a maternally imprinted gene controlling fetal growth) were less methylated in individuals exposed to famine at periconception, but there was no difference in methylation for individuals exposed to famine in late gestation [41]. Indeed, the most recent study from the same group confirms that exposure to famine in the first 10 weeks of gestation shows a greater signature in the adult blood methylome (58–59 years) compared to exposure later in gestation [42]. Whilst the exact mechanisms are so far unknown, one concept being debated is the "thrifty epigenome" hypothesis. This proposes that early life insults mold an epigenetic signature to program a phenotype that is "adapted" to the intrauterine environment, which may be problematic if the environment into which the child is born then changes [43]. Under this hypothesis, malnutrition in pregnancy could program "thrifty" epigenotypes, designed to reduce metabolic rate and store energy in an attempt to adapt to a nutritionally poor environment, but may subsequently trigger symptoms of metabolic disease if the environment changes to one of relative nutritional abundance.

MATERNAL NUTRITION EXPOSURES AND THE INFANT EPIGENOME

There are a variety of factors, both nutrition-related and otherwise, which may impact the infant epigenome *in utero* through maternal exposure (Figure 22.3). Of these exposures, particular attention has been paid to the role of one-carbon metabolites in the periconceptional period and during embryonic development [44]. In the following section, we give an overview of one-carbon metabolism, and discuss evidence for the influence of one-carbon metabolites on the infant epigenome. We also briefly review other maternal nutrition exposures that have generated research in a human intergenerational setting: maternal BMI, polyunsaturated fatty acids, and vitamin C.

ONE-CARBON METABOLISM

The one-carbon pool is composed of one-carbon units and the two main carriers that activate, transport, and transfer these units: tetrahydrofolate (THF) and S-adenosyl
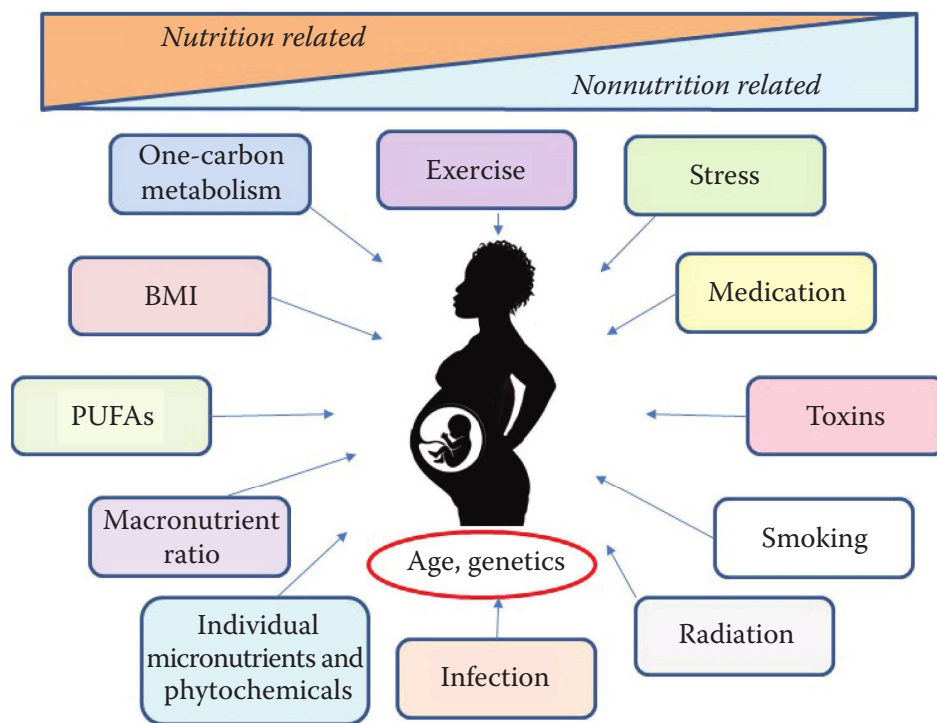
FIGURE 22.3 Potential maternal factors impacting the fetal epigenome.

methionine (SAM). One-carbon units are used as substrates for a whole range of intricate biochemical processes, including cellular biosynthesis, redox status regulation, and genome maintenance through the regulation of nucleotide pools. However, it is their role in transmethylation reactions that is central to the notion of diet-epigenome interplay, and this requires an understanding of how folate, methionine, homocysteine, transsulfuration, and transmethylation metabolic pathways interact (Figure 22.4).

S-adenosyl methionine (SAM) methylates a wide variety of acceptors in reactions catalyzed by methyl transferases. Over 200 methylation reactions are required for transcription, translation, protein localization, and signaling purposes [45], but it is the methylation of cytosine bases and amino acids on histone tails that play a role in epigenetics. The donation of SAM's methyl group forms S-adenosyl homocysteine (SAH), which is further hydrolyzed to homocysteine (Hcy), where it is maintained in an equilibrium state that thermodynamically favors SAH over Hcy [46]. A buildup of Hcy results in an increase in SAH, which in turn impedes methylation reactions since SAH competes with SAM for the active site on methyl transferase enzymes [47]. The SAM:SAH ratio is therefore often used as a proxy indicator of methylation potential [48]. In order to maintain favorable methylation conditions, Hcy has to be removed from the system. One way in which this can happen is by accepting a methyl group to form methionine, which can then in turn be condensed with ATP to form SAM and continue the cycle. Hcy can also be removed through its irreversible degradation to cystathionine and cysteine in the transsulfuration pathway requiring vitamin B_6.

The methylation of Hcy to methionine uses two distinct pathways. The major pathway is the vitamin B_{12}-dependent reaction involving folate metabolic pathways. A stepwise reduction of dietary folates and folic acid forms tetrahydrofolate (THF).
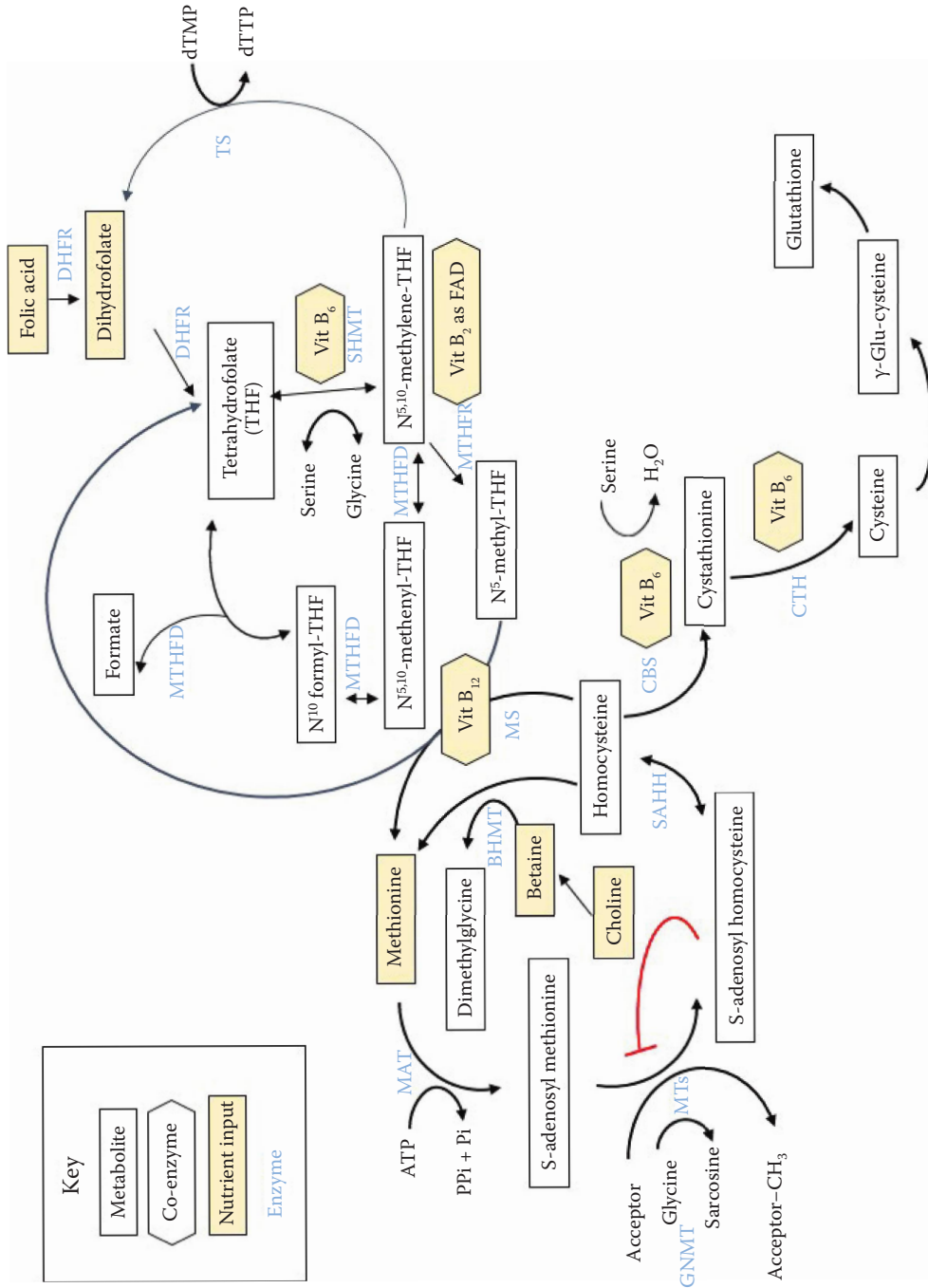
FIGURE 22.4   A simplified summary of one-carbon metabolism. BHMT, betaine homocysteine methyltransferase; CBS, cystathionine-beta-synthase; CTH, cystathionine gamma-lyase; DHFR, dihydrofolate reductase; dTMP, deoxythymidine monophosphate; dTTP, deoxythymidine triphosphate; FAD, flavin adenine dinucleotide; GNMT, glycine N-methyltransferase; MAT, methionine adenosyltransferase; MS, methionine synthase; MT, methyl transferases; MTHF, methylenetetrahydrofolate reductase; MTHFD, methylenetetrahydrofolate dehydrogenase; SAHH, S-adenosyl homocysteine hydrolase; SHMT, serine hydroxymethyltransferase; TS, thymidylate synthase.

The one-carbon units attach to the N^5 or N^{10} position, or bridge the two. THF can be converted into $N^{5,10}$-methylene-THF using a methyl group primarily from serine with vitamin B_6 as a cofactor. $N^{5,10}$-methylene-THF is either used in the thymidylate synthase pathway for DNA synthesis and repair, interconverted to $N^{5,10}$-methenyl-THF, N^{10}-formyl-THF (used in purine synthesis) and formate (an important one-carbon donor) forms, or irreversibly reduced to N^5-methyl tetrahydrofolate (methyl-THF) through the action of 5,10-methylene THF reductase (MTHFR) and vitamin B_2 as a cofactor. It is methyl-THF that then provides the methyl group for the methylation of Hcy, and in the process regenerates THF. The alternative pathway for the methylation of Hcy, predominantly used in the liver and kidneys, uses the methyl group from betaine, a product formed through the oxidation of choline. This description and Figure 22.4 give a very simplified overview of the complex and interlinking metabolic pathways involved, some of which occur with different enzymes in different cellular compartments [49]. However, this brief overview helps explain why deficiencies in the key methyl donors of choline, betaine, or folate, alongside deficiencies in vitamins B_2, B_6, and B_{12}, which play essential coenzyme roles, may disrupt the metabolic pathways that are responsible for DNA methylation. Genetic variants in proteins coding for enzymes used in key one-carbon metabolic processes can also affect activity at critical nodes and alter the flow of metabolites [50].

STUDIES INVESTIGATING ONE-CARBON METABOLITES AND EPIGENETICS

Perhaps the most famous animal experiments demonstrating how epigenetic changes driven by maternal diet in pregnancy can dramatically alter phenotype in the offspring come from the agouti mouse. In one experiment, pregnant dams were fed a diet that varied in methyl donor content (folic acid, choline, betaine, and vitamin B_{12}). Their isogenic pups showed variable methylation at an intracisternal A particle (IAP), a retrotransposon upstream of the *agouti* gene that is a metastable epiallele. The degree of methylation at this locus altered expression of the *agouti* gene, resulting in permanent phenotypic differences. The most obvious change was in fur color, but differences were also found in appetite, adiposity, and glucose tolerance, factors highly relevant to life-long chronic disease risk [19,27]. A similar experiment in a different strain of kinky-tailed mice showed that methyl donor content of the maternal diet also altered methylation at an IAP on the *Axin* gene, producing pups with varying levels of tail kink [51].

Human studies exploring associations between individual one-carbon metabolites and epigenetic effects are few and far between, especially those involving an intergenerational study design. Homocysteine and vitamin B_6 are two examples that are described in human cross-sectional studies assessing immediate effects on health at one time point within a single generation. Hyperhomocysteinemia has been associated with hypomethylation in several studies [52] supporting the hypothesis that epigenetic mechanisms may play a partial role in observed associations between hyperhomocysteinemia and decreased cardiovascular health via inflammation, free radical formation, and atherogenesis [53]. Low vitamin B_6 levels may decrease $N^{5,10}$-methenyl-THF through reduced serine hydroxymethyltransferase (SHMT) activity (Figure 22.4), stressing the thymidylate synthase pathway for DNA repair and incorporating more

uracil into DNA. This reduces genome stability through increased chromosome breaks, which, in turn, have been associated with increased levels of global DNA hypomethylation alongside a putative increased risk of tumor development [54].

Of the human studies using an intergenerational design, folate, vitamin B_{12}, and choline have been most frequently investigated. Due to its role in preventing neural tube defects and involvement in some cancer therapies, folic acid has been extensively reviewed [55]. Increased periconceptional folic acid consumption has been associated with increased methylation in infants at a differentially methylated region (DMR) of the *IGF2* gene, although these associations have not been established consistently [56]. Additional studies have found that maternal folate levels in late pregnancy influence offspring methylation patterns in the imprinting control region of *IGF2* [57] and also in the imprinting regulator *ZFP57* [58]. Despite the timing of maternal folate measurements, these findings support the importance of nutritional status around conception, since many DMRs appear to have their methylation pattern established prior to gastrulation [59]. There are fewer studies assessing the effect of maternal vitamin B_{12} status on the infant epigenome. However, preliminary evidence suggests that maternal vitamin B_{12} levels at first antenatal visit are inversely correlated with infant cord blood global methylation levels [60]. For choline, one intervention study investigated the effect of supplementing mothers in the third trimester with daily choline of 480 mg versus 930 mg [61]. Higher maternal choline intake was associated with increased methylation of *CRH* and *NR3C1* promoter regions within fetal placental tissue.

The complex interlinking metabolic pathways involved in one-carbon metabolism may be more appropriately investigated by exploring the joint effects of relevant methyl donors and coenzymes. Researchers in The Gambia have been exploring human diet–epigenome interactions by exploiting a natural experimental design, in which fluctuations in energy balance and maternal nutritional exposures display a distinct bimodal seasonal pattern. The Gambia experiences a rainy ("hungry") season from July to September (Figure 22.5a), a planting season with increased energy expenditure, depleted food stores, and peaks of malarial and diarrheal diseases. The dry ("harvest") season occurs from February to April (Figure 22.5b), when

(a)   (b)
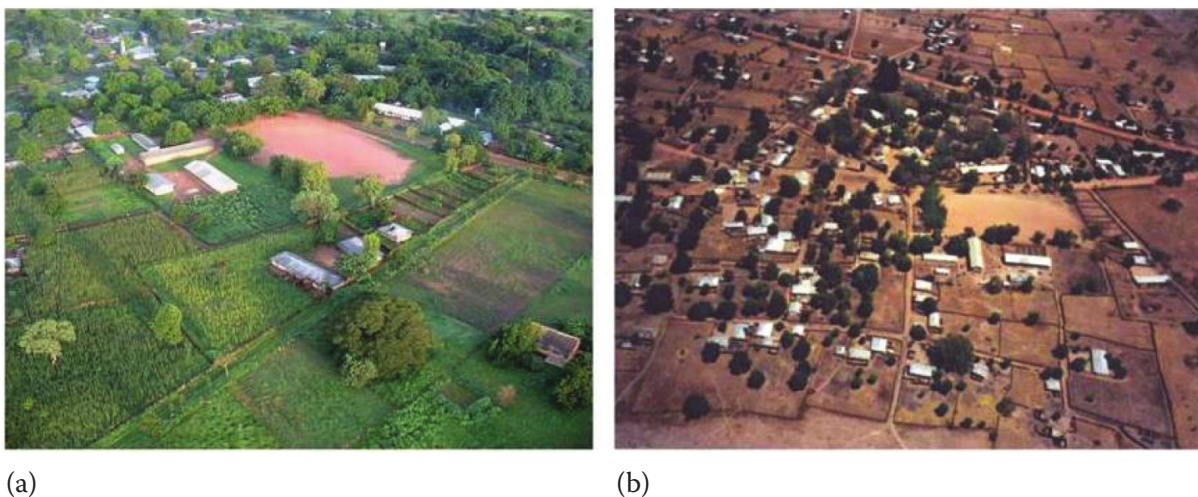
FIGURE 22.5   The Gambian (a) rainy/"hungry" and (b) dry/harvest seasons in Keneba, West Kiang. (Photo by Andrew M. Prentice.)

harvesting takes place, leading to improved food security. Despite overall household food stores being more replete in the dry season, the rainy season offers the increased availability of certain micronutrients made available from rain-fed green leafy vegetables. It has been known for several decades that Gambian children born during the rainy season are up to 10 times more likely to die prematurely in young adulthood than those born during the dry season [62]. However, the processes underlying these findings have yet to be fully delineated, and nutrition-related epigenetic regulation in early embryonic development is emerging as one plausible contributory mechanism. A recent study compared women conceiving in the peak of the dry and the peak of the rainy season [24]. Maternal periconceptional plasma concentrations of folate, vitamin B_2, methionine, betaine, and the SAM:SAH and betaine:dimethyl glycine ratios were higher in the rainy season, and concentrations of vitamin B_{12}, homocysteine, and SAH were lower. This suggested that the maternal metabolome during the rainy season contained higher concentrations of methyl donors and exhibited a higher methylation potential than the dry season metabolome. Indeed, offspring of these rainy season conceptions had higher levels of CpG methylation at six MEs in peripheral blood monocytes compared to the offspring of those conceived in the dry season, with similar patterns in hair follicle DNA [24]. A subsequent study details the genome-wide search for MEs susceptible to the periconceptional environment using two independent approaches [63]. Both found *VTRNA2-1* as their top hit, an imprinted gene that was also found to exhibit increased hypomethylation among Gambian offspring conceived in the dry season compared to the rainy season. Taken together, the results support the hypothesis that maternal nutrition during the periconceptional period can impact the infant epigenome in humans and has triggered further research into the possible phenotypic consequences.

POLYUNSATURATED FATTY ACIDS

Several studies indicate that increased consumption of ω-3 PUFAs is associated with reduced homocysteine levels [64]. In the context of early life nutrition, maternal fatty acid intake influences infant fatty acid composition via placental transfer and breast milk, with maternal PUFA intake potentially affecting infant appetite control, neuroendocrine function, and metabolic programming [65]. Investigating whether one-carbon pathways and epigenetic mechanisms are involved is of increasing interest. In rodent models, providing a maternal diet with excess folic acid and restricted vitamin B_{12} (designed to reflect the nutritional situation in much of rural India) is associated with increased maternal oxidative stress, lower level of offspring placental and brain docosahexaenoic acid, and decreased placental global methylation [66]. However, maternal supplementation with PUFAs seems to partially ameliorate the disrupted one-carbon metabolism associated with this diet [67]. One potential mechanism is through PUFAs upregulating enzymes responsible for the methylation of homocysteine to methionine [68], which in theory could increase the SAM:SAH ratio and help overcome the impact of reduced MTHFR activity brought about by the restricted vitamin B_{12}. An additional recent hypothesis describes how a decreased maternal PUFA intake may increase the availability of methyl groups for DNA methylation since there would be reduced PUFAs for phosphatidylcholine synthesis and therefore less demand for methyl groups for this

pathway. The disruption to the one-carbon pathways brought about by low PUFA levels is thought to alter epigenetic programming of placental genes, increasing the risk of aberrant methylation in offspring and adverse pregnancy outcome [69]. Human evidence at the intergenerational level, however, is currently limited to a study by Lee et al. [70]. Here, mothers were given either 400 mg of ω-3 PUFA (n = 131) or placebo (n = 130) daily from approximately 18 weeks gestation to birth. Among mothers supplemented with PUFAs, there was increased global methylation (measured at LINE-1 repetitive elements) in offspring of mothers who smoked (n = 26 case, 26 control) [71], and increased offspring *IGF2* promoter 3 methylation in preterm infants (n = 16 case, 20 control) [70].

Body Mass Index

Maternal body mass index (BMI) shows some association with infant epigenetic patterns [72]. However, it can be difficult to define exactly what BMI represents as an exposure, since it does not always correlate well with body composition [73] and is a proxy for a wide range of potential metabolic disorders [74]. In an intergenerational setting, both maternal underweight and overweight is associated with epigenetic patterns in infant cord blood [72]. The mechanism by which BMI influences epigenetics is so far unknown. Although SAM has been shown to be independently, positively associated with fat mass in older adults, alongside a higher rate of converting methionine into SAM in obese individuals, the absence of a rise in the SAM:SAH ratio means conclusions are hard to draw [75]. Paternal BMI may also be a relevant exposure to consider, since neonates born to obese fathers in a U.S. study showed hypomethylation at some loci within imprinted genes (*MEST*, *PEG3*, *NNAT*), independent of maternal obesity [76].

Vitamin C and Demethylation

As outlined earlier, the periconceptional period is a time of widespread remodeling of the offspring epigenome, including rapid erasure (demethylation) of parental epigenetic marks and subsequent remethylation (Figure 22.2). In demethylation, 5-methylcytosine (5mC) can be sequentially oxidized to different states by ten-eleven translocation (TET) dioxygenases that use vitamin C (ascorbate) as a cofactor [77]. *In vitro* studies, for example, indicate that adding vitamin C to mouse or human embryonic stem cells increases activity of TET enzymes, with active demethylation seen in the germline and associated gene expression changes [78]. Vitamin C is therefore an important factor to consider in epigenetic processes during early embryogenesis, although *in vivo* studies will be required to explore whether vitamin C deficiency could play a role in aberrant demethylation.

In this section we have identified some of the key ways that maternal nutrition might influence the offspring epigenome. Other exposures, both nutrition-related and otherwise, may also be involved. These include maternal stress [79], toxin exposure [80], maternal hyperglycemia [81], the microbiome [82], dietary polyphenols [83], vitamin D [84], vitamin A [85], and infection [86]. As evidence builds, it will be necessary to explore how these multiple exposures work together to influence the

infant epigenome and to more clearly define the potential consequences for human health. It is to the latter point that we now turn.

PHENOTYPES ASSOCIATED WITH THE IMPACT OF MATERNAL EXPOSURES ON THE INFANT EPIGENOME

Most population-based epigenetic studies to date have been limited to exploring either associations between maternal exposures and infant epigenetic patterns, or between infant epigenetic patterns and later phenotypes. Evidence linking maternal exposure to infant phenotypes with epigenetics as a mediating mechanism is scarce.

The first category of associations, linking maternal exposures to infant epigenetic patterns, was reviewed earlier. A number of studies linking infant epigenetic patterns with later phenotypes have focused on imprinted genes as candidate loci, due to their role in fetal growth [17] and a range of diseases [3,20]. It has been suggested that imprinted genes may be more vulnerable to the effects of epigenetic dysregulation, since epigenetic mechanisms underpin their monoallelic expression. On chromosome region 11p15.5 there are two imprinting control regions (ICRs) of interest: the H19/IGF2 domain (ICR1) and the KCNQ1/CDKN1C domain (ICR2). Hypomethylation of ICR1 and hypermethylation of ICR2 are associated with Russell-Silver syndrome (RSS; an undergrowth disorder), whereas hypermethylation of ICR1 and hypomethylation of ICR2 are associated with Beckwith-Wiedemann syndrome (BWS; an overgrowth disorder) [3]. In addition to the aberrant methylation found at ICR1 and ICR2, some studies suggest patients with RSS and BWS show methylation defects at multiple gene loci [87]. Evidence for associations between growth-related phenotypes (other than RSS and BWS) and methylation patterns at several imprinted gene loci is slowly accumulating. Other phenotypes investigated to date include intrauterine growth restriction [88], small for gestational age [89], birth weight [90] and later adiposity [91].

Several studies describe associations between maternal one-carbon metabolites and infant growth-related outcomes, for example, vitamin B_{12} [92], folate [93], and homocysteine [94]. What then is the evidence that these associations are mediated through epigenetic mechanisms? Preliminary evidence from the Dutch Hunger Winter, as described earlier, suggests that exposure to famine in pregnancy, particularly around conception, is associated with differential methylation in genes linked to growth, and development [41], and that famine exposure is also related to a wide range of offspring cognitive health and cardiometabolic risk factors six decades later [95]. It is, however, difficult to establish the direction of causality, since disease states can also influence the epigenome [96]. This issue of reverse causality is particularly pertinent to studies using a retrospective cohort design. Stronger evidence comes from prospective cohorts, such as the Newborn Epigenetics Study (NEST) in the United States. Hoyo et al. describe a positive association between maternal folate levels in the first trimester and birth weight [90]. Increased maternal folate was also associated with increased methylation at *MEG3*, *PLAGL1*, and *PEG3* in infant cord blood, and decreased methylation at *IGF2*. Five differentially methylated sites were associated with birth weight, and it was speculated that the association seen between maternal

folate and birth weight could be mediated by differential methylation at *MEG3*, *H19*, and *PLAGL1*. In a more recent study from the same cohort, McCullough et al. found maternal plasma concentrations of homocysteine in the first trimester were inversely associated with birth weight, particularly in males [97]. Children born to mothers with the highest quartile of plasma B_{12} showed lower weight gain between birth and 3 years. However, only maternal vitamin B_6 was positively associated with cord methylation at a DMR from *MEG3*. A further example comes from Godfrey et al., who found that higher methylation of *RXRA* and *eNOS* in umbilical cord tissue was associated with offspring adiposity at age 9, and that higher *RXRA* methylation was also associated with lower maternal carbohydrate intake in early pregnancy [91].

Given the scarcity of existing data and the many potential confounders in cohort studies, there is not yet any clear consensus on the extent to which epigenetics is involved in the etiology of suboptimal growth, particularly in the more widely investigated phenotypes of IUGR and SGA [98]. Similarly, although prepregnancy folic acid supplementation is known to prevent neural tube defects, the extent to which this protection is epigenetically mediated is not yet clear [56]. In the next few years, a number of prospective cohort studies should shed further light on the links between early life exposures and phenotypes, mediated by epigenetic mechanisms. In the meantime, we can use the literature from the range of studies described earlier to speculate that periconceptional nutritional insults may influence epigenetic mechanisms in the infant, causing disruption at multiple loci and potentially leading to a broad spectrum of phenotypic consequences in later life.

NEXT STEPS

The field of epigenetic epidemiology is in its infancy. Future challenges include the need to consider a much larger number of candidate genetic regions (a challenge that will be met as rapid improvements in technology facilitate epigenome-wide association studies with ever increasing genomic resolution); the need to focus on tissue types most appropriate to the phenotype of interest; and the need to integrate data on genetic variation, gene expression, and other epigenetic mechanisms such as histone modifications and miRNAs. Investigations will also need to assess the extent to which postnatal exposures may interact with periconceptional effects. Finally, although we focus on the maternal periconceptional environment in this chapter, there is growing interest in the potential for paternal exposures to affect the infant epigenome transgenerationally via sperm methylation profiles [99]. Despite limited human evidence to date, this is an area of great potential public heath relevance.

Once disrupted epigenetic patterns can be better mapped onto phenotypes, the significance for public health impact becomes more tangible. Can we design and test nutritional interventions to "correct" a suboptimal maternal metabolome, hence reducing patterns of aberrant methylation in the infant epigenome and reducing the burden of ill health? Despite the current complexities of disentangling the web of multiple exposures influencing the infant epigenome, as technology advances, collaborations grow, and knowledge is gained, the field of nutritional epigenetics has the potential to impact the health of children not only within the first 1,000 days but also across multiple generations.

REFERENCES

1. Egger G, Liang G, Aparicio A et al. Epigenetics in human disease and prospects for epigenetic therapy. *Nature* 2004; 429:457–63.
2. Esteller M. Epigenetics in cancer. *N Engl J Med* 2008; 358:1148–59.
3. Piedrahita JA. The role of imprinted genes in fetal growth abnormalities. *Birth Defects Res A Clin Mol Teratol* 2011; 91:682–92.
4. Jaenisch R, Bird A. Epigenetic regulation of gene expression: How the genome integrates intrinsic and environmental signals. *Nat Genet* 2003; 33(Suppl):245–54.
5. Wang Y, Leung FCC. An evaluation of new criteria for CpG islands in the human genome as gene markers. *Bioinformatics* 2004; 20:1170–7.
6. Illingworth RS, Bird AP. CpG islands: "A rough guide." *FEBS Lett* 2009; 583:1713–20.
7. Hill PWS, Amouroux R, Hajkova P. DNA demethylation, Tet proteins and 5-hydroxy-methylcytosine in epigenetic reprogramming: An emerging complex story. *Genomics* 2014; 104:324–33.
8. Messerschmidt DM, Knowles BB, Solter D. DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes Dev* 2014; 28:812–28.
9. Kouzarides T. Chromatin modifications and their function. *Cell* 2007; 128:693–705.
10. Klose RJ, Bird AP. Genomic DNA methylation: The mark and its mediators. *Trends Biochem Sci* 2006; 31:89–97.
11. Richards EJ, Elgin SC. Epigenetic codes for heterochromatin formation and silencing. *Cell* 2002; 108:489–500.
12. Mattick JS, Makunin I V. Non-coding RNA. *Hum Mol Genet* 2006; 15(Spec No 1):R17–29.
13. Guil S, Esteller M. DNA methylomes, histone codes and miRNAs: Tying it all together. *Int J Biochem Cell Biol* 2009; 41:87–95.
14. Sato F, Tsuchiya S, Meltzer SJ et al. MicroRNAs and epigenetics. *FEBS J* 2011; 278:1598–609.
15. Langley-Evans SC. Nutrition in early life and the programming of adult disease: A review. *J Hum Nutr Diet* 2015; 28:1–14.
16. Seisenberger S, Peat JR, Hore TA et al. Reprogramming DNA methylation in the mammalian life cycle: Building and breaking epigenetic barriers. *Philos Trans R Soc Lond B Biol Sci* 2013; 368:20110330.
17. Reik W, Walter J. Genomic imprinting: Parental influence on the genome. *Nat Rev Genet* 2001; 2:21–32.
18. Smallwood SA, Kelsey G. De novo DNA methylation: A germ cell perspective. *Trends Genet* 2012; 28:33–42.
19. Waterland RA, Michels KB. Epigenetic epidemiology of the developmental origins hypothesis. *Annu Rev Nutr* 2007; 27:363–88.
20. Ishida M, Moore GE. The role of imprinted genes in humans. *Mol Aspects Med* 2013; 34:826–40.
21. Moore GE, Ishida M, Demetriou C et al. The role and interaction of imprinted genes in human fetal growth. *Philos Trans R Soc B Biol Sci* 2015; 370:20140074.
22. Rakyan VK, Blewitt ME, Druker R et al. Metastable epialleles in mammals. *Trends Genet* 2002; 18:348–51.
23. Waterland RA, Kellermayer R, Laritsky E et al. Season of conception in rural Gambia affects DNA methylation at putative human metastable epialleles. *PLoS Genet* 2010; 6:e1001252.
24. Dominguez-Salas P, Moore SE, Baker MS et al. Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. *Nat Commun* 2014; 5:3746.

25. Lander ES, Linton LM, Birren B et al. Initial sequencing and analysis of the human genome. *Nature* 2001; 409:860–921.
26. Slotkin RK, Martienssen R. Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* 2007; 8:272–85.
27. Waterland RA, Jirtle RL. Transposable elements: Targets for early nutritional effects on epigenetic gene regulation. *Mol Cell Biol* 2003; 23:5293–300.
28. Barker DJ, Osmond C, Golding J et al. Growth in utero, blood pressure in childhood and adult life, and mortality from cardiovascular disease. *BMJ* 1989; 298:564–7.
29. Barker DJ, Gluckman PD, Godfrey KM et al. Fetal nutrition and cardiovascular disease in adult life. *Lancet (London, England)* 1993; 341:938–41.
30. Curhan GC, Chertow GM, Willett WC et al. Birth weight and adult hypertension and obesity in women. *Circulation* 1996; 94:1310–5.
31. Eriksson JG, Forsen T, Tuomilehto J et al. Catch-up growth in childhood and death from coronary heart disease: Longitudinal study. *BMJ* 1999; 318:427–31.
32. Lumey LH, Stein AD, Kahn HS et al. Cohort profile: The Dutch Hunger Winter families study. *Int J Epidemiol* 2007; 36:1196–1204.
33. Smith CA. The effect of wartime starvation in Holland upon pregnancy and its product. *Am J Obstet Gynecol* 1947; 53:599–608.
34. Roseboom TJ, van der Meulen JH, van Montfrans GA et al. Maternal nutrition during gestation and blood pressure in later life. *J Hypertens* 2001; 19:29–34.
35. Ravelli GP, Stein ZA, Susser MW. Obesity in young men after famine exposure in utero and early infancy. *N Engl J Med* 1976; 295:349–53.
36. Susser E, Neugebauer R, Hoek HW et al. Schizophrenia after prenatal famine. Further evidence. *Arch Gen Psychiatry* 1996; 53:25–31.
37. St Clair D, Xu M, Wang P et al. Rates of adult schizophrenia following prenatal exposure to the Chinese famine of 1959–1961. *JAMA* 2005; 294:557–62.
38. Li Y, He Y, Qi L et al. Exposure to the Chinese famine in early life and the risk of hyperglycemia and type 2 diabetes in adulthood. *Diabetes* 2010; 59:2400–6.
39. Rao S, Yajnik CS, Kanade A et al. Intake of micronutrient-rich foods in rural Indian mothers is associated with the size of their babies at birth: Pune Maternal Nutrition Study. *J Nutr* 2001; 131:1217–24.
40. Yajnik CS, Deshpande SS, Jackson AA et al. Vitamin B_{12} and folate concentrations during pregnancy and insulin resistance in the offspring: The Pune Maternal Nutrition Study. *Diabetologia* 2008; 51:29–38.
41. Heijmans BT, Tobi EW, Stein AD et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc Natl Acad Sci USA* 2008; 105:17046–9.
42. Tobi EW, Slieker RC, Stein AD et al. Early gestation as the critical time-window for changes in the prenatal environment to affect the adult human blood methylome. *Int J Epidemiol* 2015; 44:1211–23.
43. Stöger R. The thrifty epigenotype: An acquired and heritable predisposition for obesity and diabetes? *BioEssays* 2008; 30:156–66.
44. Steegers-Theunissen RPM, Twigt J, Pestinger V et al. The periconceptional period, reproduction and long-term health of offspring: The importance of one-carbon metabolism. *Hum Reprod Update* 2013; 19:640–55.
45. Lu SC, Mato JM. S-adenosylmethionine in liver health, injury, and cancer. *Physiol Rev* 2012; 92:1515–42.
46. Pajares MA, Pérez-Sala D. Betaine homocysteine S-methyltransferase: Just a regulator of homocysteine metabolism? *Cell Mol Life Sci* 2006; 63:2792–803.
47. Scotti M, Stella L, Shearer EJ et al. Modeling cellular compartmentation in one-carbon metabolism. *Wiley Interdiscip Rev Syst Biol Med* 2013; 5:343–65.

48. Mason JB. Biomarkers of nutrient exposure and status in one-carbon (methyl) metabolism. *J Nutr* 2003; 133:941S–947.

49 Stover PJ, Field MS. Trafficking of intracellular folates. *Adv Nutr* 2011; 2:325–31.

50. Fredriksen A, Meyer K, Ueland PM et al. Large-scale population-based metabolic phenotyping of thirteen genetic polymorphisms related to one-carbon metabolism. *Hum Mutat* 2007; 28:856–65.

51. Waterland RA, Dolinoy DC, Lin J-R et al. Maternal methyl supplements increase off-spring DNA methylation at Axin Fused. *Genesis* 2006; 44:401–6.

52. Zhou S, Zhang Z, Xu G. Notable epigenetic role of hyperhomocysteinemia in atherogenesis. *Lipids Health Dis* 2014; 13:134.

53. Beard RS, Bearden SE. Vascular complications of cystathionine β-synthase deficiency: Future directions for homocysteine-to-hydrogen sulfide research. *Am J Physiol Heart Circ Physiol* 2011; 300:H13–26.

54. Huang JY, Butler LM, Wang R et al. Dietary intake of one-carbon metabolism-related nutrients and pancreatic cancer risk: The Singapore Chinese Health Study. *Cancer Epidemiol Biomarkers Prev* 2016; 25:417–24.

55. Crider KS, Yang TP, Berry RJ et al. Folate and DNA methylation: A review of molecular mechanisms and the evidence for folate's role. *Adv Nutr* 2012; 3:21–38.

56. Gonseth S, Roy R, Houseman EA et al. Periconceptional folate consumption is associated with neonatal DNA methylation modifications in neural crest regulatory and cancer development genes. *Epigenetics* 2015; 10:1166–76.

57. Ba Y, Yu H, Liu F et al. Relationship of folate, vitamin B_{12} and methylation of insulin-like growth factor-II in maternal and cord blood. *Eur J Clin Nutr* 2011; 65:480–5.

58. Amarasekera M, Martino D, Ashley S et al. Genome-wide DNA methylation profiling identifies a folate-sensitive region of differential methylation upstream of ZFP57-imprinting regulator in humans. *FASEB J* 2014; 28:4068–76.

59. Woodfine K, Huddleston JE, Murrell A. Quantitative analysis of DNA methylation at all human imprinted regions reveals preservation of epigenetic stability in adult somatic tissue. *Epigenetics Chromatin* 2011; 4:1.

60. McKay JA, Groom A, Potter C et al. Genetic and non-genetic influences during pregnancy on infant global and site specific DNA methylation: Role for folate gene variants and vitamin B_{12}. *PLoS One* 2012; 7:e33290.

61. Jiang X, Yan J, West AA et al. Maternal choline intake alters the epigenetic state of fetal cortisol-regulating genes in humans. *FASEB J* 2012; 26:3563–74.

62. Moore SE, Cole TJ, Poskitt EME et al. Season of birth predicts mortality in rural Gambia. *Nature* 1997; 388:434.

63. Silver MJ, Kessler NJ, Hennig BJ et al. Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. *Genome Biol* 2015; 16:118.

64. Huang T, Zheng J, Chen Y et al. High consumption of Ω-3 polyunsaturated fatty acids decrease plasma homocysteine: A meta-analysis of randomized, placebo-controlled trials. *Nutrition* 2011; 27:863–7.

65. Kabaran S, Besler HT. Do fatty acids affect fetal programming? *J Heal Popul Nutr* 2015; 33:14.

66. Roy S, Kale A, Dangat K et al. Maternal micronutrients (folic acid and vitamin B(12)) and omega 3 fatty acids: Implications for neurodevelopmental risk in the rat offspring. *Brain Dev* 2012; 34:64–71.

67. Khot V, Kale A, Joshi A et al. Expression of genes encoding enzymes involved in the one carbon cycle in rat placenta is determined by maternal micronutrients (folic acid, vitamin B_{12}) and omega-3 fatty acids. *Biomed Res Int* 2014; 2014:613078.

68. Huang T, Hu X, Khan N et al. Effect of polyunsaturated fatty acids on homocysteine metabolism through regulating the gene expressions involved in methionine metabolism. *Sci World* J 2013; 2013:1–8.

69. Khot V, Chavan-Gautam P, Joshi S. Proposing interactions between maternal phospholipids and the one carbon cycle: A novel mechanism influencing the risk for cardiovascular diseases in the offspring in later life. *Life Sci* 2015; 129:16–21.

70. Lee H-S, Barraza-Villarreal A, Biessy C et al. Dietary supplementation with polyunsaturated fatty acid during pregnancy modulates DNA methylation at IGF2/H19 imprinted genes and growth of infants. *Physiol Genomics* 2014; 46:851–7.

71. Lee H-S, Barraza-Villarreal A, Hernandez-Vargas H et al. Modulation of DNA methylation states and infant immune system by dietary supplementation with ω-3 PUFA during pregnancy in an intervention study. *Am J Clin Nutr* 2013; 98:480–7.

72. Sharp GC, Lawlor DA, Richmond RC et al. Maternal pre-pregnancy BMI and gestational weight gain, offspring DNA methylation and later offspring adiposity: Findings from the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* 2015; 44:1288–304.

73. Wells JCK, Fewtrell MS. Measuring body composition. *Arch Dis Child* 2006; 91:612–7.

74. Müller MJ, Lagerpusch M, Enderle J et al. Beyond the body mass index: Tracking body composition in the pathogenesis of obesity and the metabolic syndrome. *Obes Rev* 2012; 13(Suppl 2):6–13.

75. Elshorbagy AK, Nijpels G, Valdivia-Garcia M et al. S-adenosylmethionine is associated with fat mass and truncal adiposity in older adults. *J Nutr* 2013; 143:1982–8.

76. Soubry A, Murphy SK, Wang F et al. Newborns of obese parents have altered DNA methylation patterns at imprinted genes. *Int J Obes* 2015; 39:650–7.

77. Young JI, Züchner S, Wang G. Regulation of the epigenome by vitamin C. *Annu Rev Nutr* 2015; 35:545–64.

78. Blaschke K, Ebata KT, Karimi MM et al. Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature* 2013; 500:222–6.

79. Babenko O, Kovalchuk I, Metz GAS. Stress-induced perinatal and transgenerational epigenetic programming of brain development and mental health. *Neurosci Biobehav Rev* 2014; 48:70–91.

80. Anway MD, Skinner MK. Epigenetic transgenerational actions of endocrine disruptors. *Endocrinology* 2006; 147:S43–49.

81. El Hajj N, Schneider E, Lehnen H et al. Epigenetics and life-long consequences of an adverse nutritional and diabetic intrauterine environment. *Reproduction* 2014; 148:R111–20.

82. Davie JR. Inhibition of histone deacetylase activity by butyrate. *J Nutr* 2003; 133:2485S–93S.

83. Fang M, Chen D, Yang CS. Dietary polyphenols may affect DNA methylation. *J Nutr* 2007; 137:223S–8S.

84. Pereira F, Barbáchano A, Singh PK et al. Vitamin D has wide regulatory effects on histone demethylase genes. *Cell Cycle* 2012; 11:1081–9.

85. Feng Y, Zhao L-Z, Hong L et al. Alteration in methylation pattern of GATA-4 promoter region in vitamin A-deficient offspring's heart. *J Nutr Biochem* 2013; 24:1373–80.

86. Claycombe KJ, Brissette CA, Ghribi O. Epigenetics of inflammation, maternal infection, and nutrition. *J Nutr* 2015; 145:1109S–15S.

87. Azzi S, Rossignol S, Steunou V et al. Multilocus methylation analysis in a large cohort of 11p15-related foetal growth disorders (Russell Silver and Beckwith Wiedemann syndromes) reveals simultaneous loss of methylation at paternal and maternal imprinted loci. *Hum Mol Genet* 2009; 18:4724–33.

88. Einstein F, Thompson RF, Bhagat TD et al. Cytosine methylation dysregulation in neonates following intrauterine growth restriction. *PLoS One* 2010; 5:e8887.

89. Bouwland-Both MI, van Mil NH, Stolk L et al. DNA methylation of IGF2DMR and H19 is associated with fetal and infant growth: The generation R study. *PLoS One* 2013; 8:e81731.

90. Hoyo C, Daltveit AK, Iversen E et al. Erythrocyte folate concentrations, CpG methylation at genomically imprinted domains, and birth weight in a multiethnic newborn cohort. *Epigenetics* 2014; 9:1120–30.

91. Godfrey KM, Sheppard A, Gluckman PD et al. Epigenetic gene promoter methylation at birth is associated with child's later adiposity. *Diabetes* 2011; 60:1528–34.

92. Rush EC, Katre P, Yajnik CS. Vitamin B_{12}: One carbon metabolism, fetal growth and programming for chronic disease. *Eur J Clin Nutr* 2014; 68:2–7.

93. van Uitert EM, Steegers-Theunissen RPM. Influence of maternal folate status on human fetal growth parameters. *Mol Nutr Food Res* 2013; 57:582–95.

94. Hogeveen M, Blom HJ, den Heijer M. Maternal homocysteine and small-for-gestational-age offspring: Systematic review and meta-analysis. *Am J Clin Nutr* 2012; 95:130–6.

95. de Rooij SR, Painter RC, Roseboom TJ et al. Glucose tolerance at age 58 and the decline of glucose tolerance in comparison with age 50 in people prenatally exposed to the Dutch famine. *Diabetologia* 2006; 49:637–43.

96. Relton CL, Davey Smith G. Two-step epigenetic Mendelian randomization: A strategy for establishing the causal role of epigenetic processes in pathways to disease. *Int J Epidemiol* 2012; 41:161–76.

97. McCullough LE, Miller EE, Mendez MA et al. Maternal B vitamins: Effects on offspring weight and DNA methylation at genomically imprinted domains. *Clin Epigenetics* 2016; 8:8.

98. Toure DM, Baccaglini L, Opoku ST et al. Epigenetic dysregulation of Insulin-like growth factor (IGF)-related genes and adverse pregnancy outcomes: A systematic review. *J Matern Fetal Neonatal Med* 2016; 18:1–11.

99. Soubry A. Epigenetic inheritance and evolution: A paternal perspective on dietary influences. *Prog Biophys Mol Biol* 2015; 118:79–85.

ANNEX 3.6   Magazine article: Epigenetics, Nutrition and Human Health

RESEARCH PAPER COVER SHEET

SECTION A – Student Details

Student	Philip James
Principal Supervisor	Dr Matt Silver
Thesis Title	Maternal nutrition, seasonality and epigenetics: an exploration of one-carbon metabolism and a novel nutritional supplement design in The Gambia.

SECTION B – Paper already published

Where was the work published?	Sight and Life Magazine		
When was the work published?	10th July, 2015		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion	N/A		
Have you retained the copyright for the work?*	No	Was the work subject to academic peer review?	No, editorial only

 If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

From: Klaus Kraemer <klaus.kraemer@sightandlife.org>
Sent: 25 October 2018 03:28
To: Philip James <Philip.James@lshtm.ac.uk>
Subject: RE: Permission to re-use a Sight and Life article in a thesis

Dear Philip,
Greetings from Mumbai where we will have another EPC today on aflatoxin.
We will certainly grant permission to include your contribution to the magazine in your thesis. Good luck with everything.
Best wishes,
Klaus

Dr Klaus Kraemer
Managing Director, *Sight and Life*
Adjunct Professor, *Johns Hopkins Bloomberg School of Public Health*
PO Box 2116, 4002 Basel, Switzerland
phone +41 61 815 8756 · mobile +41 79 820 0428
email klaus.kraemer@sightandlife.org

sightandlife.org · Facebook · Twitter · LinkedIn · Instagram · YouTube
Read our latest issue of the *Sight and Life* magazine!

Multi-authored Work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper.

Andrew Prentice and I were originally approached by Sight and Life to write a short magazine article on epigenetics and nutrition. I drafted the chapter, and incorporated comments from Matt Silver and Andrew Prentice for the final version.

Student Signature:

Date: 16th October 2018

Supervisor Signature:

Date: 16th October 2018

Epigenetics, Nutrition and Human Health

**Philip James, Matt Silver,
and Andrew Prentice**
MRC International Nutrition Group, The Gambia

Imagine yourself at the scene of a crime where you need to determine the age of a victim or perpetrator. If you are lucky, you will have access to skin or dental tissue, or perhaps anthropometry measures, all of which may help determine an approximate age. However, recent research suggests that you could instead obtain an astonishingly accurate measure of chronological age with an "epigenetic clock" that uses a very small number of epigenetic marks in the genome.[1] This would allow you to pin down a sample's age to within a few months, irrespective of the tissue from which it was obtained.

The field of epigenetics is currently attracting a lot of attention from scientists and the wider public. Epigenetic processes describe changes to the genome that can alter gene expression without changing the underlying DNA sequence[2] (**Figure 1**). One such mechanism is DNA methylation of cytosine bases at CpG dinucleotide sites, and there is strong evidence that this can be influenced by a diverse array of intrinsic and environmental factors, including age, disease, stress, exposure to pollutants, and nutrition. Furthermore, epigenetic marks have been associated with a range of diseases affecting health throughout the life course, including cancers, and neurological and metabolic disorders.[3] Together, these observations suggest that our epigenomes carry a "cellular memory" of environmental insults, with the potential for lasting effects on health and disease. Epigenetic changes at certain locations are also believed to be heritable, raising the possibility of trans-generational effects that cannot be explained by standard Mendelian genetics.[4]

Diet and epigenetics in The Gambia

Our group is exploring human diet-epigenome interactions by exploiting an "experiment of nature" in rural Gambia whereby

FIGURE 1: Epigenetic mechanisms of DNA modification. Reproduced with permission from Yan MS-C, Matouk CC, Marsden PA.[5]

1 **DNA methylation**
Cytosine DNA methylation

2 **Histone code**

H2B
H2A   H4
H3
H1

RNA strand

3 **RNA-based mechanisms**

Nucleosome
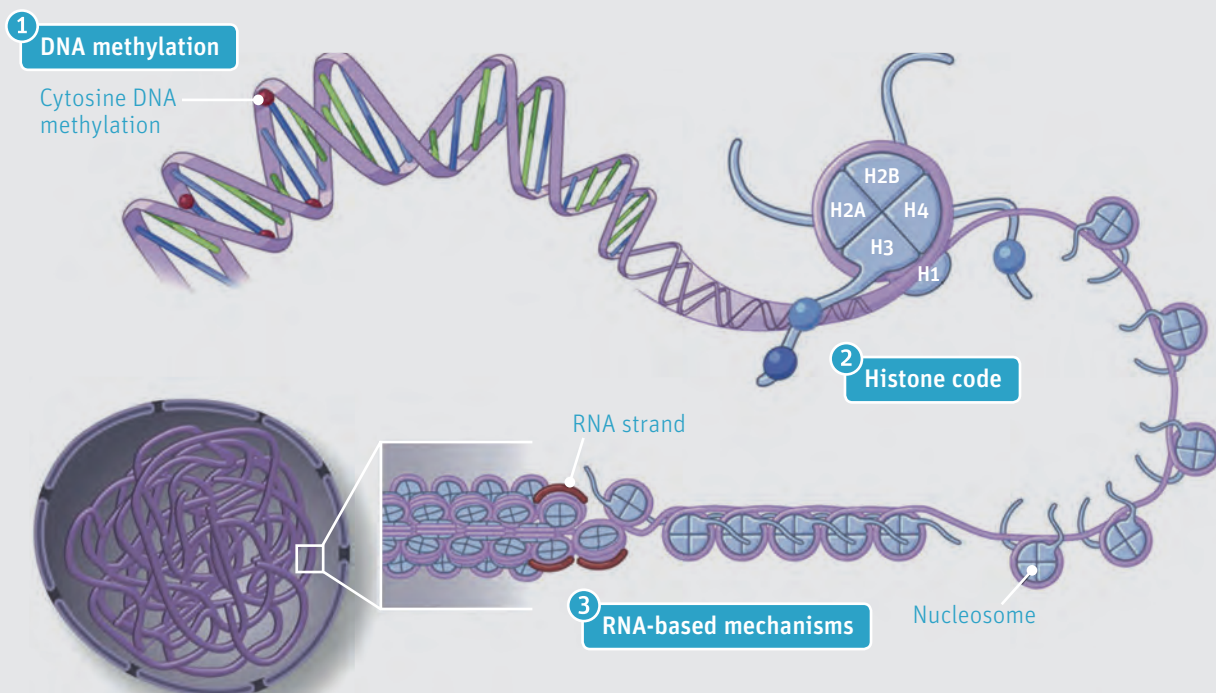
FIGURE 2: Keneba in the rainy and dry seasons

fluctuations in energy balance and maternal nutrition show a distinct bimodal seasonal pattern (**Figure 2**). Our study population experiences a rainy ("hungry") season from July to September, with increased energy expenditure through agricultural work, depleted food stores, and peaks of malarial and diarrheal diseases. The dry ("harvest") season occurs from February to April, when harvesting takes place, leading to improved food security.

"Epigenomes carry a 'cellular memory' of environmental insults, with the potential for lasting effects on health and disease"

Almost 20 years ago, we uncovered strong evidence that the season when a child is born has a profound effect on lifelong health. Gambian children born during the rainy season are up to 10 times more likely to die prematurely in young adulthood.[6] Since then, pieces of the puzzle are starting to fall into place, with nutrition-related epigenetic regulation in the early embryo emerging as a highly plausible candidate mechanism.

Five years ago, in partnership with Rob Waterland at Baylor College of Medicine in Houston, we demonstrated that season of conception predicts DNA methylation at certain genomic loci known as metastable epialleles (MEs). These are CpG sites whose methylation state varies between individuals, but where variation is correlated across tissues originating from all germ layers in a single individual[7] – indicating that the marks must have been laid down in the first few hours after conception before cell types start to specialize. This period in the very early embryo is when the methylome is globally reprogrammed – a period crucial to development.[8] We therefore use MEs as a device

to study the influence of a mother's nutrition on the epigenome of the baby at the time of conception.

"We use metastable epialleles as a device to study the influence of a mother's nutrition on the epigenome of the baby at the time of conception"

We have since shown that a mother's levels of several key nutrients vary by season and predict DNA methylation at six MEs in their offspring.[9] These nutrients play a role in 1-carbon metabolism, a biological pathway crucial for the provision of methyl (CH3) groups required for DNA methylation (**Figure 3**). The two main carriers that activate, transport and transfer these methyl groups are tetrahydrofolate (THF) and S-adenosylmethionine (SAM). While 1-carbon units are used as substrates for a whole range of intricate biochemical processes (including cellular biosynthesis, redox status regulation and genome maintenance through the regulation of nucleotide pools), it is their role in cytosine and histone methylation that is central to the interplay between diet and the epigenome.

Linking diet, epigenetics and health

Our latest research has identified another ME that is sensitive to the periconceptional environment in Gambian infants.[11] The associated gene has been implicated in the regulation of immune function and is a putative tumor suppressor, suggesting a potential epigenetic pathway linking a nutritional insult affecting the very early embryo to some serious outcomes in later life. This requires rigorous testing, but work in animal models has already demon-

FIGURE 3: An overview of 1-carbon metabolism. Reproduced with permission from Dominguez-Salas P, Moore SE, Cole D et al.[10]
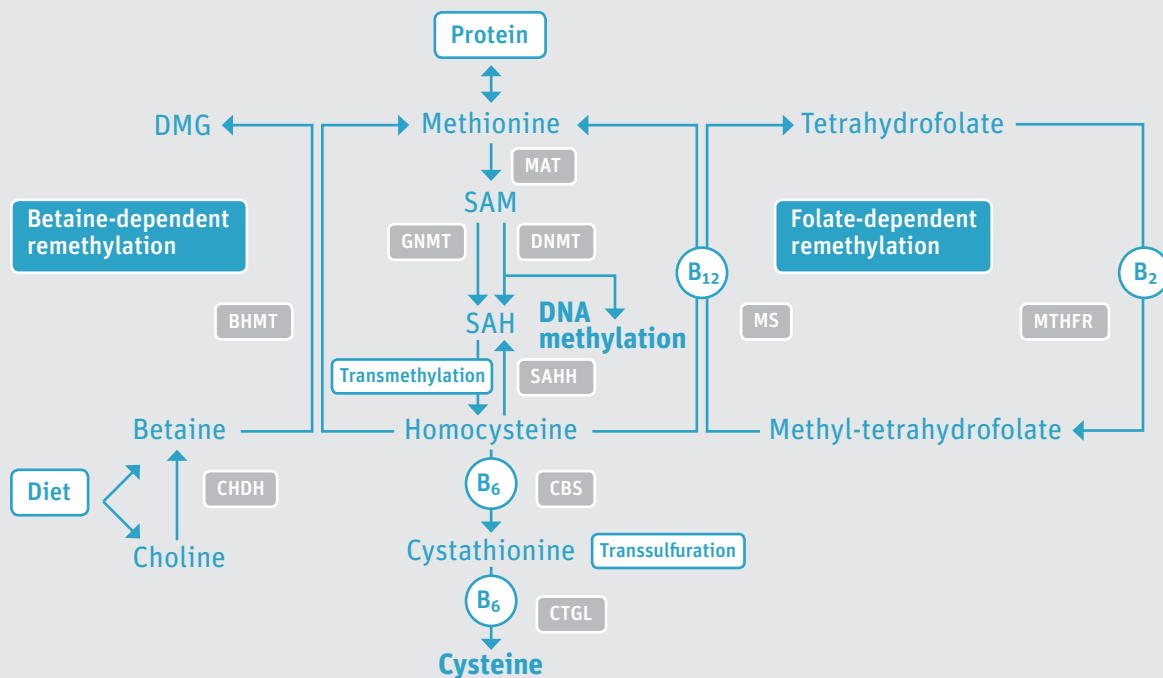
BHMT, betaine-homocysteine methyltransferase; B_2, vitamin B_2; B_6, vitamin B_6; B_{12}, vitamin B_{12}; CBS, cystathionine-β-synthase; CHDH, choline dehydrogenase; CTGL, cystathionine-γ-lyase; DMG, dimethylglycine; DNMT, DNA methyltransferases; GNMT, glycine N-methyltransferase; MAT, methionine adenosyltransferase; MTHFR, methylenetetrahydrofolate reductase; MS, methionine synthase; SAH, S-adenosylhomocysteine; SAHH, S-adenosylhomocysteine hydrolase; SAM, S-adenosylmethionine

strated that maternal diet can influence the offspring epigenome, with subsequent dramatic effects on phenotype. In the case of the Agouti mouse, pregnant dams fed a diet rich in methyl donor micronutrients (vitamin B_{12}, folic acid, betaine and choline) produced offspring with increased methylation at the agouti locus, leading to fewer obese yellow offspring and more lean brown offspring (**Figure 4**) – characteristics that persisted into adult life, with associated differences in appetite, adiposity and glucose tolerance.[12,13]

Implications for the future
Our Gambian studies offer the first-in-human evidence that periconceptional nutrition can affect the epigenome of the fetus. Just as the epigenetic clock shows that our genome carries an epigenetic signature of the ageing process, it seems that it also bears the hallmark of nutritional exposures at the very start of life.

"Our Gambian studies offer the first-in-human evidence that periconceptional nutrition can affect the epigenome of the fetus"

The next task is to characterize more clearly how this affects phenotype. To what extent does disrupted methylation affect gene expression? How might these effects influence lifelong risk of morbidity and mortality? Most importantly, from a translational perspective, can epigenetic errors be corrected by optimizing the maternal metabolome through periconceptional nutritional supplementation?

The best known periconceptional nutrition supplement to prevent neural tube defects is folic acid. While the mechanism in this particular example has not yet been conclusively described, epigenetics is a strong contender. Given current research, we believe it is possible to speculate that other forms of periconceptional supplementation might positively influence the epigenome of the unborn child, leading to lifelong health gains across a whole variety of phenotypes.

Clearly the process of validating key findings, further delineating the mechanisms involved, and translating the science into practical public health policy requires a large collaborative effort, along with advances in technology and access to funding. While it is not possible to predict the precise direction research will take in this rapidly evolving field, it seems likely that nutritional considerations will remain paramount in the development of epigenetic therapies to improve health over the life course.

FIGURE 4: The effect of maternal diet on offspring phenotype of the Agouti mouse. Coat color corresponds to methylation levels at the agouti locus. Reproduced with permission from the American Society for Microbiology, Waterland et al.[12]

Correspondence: *Philip James,*

MRC International Nutrition Group, London School of Hygiene & Tropical Medicine, Keppel Street, London, WC1E 7HT, UK.

Email: philip.james@lshtm.ac.uk

References

01. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol 2013;14(115).

02. Tammen SA, Friso S, Choi S-W. Epigenetics: the link between nature and nurture. Mol Aspects Med 2013;34(4):753–64.

03. Desai M, Jellyman JK, Ross MG. Epigenomics, gestational programming and risk of metabolic syndrome. Int J Obes 2015;(Feb):1–9.

04. Daxinger L, Whitelaw E. Understanding transgenerational epigenetic inheritance via the gametes in mammals. Nat Rev Genet 2012;13(3):153–162.

05. Yan MS-C, Matouk CC, Marsden PA. Epigenetics of the vascular endothelium. J Appl Physiol 2010;109(3):916–26.

06. Moore SE, Cole TJ, Poskitt EME et al. Season of birth predicts mortality in rural Gambia. Nature 1997;388(6641):434.

07. Rakyan VK, Blewitt ME, Druker R et al. Metastable epialleles in mammals. Trends Genet 2002;18(7):348–51.

08. Lee HJ, Hore TA, Reik W. Reprogramming the Methylome: Erasing Memory and Creating Diversity. Cell Stem Cell 2014;14(6):710–719.

09. Dominguez-Salas P, Moore SE, Baker MS et al. Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. Nat Commun 2014;5:1–7.

10. Dominguez-Salas P, Moore SE, Cole D et al. DNA methylation potential: dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women. Am J Clin Nutr 2013;97(6):1217–27.

11. Silver M, Kessler N, Hennig B, et al. Independent genomewide screens identify the tumor suppressor VTRNA2-1 as a human epiallele responsive to periconceptional environment. Genome Biol. 2015. 16:118

12. Waterland RA, Jirtle RL. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. Mol Cell Biol 2003;23(15):5293–300.

13. Waterland RA, Michels KB. Epigenetic epidemiology of the developmental origins hypothesis. Annu Rev Nutr 2007;27:363–88.

ANNEX 4.1   Original methods overview and analysis plan of modelling Gambian maternal nutritional plasma concentration data using a constraint-based platform (RECON-III)

The below provides information of the planned collaboration with Luxembourg Centre for Systems Biomedicine (discontinued as the model one-carbon pathways required some more development). Note that the model referred to in the text below, RECON-III, has since been made public and is termed Recon3D. Full details can be found in Brunk *et al.* (2018)[1].

1.   Introduction to constraint-based modelling

1.1   Overview

Systems biology is an integrative discipline that analyses complex datasets to help generate hypotheses which can be experimentally validated and used to improve computer modelling in an iterative fashion. It uses *in silico* analysis of reconstructed networks and 'omics' datasets. We will use Recon-III (Figure 1), the most comprehensive genome-wide reconstruction of human metabolism to date, representing over 80,000 reactions in more than 20 organs / cellular compartments.
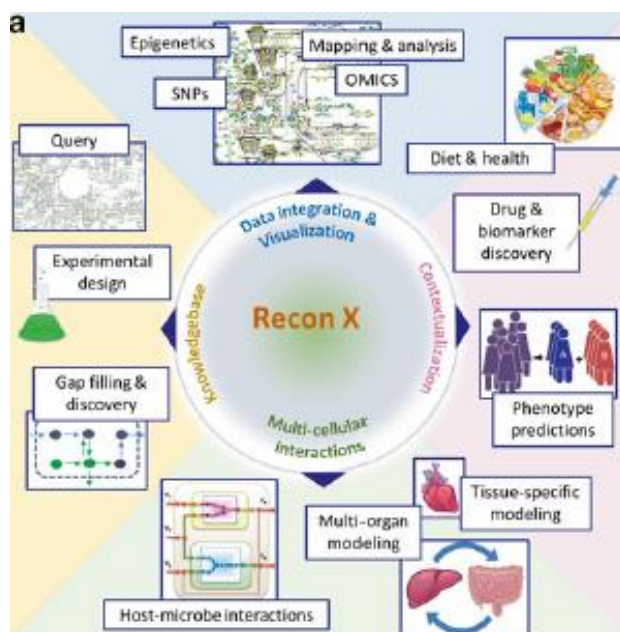
Figure 1: The Recon generation of metabolic modelling. **Source:** Aurich MK, Thiele I. *Methods Mol Biol.* 2016; **1386:** 253–81[2]. Reproduced with permission.

The planned modelling approach uses flux balance analysis (FBA), a method within constraint-based modelling. Fundamental principles of this approach are outlined below. The planned analysis will generate a nuanced understanding of how the overall flow of

metabolites through 1C pathways varies across the year and help us predict how a potential nutritional supplement may impact the metabolome.

1.2 Constraint-based modelling

Constraint-based modelling approach uses stoichiometry and resulting mass balance equations to generate a map of the flux of metabolites through a network of reactions.

Key concepts:

- The law of mass conservation states that in an isolated system mass cannot be created or destroyed by reactions.
- Stoichiometric coefficients in an equation represent the proportion of substrates and products involved in that reaction. They constrain mass balance to ensure the total amount of any compound being produced equals the total amount being consumed in a steady state.
- 'Flux' can be defined as the rate that metabolites are consumed and produced in a metabolic pathway, regulated by enzymes. Each reaction is linked to a gene product via 'gene-reaction-protein associations', with Boolean rules determining whether multiple enzymes are required for a reaction ('AND') or whether the reaction can be catalysed by isozymes ('OR').
- Since metabolic reactions are so fast in comparison to the timescale of cell growth or gene regulation a steady state assumption is adopted.

Because FBA models metabolism in a steady state it is different from models requiring enzymatic kinetic data, which are difficult to obtain *in vivo* at the population level. The models using kinetic data involve series of differential equations which would be time-consuming (if not impossible) to solve in the context of genome-wide metabolic reconstructions.

Creating a **stoichiometric matrix** is a central step in flux balance analysis, notated as Matrix **S** (*m* metabolites x *n* reactions). Figure 2 below illustrates how a stoichiometric matrix is built up from information on metabolic reactions. Part a) represents the first few steps of glycolysis. Part b) shows how this can be portrayed in stoichiometric matrix form (matrix S). Each metabolite is given a row, and each reaction a column. If a metabolite participates in a reaction its stoichiometric coefficients are entered: negative coefficients if consumed and positive coefficients if produced. If one metabolite is present in different cellular compartments then each compartment is also given a separate row. Each reaction is given an upper and lower bound to limit flux rates and determine if the reaction is reversible or not (part c). As many reactions as possible are bounded to experimental data, to constrain the model to the most realistic context.

The stoichiometric matrix can then be used to form flux balance equations using two vectors: Vector **V** detailing the flux of each metabolite through all its reactions (length ***n)***, and vector **X** with the concentration of metabolites (length ***m***). The equations are arranged so that the

overall change in concentration of a metabolite per unit time is determined by calculating its net production or consumption (from matrix S) and associated flux (rate from vector V) for each reaction it participates in. Matrix algebra can be used to show that the change in concentration of metabolites over time (vector X) can be calculated by the dot product of matrix S with vector V (Figure 3).

$$\frac{d[\mathbf{X}]}{dt} = \mathbf{S} \bullet \mathbf{V}$$

In a steady state d$\mathbf{X}$/dt=0, so in a steady state model $\mathbf{S} \bullet \mathbf{V} = \mathbf{0}$
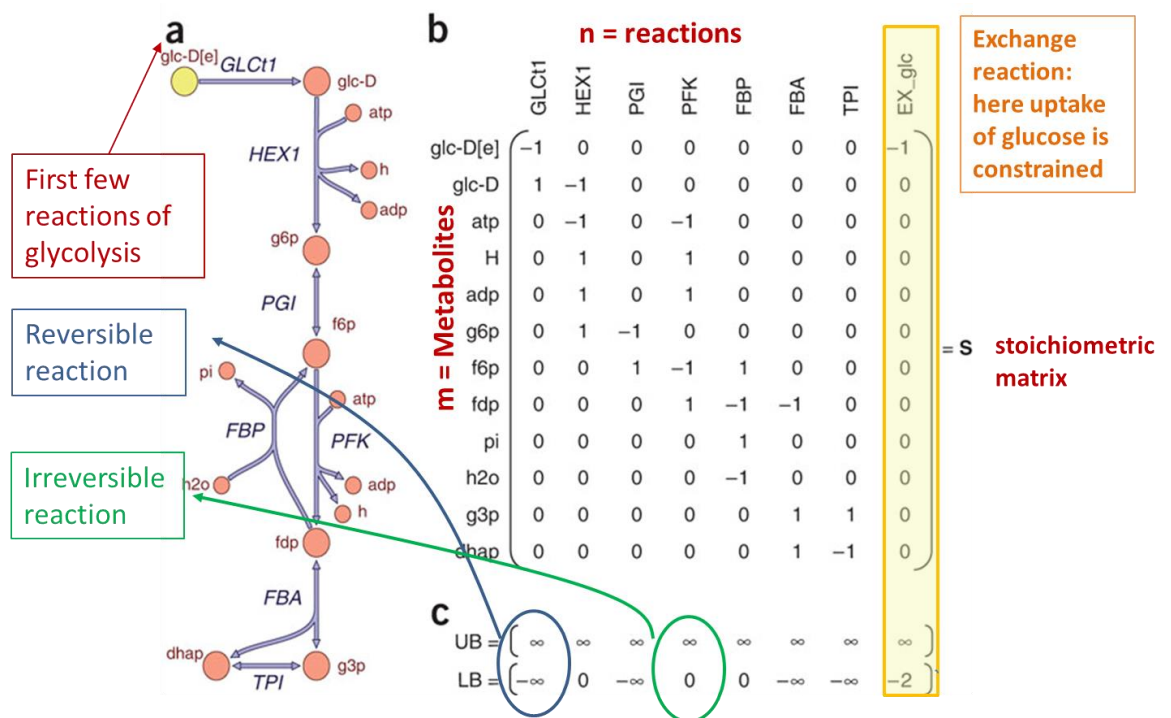
Figure 2: Creating a stoichiometric matrix. Adapted from Becker SA et al. *Nat Protoc.* 2007;**2:** 727–38[3]. Reproduced with permission.

Vector **X** = change in concentration of metabolites

Stoichiometric Matrix **S**

Vector **V** = fluxes through each reaction

		GLCt1	HEX1	PGI	PFK	FBP	FBA	TPI	EX_glc
d[glc-D(e)]/dt	glc-D[e]	−1	0	0	0	0	0	0	−1
d[glc-D]/dt	glc-D	1	−1	0	0	0	0	0	0
d[atp]/dt	atp	0	−1	0	−1	0	0	0	0
d[H]/dt	H	0	1	0	1	0	0	0	0
d[adp]/dt	adp	0	1	0	1	0	0	0	0
d[g6p]/dt	g6p	0	1	−1	0	0	0	0	0
d[f6p]/dt	f6p	0	0	1	−1	1	0	0	0
d[fdp]/dt	fdp	0	0	0	1	−1	−1	0	0
d[pi]/dt	pi	0	0	0	0	1	0	0	0
d[h2o]/dt	h2o	0	0	0	0	−1	0	0	0
d[g3p]/dt	g3p	0	0	0	0	0	1	1	0
d[dhap]/dt	dhap	0	0	0	0	0	1	−1	0

V *GLCt1*
V *HEX1*
V *PGI*
V4
V5
V6
V7
V8

Figure 3: Change in metabolite concentration as the dot product of matrix S with vector V. Adapted from Becker SA et al. *Nat Protoc.* 2007;**2:** 727–38[3]. Reproduced with permission.

We can therefore rearrange our mass balance equations so that **S•V =0** and form a **system of linear equations** (Figure 4). These linear equations impose constraints on the flow of metabolites through the network. Note that in some models the equations also subtract a term describing the dilution of the metabolite due to biomass growth, but this is often ignored as it is usually very small.

For metabolite **glc-D[e]** the overall mass balance (flux) equation is:

$-V_{GLCt1} - V_{EX\_glc} = 0$

For metabolite **fdp** the overall mass balance (flux) equation is:

$V_{PFK} - V_{FBP} - V_{FBA} = 0$

The overall mass balance equations show total rate of consumption = total rate of production for each metabolite
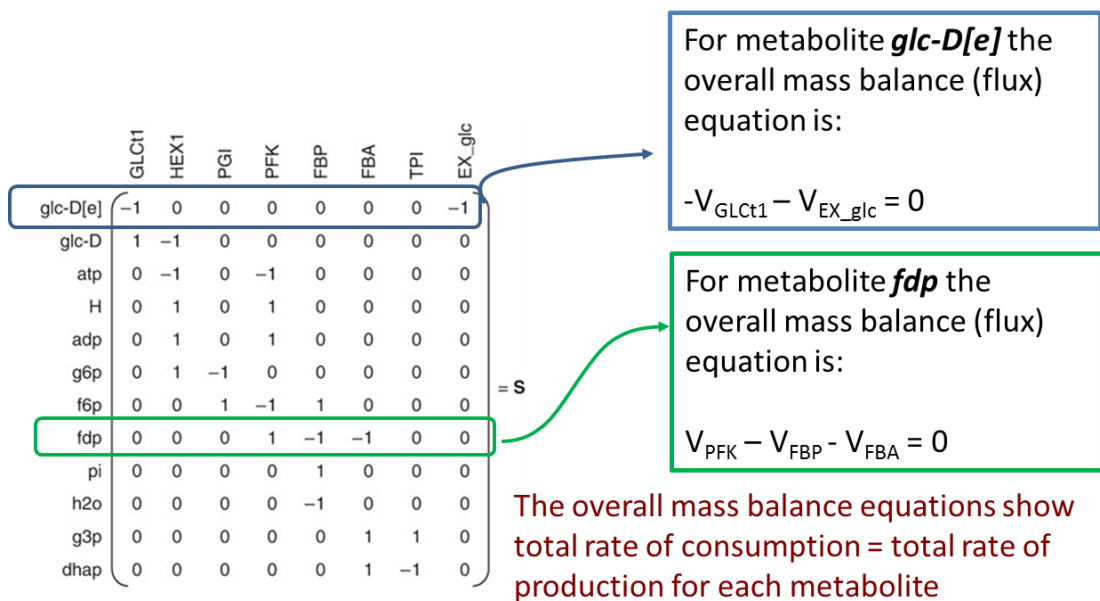
Figure 4: The creation of mass balance equations in a steady state
Adapted from Becker SA et al. *Nat Protoc.* 2007;**2:** 727–38[3]. Reproduced with permission.

Under a given set of constraints (stoichiometry and bounds) we can define the **allowable flux distribution** of the system. We then define our **objective function** (our 'problem' to investigate) *e.g.* maximise production of a certain metabolite. A collection of functions used within Matlab helps us solve the equations to predict our objective function. This produces a vector of weights describing how much each reaction contributes towards the objective. In complex metabolic models our system is under-determined, meaning we have more reactions than compounds (n>m), so there is (usually) never one unique solution. We therefore use linear programming to then identify the flux distribution that maximises or minimises the objective function within the allowable flux distribution space (Figure 5).

Once we have solved the equations to achieve our objective function we can explore the solution space in more detail:

- **Flux variability analysis** allows us to ascertain the minimum and maximum flux through each reaction that still helps us achieve our objective function under the given constraints.
- **Sampling analysis:** this captures the solution space in higher resolution. Within the solution space generated under specific constraints random samples of flux distributions are taken, generating a histogram of flux distributions. The extremes of these histograms are what flux variability analysis (above) calculates.

We define the **allowable flux distribution** of the system using constraints (from our lab data) and our linear equations (from stoichiometry)

We define our **objective function** (our 'problem' to investigate) *e.g.* maximise flux of SAM-> SAH

We solve the equations to predict our objective function (produces a vector of weights describing how much each reaction contributes towards the objective).

In complex metabolic models there is (usually) never one unique solution. Linear programming identifies the flux distribution that maximises or minimises the objective function.

Figure 5: Using linear programming to predict equations meeting the objective function. Adapted from Orth JD, et al. *Nat Biotechnol.* 2010;**28:** 245–8[4]. Reproduced with permission.

2. Preparatory work

2.1 Dietary input file

RECON-III models nutrients available for absorption in the intestine. The nutrients are not forced to be taken up, but the model can decide what is required to achieve a specific objective function. The dietary input is a vector with a set of exchange reactions. Exchange reactions describe how a metabolite enters or exits a cell / compartment. In this case the exchange reactions allow the metabolites to move from the intestine model compartment to the portal vein compartment if required. We have already created a 'Gambia diet' based on foods known to be available from Gambian food tables. We restricted the calorie intake to 1830 kcal / day and corrected the intakes to represent the realistic macronutrients ratios found in the region (high carbohydrate intakes). We set intakes of 1-carbon related dietary inputs to those recorded in recent food intake calculations from Dominguez-Salas *et al.* (2013)[5].

We decided to use a single diet based on mean annual dietary intakes rather than vary the diet by season. This is because for now we are more confident in the plasma concentration data as the more direct measures of nutritional status. We know that dietary intakes do not map onto plasma concentrations in simple ways, there are large measurement errors, and we also know the model has to simplify reality and does not model bioavailability, absorption rates, nutrient-nutrient interactions etc. Therefore at the outset we would prefer to see changes in flux by season / month based purely on plasma concentration data, rather than a combination of plasma with dietary input changes. We could manipulate the diet by season in later analysis to see the impact, but as a secondary objective.

Ideally we need to add three additional metabolites as exchange reactions for the dietary exchange vector.

- Methylcobalamin (VMH ID: C06453)
- Betaine (VMH ID: glyb)
- Choline (VMH ID: chol)

All other metabolites are captured in the Virtual Metabolic Human (VMH) database[1]. Although these three metabolites are currently missing in the VMH dietary flux calculation, we can manually calculate them since we have data on intakes per day. We can convert these data into mmol/(human*day) using molecular mass information.

The flux for methylcobalamin can be calculated by taking the overall vitamin B12 intake and splitting it between methylcobalamin and adenosylcobalamin. So far VMH only calculates exchange reactions for adenosylcobalamin (adocbl).

---

[1] VMH is the Luxembourg university metabolic database which RECON-III draws from.

2.2   Plasma concentration file

As the default the plasma nutrient metabolites represented in RECON-III will be constrained to average European concentrations. However, we will over-ride concentrations for the metabolites we have plasma concentration data on, inputting that laboratory plasma concentration data in micromole/L. All plasma concentration data will be forced into the model as constraints on uptake from the blood compartment. B12 and folate, however, involve an additional manipulation step:

- Most B12 circulating in plasma is methylcobalamin[6], so we will map this onto C06453 (not onto adenosylcobalamin).
- 'Folate' in plasma is a measure of total folate. However, we need to split this into the specific forms of folate in order for these to influence their specific biochemical pathways in the model. We therefore need to research the split between 10-fthf, thf and 5mthf and reflect this in the concentration input file.

2.3   Imputation of flux values for reactions with nutritional cofactors

The plasma metabolites acting as co-factors will not influence the flux distributions since they remain unchanged in the reactions. Fluxes can only be potentially altered if the plasma metabolites are substrates or products in a reaction. Attempting to model the co-factors in half equations would not change the flux distributions as they would simply cancel out. These co-factors relate to:

- Vitamin B6 via pyridoxal 5'phosphate. This is a co-factor in GHMT (EC 2.1.2.1), in CYSTS (EC 4.2.1.22) and in CYSTGL (EC 4.4.1.1).
- Vitamin B12 via methylcobalamin. This is a co-factor for METS (EC 2.1.1.13).
- Vitamin B2 via FMN and FAD. FAD is a co-factor for MTHFR (EC 1.5.1.20).

Since these co-factors are a potentially key focus of interest we need to manually impute the fluxes based on the co-factor plasma concentration data.

- Firstly using the overall population data determine the maximum and minimum flux of the reactions in question. We scale the range from 0-100%.
- We then obtain the minimum and maximum concentration of the relevant co-factor metabolites and scale that range from 0-100%.
- We finally impute the adjusted fluxes of the reactions based on the co-factor plasma concentration scale.

2.4   Deciding how to measure methylation potential

Currently the DNA methyltransferase reactions are not incorporated in the model. We might therefore choose to look at a variety of other reactions that use the SAM -> SAH (adomet -> ahcys) pathway and report the mean flux through these reactions as a proxy for 'methylation

potential'. A caveat is that there is no guarantee an increased flux through other SAM -> SAH reactions would mirror an increase in the DNMT reaction. Indeed, it could be argued that more methyl groups being used for other transmethylation reactions could reduce the amount available for DNA methylation (indeed, that is the reason for some trials supplementing with choline and creatine – to have a methyl group sparing effect). Having an average flux from several transmethylation reactions might help at least talk of methylation potential broadly, and it might be the best we can do for now.

2.5 Deciding the organ to isolate for detailed modeling work

Many of the modelling tools will work best if we isolate one organ to explore. The overall model is simply too large for many of the techniques (detailed in section 3). We suggest choosing an organ *a priori.* The liver is our primary suggestion as:

a) This is the organ that processes the nutrients contained in the plasma, coming from the intestine, before circulating onto other organs in the body.

b) We know that BHMT is predominantly expressed in the liver, and to a much lesser extent in the kidney. This reaction uses a methyl group from betaine to methylate homocysteine. Betaine can also be formed through the oxidation of choline. Therefore choosing the liver means our choline and betaine data can be modelled more directly.

2.6 Deciding key reactions of interest

The main reactions of interest have been selected *a priori* based on them a) being core reactions to 1-carbon metabolism or b) involving more distal metabolites that we think could influence 1-carbon metabolism (Table 1). These are selected for each organ / compartment they are found in and number ~350 of the total 80,000 reactions present in RECON-III.

Table 1: Key 1-carbon pathways chosen *a priori* for analysis

Reaction in Recon-III	Enzyme name	Reaction (EC number)	Nutrition-related hypothesis
MTHFR3	Methylenetetrahydrofolate Reductase	1.5.1.20	Increased B2 increased flux
METS	Methionine synthase	2.1.1.13	Increased B12 increased flux
METAT	Methionine Adenosyltransferase	2.5.1.6	Increased methionine increased flux
AHC	Adenosylhomocysteinase	3.3.1.1	Increased SAH increased flux and decreased flux of 2.1.1.37
DHFR	Dihydrofolate Reductase	1.5.1.3	Increased folate increase flux
GHMT2	Serine hydroxymethyltransferase,	2.1.2.1	Increased B6 increased flux
BHMT	Betaine Homocysteine Methyltransferase	2.1.1.5	Increased choline & betaine increased flux

Reaction in Recon-III	Enzyme name	Reaction (EC number)	Nutrition-related hypothesis
CYSTS	Cystathionine-Beta-Synthase	4.2.1.22	Increased B6 increased flux
MTHFD / MTHFD2	Methylenetetrahydrofolate Dehydrogenase	1.5.1.5	Decreased B2 increased flux? Increased B6 increased flux?
ADMDC	Adenosylmethionine decarboxylase	4.1.1.50	Increased flux here decreases flux of 2.1.1.37?
TMDS	Thymidylate synthase	2.1.1.45	Decreased B6 decreases flux?
PETOH	Phosphatidylethanolamine N-methyltransferase	2.1.1.17	PUFA intake increases flux? Decreases in turn flux of 2.1.1.37?
CHOL	Choline dehydrogenase	1.1.99.1	Increased choline increased flux?
BETALD	α-aminoadipic semialdehyde dehydrogenase	1.2.1.8	Increased choline increased flux?
Trans-methylation	Various	Various	Increased SAM:SAH ratio increased flux

2.7   Decide what the objective function should be

- The model is designed to enable maintenance of all essential physiological processes, and if a baby biomass component is added, to supply the needs of the growing baby.
- If we change the objective to optimise the flux of one particular pathway, e.g. a SAM->SAH pathway, we risk obtaining an unrealistic solution. The model would simply try to solve the equation to maximise that flux, and then ignore the other pathways needing to be maintained for survival. It is only when the objective is set to maintenance that we are more confident the plasma nutrients will be taken into the model.
- Therefore to gain the benefits that this systems biology model offers we should keep to the original objective function of maintenance.
- We then run different models according to months / seasons and see the flux results to identify the model that represents the best nutritional profile / flux through methylation reactions.
- In this way we're not actually saying one single reaction has to be optimized – which is beneficial as in reality we don't necessarily want to maximize flux through one single reaction.
- What will signpost us to the best model out there? From previous studies it would seem a higher methylation potential (higher SAM:SAH metabolite ratio) and higher levels of methyl donors & co-factors are associated with reduced loss of imprinting at certain genetic loci. We can use this as a starting point.

3   Preliminary analysis plan

3.1   Key objectives

Integrate maternal plasma biomarker data into RECON-III to:

A. Describe the seasonal pattern of the 1C metabolome and related metabolic pathways to explore whether the observed patterns fit in with existing knowledge of metabolism. Do the metabolites generally co-vary in ways that we expect?
B. Identify key drivers of methylation potential.

Then integrate the infant epigenomic data to:

C. Explore the extent to which the maternal metabolome is associated with infant methylation patterns at key genomic regions.
D. Model how a maternal metabolome can be optimised *e.g*. how do we move metabolomes associated with suboptimal epigenetic patterns (e.g. loss of imprinting) towards a more protective one?

3.2   Objective A: Compare fluxes of key reactions by month and season

- Input all the concentration data and obtain the distribution of fluxes through the metabolic pathways for each woman.

3.2.1 Primary analysis

- We can focus on main pathways of interest (section 2.6) and summarise the mean flux by month of plasma collection. We could initially report these by organ, however, our organ of interest (section 2.5) might be the only one we choose to finally publish based on the sheer volume of data we will get.
- We could also repeat an analysis looking at the months corresponding to the peak of the dry (Feb – Apr) and peak of the rainy season (July-Oct) (e.g. Table 2).
- We could test differences in fluxes by month / season using a one-way ANOVA for more than two groups (*i.e.* the months), or the t-test for two groups (*i.e.* the seasons). If the fluxes are not normally distributed we could consider the Kruskal Wallis and Wilcoxon Mann-Whitney tests respectively. Considering we have selected pathways *a priori* we could use a cut-off of $p < 0.05$ to determine significant differences in fluxes.

Table 2: Mean flux through key 1-carbon reactions in the liver by season

Reaction	Equation	Mean Rainy season flux	Mean Dry season flux	P value
'AHC'	'h2o[c] + ahcys[c] <=> adn[c] + hcys_L[c] '			
'BHMT'	'hcys_L[c] + glyb[c] -> dmgly[c] + met_L[c] '			
'CYSTS'	'hcys_L[c] + ser_L[c] -> h2o[c] + cyst_L[c] '			
'DHFR'	'h[c] + nadph[c] + dhf[c] <=> nadp[c] + thf[c] '			
'GHMT2r'	'thf[c] + ser_L[c] <=> h2o[c] + gly[c] + mlthf[c] '			
'GNMT'	'amet[c] + gly[c] -> h[c] + ahcys[c] + sarcs[c] '			
'METAT'	'h2o[c] + atp[c] + met_L[c] -> pi[c] + amet[c] + ppi[c] '			
'METS'	'5mthf[c] + hcys_L[c] -> h[c] + thf[c] + met_L[c] '			
'MTHFD'	'nadp[c] + mlthf[c] <=> nadph[c] + methf[c] '			
'MTHFR3'	'2 h[c] + nadph[c] + mlthf[c] -> nadp[c] + 5mthf[c] '			

- We may also consider more complex multivariate approaches, for example using principal components to look at broader-scale flux changes across multiple reactions
- We could consider taking the women in the two seasons separately and constraining two models based on population mean fluxes (*i.e.* one representative dry season and wet season model). We might need to pick more 'extreme' individuals to represent the seasonal differences. Then we could isolate one organ and perform a **flux variability analysis** (FVA) on our key reactions to see the range of possible fluxes for that reaction in the given constraints (Table 3).

Table 3: Example of flux variability analysis to determine maximum and minimum fluxes through key 1-carbon reactions by season

Reaction	Equation	Rainy season		Dry season	
		Min flux	Max flux	Min flux	Max flux
'AHC'	'h2o[c] + ahcys[c] <=> adn[c] + hcys_L[c] '				
'BHMT'	'hcys_L[c] + glyb[c] -> dmgly[c] + met_L[c] '				
'CYSTS'	'hcys_L[c] + ser_L[c] -> h2o[c] + cyst_L[c] '				
'DHFR'	'h[c] + nadph[c] + dhf[c] <=> nadp[c] + thf[c] '				
'GHMT2r'	'thf[c] + ser_L[c] <=> h2o[c] + gly[c] + mlthf[c] '				
'GNMT'	'amet[c] + gly[c] -> h[c] + ahcys[c] + sarcs[c] '				
'METAT'	'h2o[c] + atp[c] + met_L[c] -> pi[c] + amet[c] + ppi[c] '				
'METS'	'5mthf[c] + hcys_L[c] -> h[c] + thf[c] + met_L[c] '				
'MTHFD'	'nadp[c] + mlthf[c] <=> nadph[c] + methf[c] '				
'MTHFR3'	'2 h[c] + nadph[c] + mlthf[c] -> nadp[c] + 5mthf[c] '				

- Table 3 gives the range of fluxes but not a precise idea of the 'true' solution. We could therefore also perform a **sampling frame** on these reactions to generate density plots of the most likely fluxes and map the results to the main pathways using the example from Aurich *et al.* 2014[7] (Figure 7).
- Due to the number of reactions involved and the redundancy of certain metabolic pathways, the flux distributions predicted by the FVA are not unique and there exist a large number of alternate optimal solutions. Hence, we use methods such as 'sparseFBA' to get a unique solution (sparseFBA limits the number of reaction fluxes). The idea is that this will produce a clearer signal of the effect of our constraints. When we predict a unique solution for each model, we can have increased confidence that the differences we see are real, rather than just due to the existence of alternative solutions. This means we need to make sure the model is forced to use our key reactions of interest.
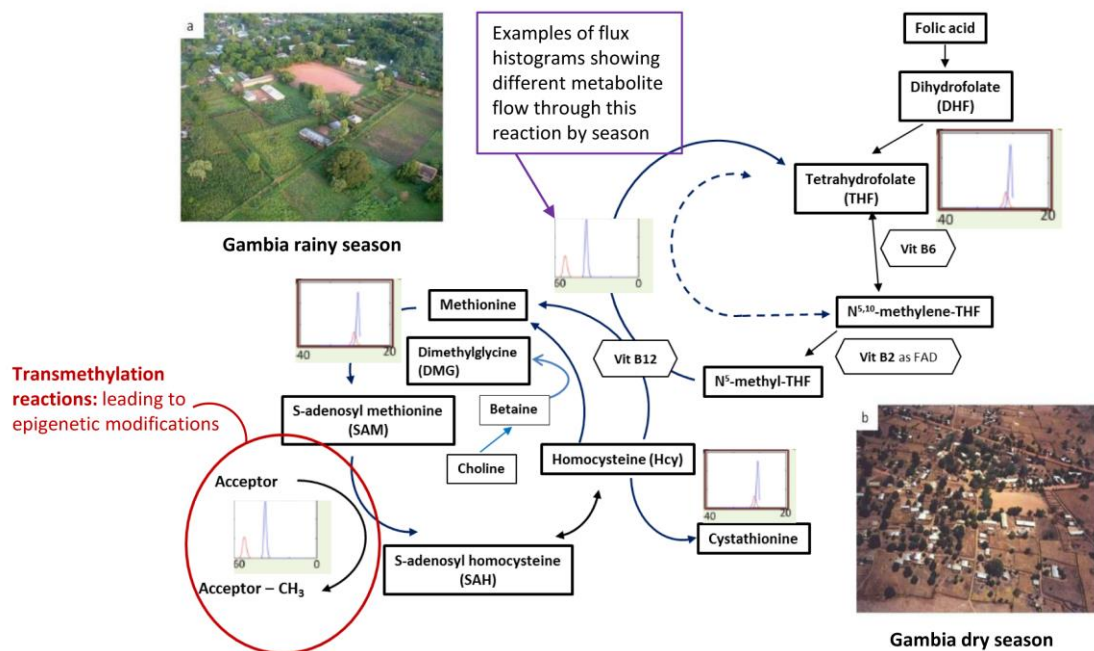
Fig. 7: An example of how we could summarise the density plots generated by sampling analysis to show flux through key pathways by season. Photo credit: Andrew M Prentice.

3.3 Objective B: Identify drivers of methylation potential

- For example, we could we create two representative seasonal models that have the mean plasma concentrations from those months, or otherwise choose the two most extreme individuals showing greatest variation by season.
- We would then either optimise models according to our objective function, or just keep the fluxes as they are from the mean values (depending on decision from section 2.7).
- If we did choose to optimise the models for an objective function, e.g. GNMT flux to represent methylation potential, we could end up with **a flux balance analysis** along the lines of Table 4 below, otherwise we just report the mean fluxes making up our models.

Table 4: Example of optimizing a cell line model to maximize flux through GNMT and the resulting vector of fluxes through other key 1-carbon pathways to achieve that.

Reaction	Equation	Flux when GNMT optimised
'AHC'	'h2o[c] + ahcys[c] ⇔ adn[c] + hcys_L[c] '	4.749650462
'BHMT'	'hcys_L[c] + glyb[c] -> dmgly[c] + met_L[c] '	0
'CYSTS'	'hcys_L[c] + ser_L[c] -> h2o[c] + cyst_L[c] '	0
'DHFR'	'h[c] + nadph[c] + dhf[c] ⇔ nadp[c] + thf[c] '	0.000370376
'GHMT2r'	'thf[c] + ser_L[c] ⇔ h2o[c] + gly[c] + mlthf[c] '	0.471297587
'GNMT'	'amet[c] + gly[c] -> h[c] + ahcys[c] + sarcs[c] '	4.749650462
'METAT'	'h2o[c] + atp[c] + met_L[c] -> pi[c] + amet[c] + ppi[c]	4.749650462
'METS'	'5mthf[c] + hcys_L[c] -> h[c] + thf[c] + met_L[c] '	4.749650462
'MTHFD'	'nadp[c] + mlthf[c] ⇔ nadph[c] + methf[c] '	4.080110894
'MTHFR3'	'2 h[c] + nadph[c] + mlthf[c] -> nadp[c] + 5mthf[c] '	4.749650462

- Regardless of whether we pick an objective function or just use the mean values from our chosen month / season / extreme individual, we can perform flux variability analysis on these models e.g. Table 5.

Table 5: Example of FVA using a cell line model where the flux through GNMT has been optimized

Reaction	Equation	Min flux	Max flux
'AHC'	'h2o[c] + ahcys[c] <=> adn[c] + hcys_L[c] '	4.749650	4.764032
'BHMT'	'hcys_L[c] + glyb[c] -> dmgly[c] + met_L[c] '	0	2.53E-13
'CYSTS'	'hcys_L[c] + ser_L[c] -> h2o[c] + cyst_L[c] '	0	0
'DHFR'	'h[c] + nadph[c] + dhf[c] <=> nadp[c] + thf[c] '	0.000370	0.000370
'GHMT2r'	'thf[c] + ser_L[c] <=> h2o[c] + gly[c] + mlthf[c] '	0.027421	7.607178
'GNMT'	'amet[c] + gly[c] -> h[c] + ahcys[c] + sarcs[c] '	4.749650	4.749650
'METAT'	'h2o[c] + atp[c] + met_L[c] -> pi[c] + amet[c] + ppi[c] '	4.749650	4.764032
'METS'	'5mthf[c] + hcys_L[c] -> h[c] + thf[c] + met_L[c] '	4.749650	4.764032
'MTHFD'	'nadp[c] + mlthf[c] <=> nadph[c] + methf[c] '	-500	500
'MTHFR3'	'2 h[c] + nadph[c] + mlthf[c] -> nadp[c] + 5mthf[c] '	4.749650462	4.764032197

3.4   Objective C and D: Integration of epigenetic data

- How we integrate the epigenetic data remains an open question pending the methylation analysis.
- We may keep the methylation and metabolome analyses separate, but use the former to inform the latter. For example, we could stratify the women as per epigenetic outcomes in their children and compare their metabolomes:
  - We could take total mean methylation across the whole genome, or across certain classes of gene, and categorise women into two or three groups.
  - We could take one imprinted gene and use a binary loss of imprinting / normal imprinting outcome to categorise the women.
  - To be developed.

References

1.   Brunk E, Sahoo S, Zielinski DC, et al. (2018) Recon3D enables a three-dimensional view of gene variation in human metabolism. *Nat. Biotechnol.* **36**, 272–281.

2.   Aurich MK & Thiele I (2016) Computational Modeling of Human Metabolism and Its Application to Systems Biomedicine. *Methods Mol. Biol.* **1386**, 253–81.

3.   Becker SA, Feist AM, Mo ML, et al. (2007) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nat. Protoc.* **2**, 727–38.

4.   Orth JD, Thiele I & Palsson BØ (2010) What is flux balance analysis? *Nat. Biotechnol.* **28**, 245–8.

5.   Dominguez-Salas P, Moore SE, Cole D, et al. (2013) DNA methylation potential: dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women. *Am. J. Clin. Nutr.* **97**, 1217–27.

6.   Green R & Miller JW (2007) Vitamin B12. In *Handbook of Vitamins*, 4th ed. [ZempleniJ, Rucker R, McCormick D, et al., editors]. Boca Ralton, FL: Taylor & Francis Group.

7.   Aurich MK, Paglia G, Rolfsson Ó, et al. (2014) Prediction of intracellular metabolic states from extracellular metabolomic data. *Metabolomics* **11**, 603–619.

ANNEX 4.2   A Primer on Fourier series regression models and how they will be used to back-extrapolate biomarkers

Seasonal trends are periodic fluctuations around a mean that can be modelled using Fourier terms in linear regression models. These models introduce sine and cosine waves to represent the seasonal fluctuation of biomarker concentrations.

Step 1: Calculate the date of measurement ('booking' / enrolment date) in radians ('θ')

- The timespan 1st January to 31st December (365 days) can also be represented as 0 to 2π radians. This is necessary in order to convert time into a cyclical variable, so that in a Fourier model, e.g. 31st December is 'adjacent' to 1st January: $\sin\theta^{31/12} \sim \sin\theta^{1/1}$
- The easiest way to do this is represent the date of booking as a value between 0 and 1 (i.e. a fraction of 365) then multiply that by 2π to express that date in radians.
- Take the example 01/03/2010 as an example of our measurement date. We subtract this date from the beginning of the year (01/01/2010) to get the number of days into the year our measurement date is: 01/03/2010 – 01/01/2010 = 59 days into the year.
- We express this as a fraction of 365 to get a value between 0 and 1: 59/365 = 0.162
- We multiply this by 2π to express this in radians (our θ value): 0.162 x 2π = 1.015 radians

Step 2: Create the Fourier terms

- Fourier terms (FTs) take the sine and cosine of our θ values.
- i.e. first pair of FTs (sin1, cos1) = sin(θ), cos(θ)
- Second pair (sin2, cos2) = sin(2θ), cos(2θ)
- Go up to 4 pairs of terms (sin4, cos4).
- These FTs create cyclical patterns that we can use to describe our seasonal trends.

Step 3: Decide on whether to transform the biomarker variable

- Assess the normality of the biomarker data by performing a skewness-kurtosis test.
- E.g. for folate the data is highly skewed, and log-transforming it creates a normally distributed variable.
- We will therefore use log-transformed folate throughout the following steps.
- Note in Chapter 4 we use the pre-adjusted log folate, after taking account of BMI, age, gestational age and inflammation. In this document we will just take log-folate and later adjust for gestational age, to keep the same methodology as

Step 4: Decide on how many pairs to FTs to keep in the model

We are going to regress our biomarker of interest (log folate) against the FTs (independent variables), successively adding in one pair of terms to find our best fit. E.g. Let us decide the seasonal trend for folate:

- Regress log folate (alone) to obtain the null model (*i.e.* no seasonal trend)
- Store the estimates from this regression ('estimates A')
- Regress log folate against the first pair of FTs (sin1, cos1)
- Store the estimates ('estimates B').
- Perform a likelihood ratio test (LRT) comparing estimates B with estimates A.
- If p<0.05 we are justified in keeping that first pair of FTs.
- If p>0.05 we have no evidence to suggest that adding in the FTs is any better than the null model. We would therefore declare there is no seasonal trend.
- Assuming there is evidence for a seasonal trend we then regress log folate with the first two pairs of FTs and store the estimates ('estimates C').
- Perform a LRT comparing estimates C with B to see whether two pairs of FTs fits better than one pair.
- We keep going in this manner until the LRT is not significant (p>0.05), signalling there is no extra value in adding the additional pair of terms.
- For each biomarker we suggest considering a maximum of 4 pairs of FTs.
- For log folate in MDEG2 we find that 3 pairs of FTs gives us the best fit.

Step 5: Visualise the seasonal trend

- It can be helpful at this stage to plot the seasonal trend.
- For example, with log folate we want to plot the fitted predictions with 3 pairs of FTs.
- To do this regress log folate against 3 pairs of FTs, and store the predicted values. These predictions, when plotted against the booking date (θ) will give the smoothed seasonal trends.
- Also store the standard error (SE) of the predicted values if you want to plot the 95% CIs. Stata has a program to store the SE of the predictions with one command. You can then generate the upper and lower intervals using your predicted values ± 1.96*SE.
- This is what we get for 3 pairs of FTs for log folate with 95% CIs plotted:
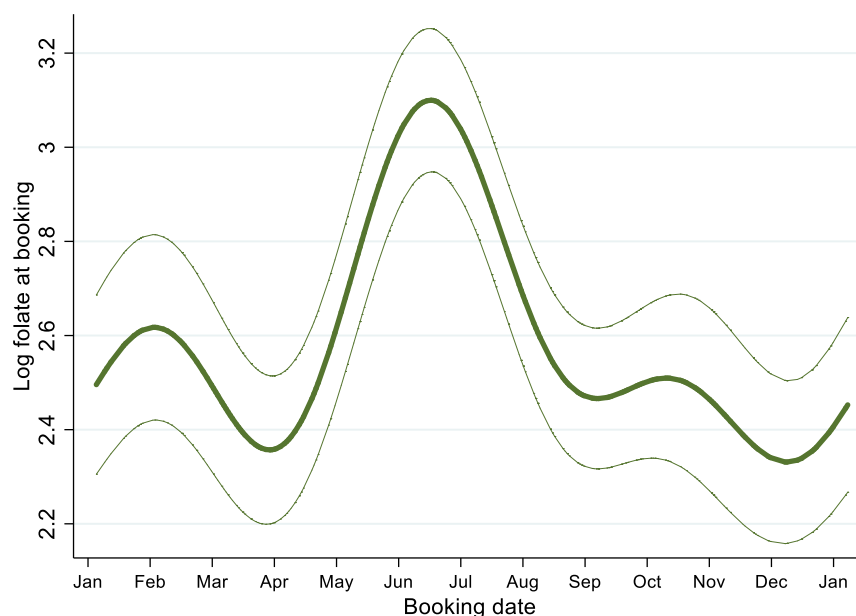
This step is what we presented in Chapter 4. For future research we would like to back-extrapolate the biomarker concentrations to predict their value at the time of conception. The following steps show how we use the above Fourier term regression models to do that.

Step 6: Write the linear equation for the seasonal trend

- Use the constant and coefficients from the linear regression model with the FTs to write the linear equations for the predicted values.
- For example this is the output we get when we regress log folate against 3 pairs of FTs:

Log folate	Coefficient	Std. Err.	t	P value	Lower 95% confidence interval	Upper 95% confidence interval
sin1	.0704035	.0459901	1.53	0.127	-0.0202481	0.1610552
cos1	-.2125245	.0463401	-4.59	0.000	-0.303866	-0.1211829
sin2	-.0415646	.0471592	-0.88	0.379	-0.1345207	0.0513915
cos2	.1484138	.0453836	3.27	0.001	0.0589577	0.2378699
sin3	.1437096	.0445225	3.23	0.001	0.0559508	0.2314684
cos3	-.0529964	.0470574	-1.13	0.261	-0.1457518	0.039759
constant	2.584036	.0326402	79.17	0.000	2.519698	2.648373

- The equation for predicted folate with 3 pairs of FTs is therefore:

Log Folate (predicted) = 2.584036 + 0.0704035*sin1 -0.2125245*cos1 -0.0415646*sin2 + 0.1484138*cos2 + 0.1437096*sin3 - 0.0529964*cos3

Where sin1 = sin(θ), sin2 = sin(2θ) etc. and θ is the date of booking in radians, as per Steps 1 and 2.

- This gives us the overall (mean) shape of the seasonal trend for the whole population.

Step 7: Obtain the intercept for the individual values

- Remember the above equation gives the overall population seasonal trend. The coefficients for the FTs describe the shape of the trajectory.
- The constant (intercept) will shift that trajectory vertically up or down according to the individual woman's actual folate concentration at the time of blood measurement (booking).
- We therefore need to obtain the intercept (I_j) for woman j based on her measured concentration at booking (log folate$_{bj}$).
- We therefore re-arrange the equation:
- Intercept (I_j) = log folate$_{bj}$ - 0.0704035*sin(θ_{bj}) +0.2125245*cos(θ_{bj})+ 0.0415646*sin($2\theta_{bj}$) - 0.1484138*cos($2\theta_{bj}$) - 0.1437096*sin($3\theta_{bj}$) + 0.0529964*cos($3\theta_{bj}$)
  where θ_j is the booking date (when folate was measured) for woman *j.*

Step 8: Obtain the date of conception in radians

- First calculate the date of conception for woman *j*
- This is done by subtracting the gestational age (days) from the date of booking.
- Then follow the process outlined in Step 1 to obtain the date of conception expressed in radians (θ_{cj}).

Step 9: Calculate the new Fourier terms

- Now calculate the new FTs using the date of conception radians.
  - $\sin1_{cj} = \sin(\theta_{cj})$
  - $\cos1_{cj} = \cos(\theta_{cj})$
  - $\sin2_{cj} = \sin(2\theta_{cj})$ etc.

Step 10: Calculate the predicted biomarker concentration at the time of conception

- Use the intercepts calculated for each individual (I_j) [step 7] , the coefficients from the linear equation [step 6] and the FTs using the date of conception radians [step 9] to calculate the predicted biomarker concentration at conception:

Predicted log folate at conception = I_j + 0.0704035*$\sin(\theta_{cj})$ -0.2125245*$\cos(\theta_{cj})$ - 0.0415646*$\sin(2\theta_{cj})$ + 0.1484138*$\cos(2\theta_{cj})$ + 0.1437096*$\sin(3\theta_{cj})$ - 0.0529964*$\cos(3\theta_{cj})$

- This takes the plasma concentration of log folate at blood measurement (booking) and then tracks back using the seasonal trajectory to the time of conception.
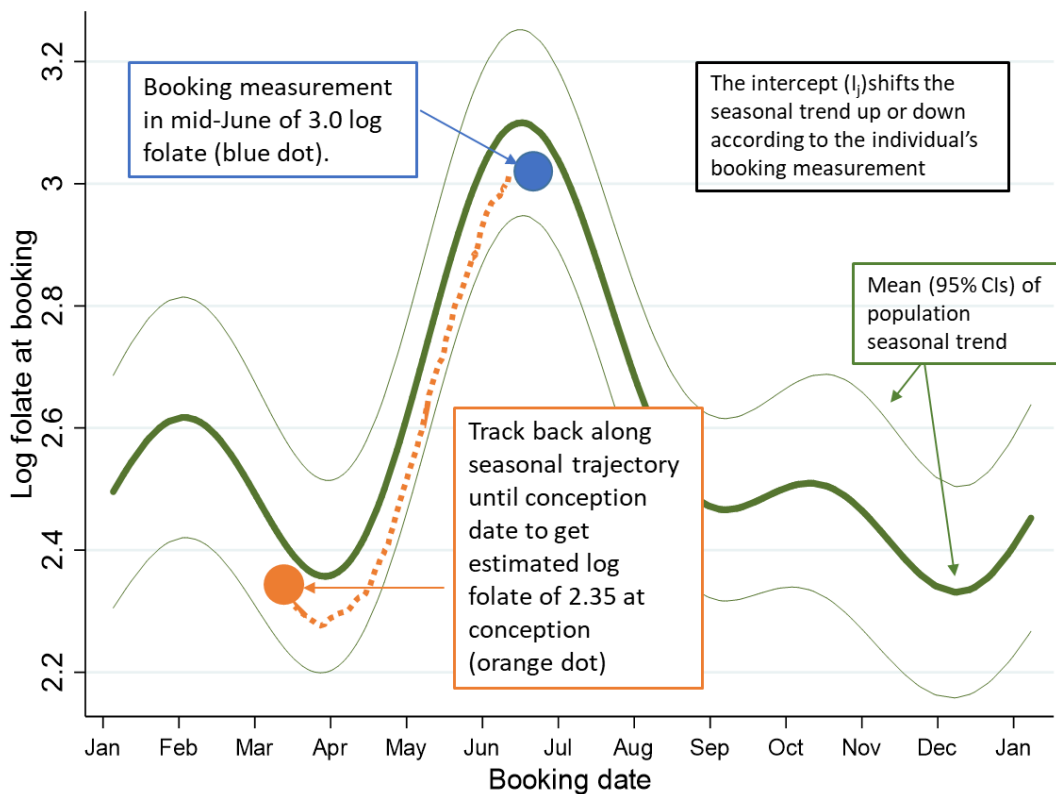
Checking the first few entries shows this has worked (using the graph in step 5 to help verify the changes make sense):

Booking date	Gestational age (days)	Conception date	Log folate at booking	Predicted log folate at conception
30 Jun 10	75	16 Apr 2010	2.901422	2.325179
14 Jul 10	103	02 Apr 2010	3.091951	2.590727
21 Jul 12	108	04 Apr 2010	3.468233	3.063819
29 Jul 10	102	18 Apr 2010	2.861686	2.655936
21 Jul 13	94	18 Apr 2010	2.183802	1.877994

- Another way to check this has worked is to plot the predicted log folate (with 3 pairs of FTs) at conception against conception date (θ), using the same method outlined in Step 5. If it has worked the graph (predicted conception concentration vs. conception date) should be almost identical to the graph in Step 5 (booking concentration vs. booking date).
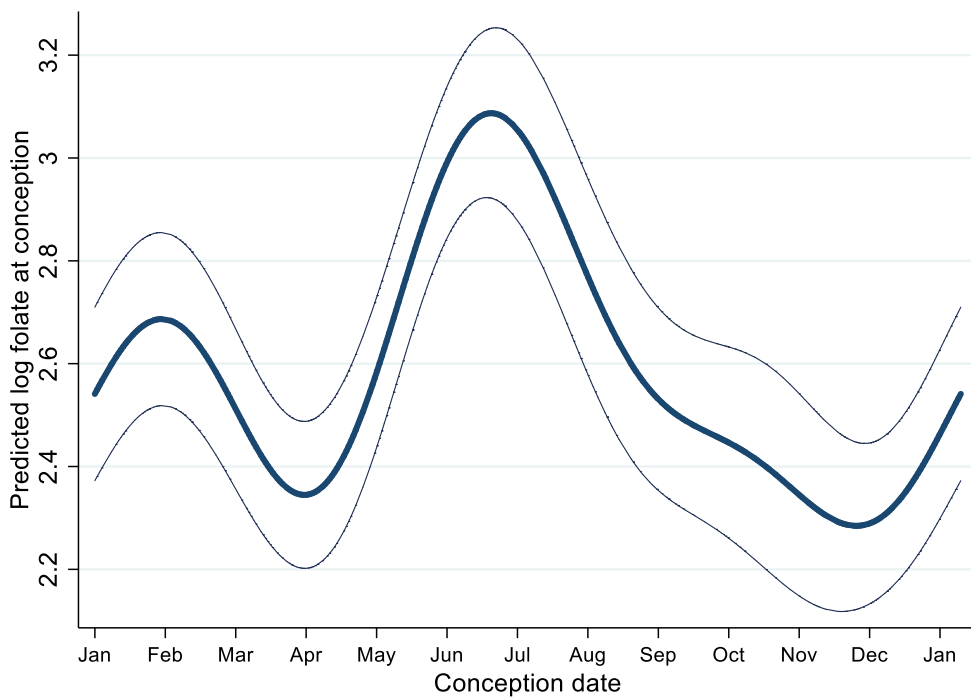
- Note that the shape of the curve is not exactly identical to the booking folate seasonal trend. This is because gestational age at booking varies slightly between women, so that there is not a straightforward translation of the means and CIs along the x-axis. (Strictly, mean gestational age at booking varies across the year).

Step 11: Adjust the predicted values at conception for gestational age

- Regress the predicted log folate at conception values against gestational age and store the residuals. This is because stage of pregnancy can influence biomarker concentrations through haemodilution and other physiological processes.

Predicted folate at conception	Coefficient	Std. Err.	t	P value	Lower 95% confidence interval	Upper 95% confidence interval
Gestational age (days)	-.0001341	.0014913	-0.09	0.928	-0.0030732	0.002805
Constant	2.593367	.1510821	17.17	0.000	2.295606	2.891128

- Add the residuals to the intercept (2.593367) to give the final predicted values adjusted for gestational age.
- Finally, calculate the exponentials (of natural log) of the predicted adjusted log folate at conception.
- Example of first few entries below:

Predicted folate at conception	Residuals	Intercept	Predicted log folate at conception adjusted for gestational age	Predicted folate at conception adjusted for gestational age (nmol/L)
2.325179	-.2581314	2.593367	2.335236	10.33189
2.590727	.0111711	2.593367	2.604538	13.52497
3.063819	.4849338	2.593367	3.078301	21.72146
2.655936	.0762461	2.593367	2.669613	14.43438
1.877994	-.7027685	2.593367	1.890599	6.623332

- Plotting the final shape of the seasonal trend of predicted log folate at conception adjusted for gestational age confirms that gestational age does not have a large effect on our biomarker concentrations.
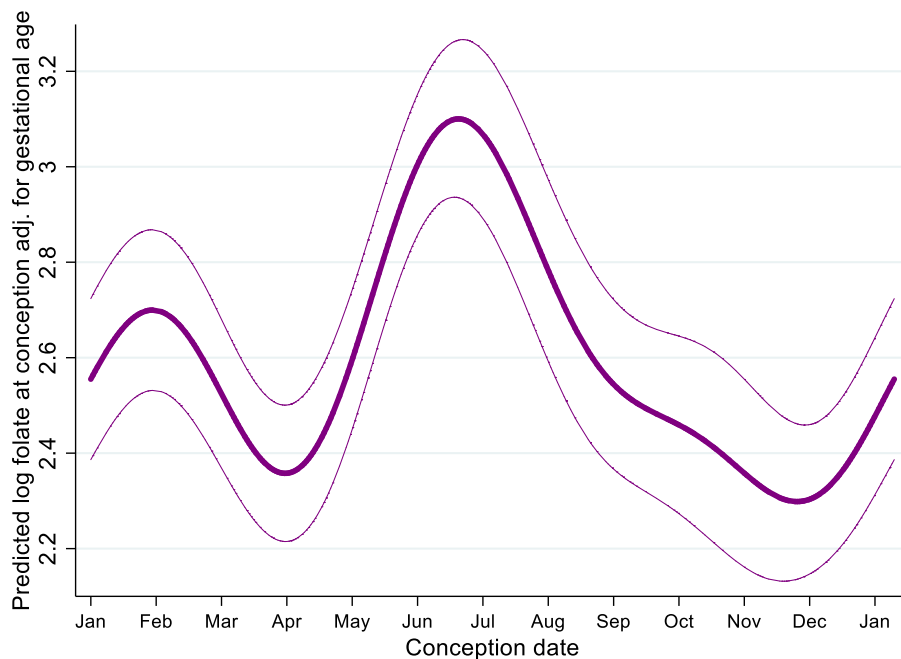
A comparison with the MDEG Main Study 2014 method

- For a simplified summary of the methodology refer to the Dominguez-Salas *et al.* (2014; PMID 24781383) paper where the first section of the methodology explains the back-extrapolation. In the words of the statistician: "To estimate biomarker concentrations at the time of conception we back-extrapolated along a trajectory parallel to the seasonal patterns (fitted by Fourier regression) …To account for pregnancy-mediated changes in biomarkers the values were further adjusted for the gestational age of the infant at the time of measurement."
- Note, however, for the paper the MDEG1 main study biomarkers were back-extrapolated using the Fourier term coefficients from a separate dataset (the 'Indicator Group').
- The difference in this method is that we are using the seasonal trend from the same dataset (MDEG-2) to back-extrapolate the biomarkers; there is no indicator group.
- This is mainly because we have an expanded set of biomarkers for MDEG-2, many of which were not included in the Indicator Group.
- The Indicator Group comprises repeated monthly measures from a cohort of 30 women followed for one year. The MDEG-2 dataset comprises cross-sectional information from different women each month over one year.
- Why are we justified in back-extrapolating along a seasonal trend trajectory generated from the same dataset?
  - Firstly, when we compare the seasonal trends from the MDEG-2 data with the Indicator Group data the patterns are very similar, with the peaks and troughs of the curves corresponding fairly well (results in Chapter 4).
  - Secondly, inspecting the results of the Fourier term random effects linear regression models of the Indicator Group shows that the predicted fit explains a much greater proportion of intra-individual variation than inter-individual variation. This suggests the seasonal differences within an individual are much stronger than the between-person differences. Therefore the MDEG2 data can still be used to capture seasonal differences even though we do not have repeated measures in the same person.

ANNEX 4.3 Core one-carbon biomarker details on metabolism and factors affecting plasma concentration

Nutrient	Main role in one-carbon pathways	Correlation between plasma biomarker & intake data	Do plasma levels change in pregnancy?	Stability on storage and free-thaw cycles
Folate	The folate assay provides information on tetrahydrofolate (THF), which is reduced to methylene-THF, then to 5-methyl-THF, the form which donates its methyl group to homocysteine using vitamin B12[1]. The assay used in MDEG-2 is a total folate assay. Its limitation is that it shows a strong bias for 5-methyl-THF, so it does not perform well when the population in enriched with the MTHFR 677C->T genetic variant, but this allele is present in only 3.2% of the Gambian population (Keneba Biobank data).	Conflicting views: some document no correlation[2] whereas others suggest it can reflect recent intake[3].	Conflicting views: Slight decrease in serum folate over pregnancy, after 1st trimester[4]. Slight increase in plasma folate gestational week (GW) 9-16[5] Folate status alters the change in concentration during pregnancy in betaine, DMG and choline[6]	Stable for storage at -70°C[7,8]. Stable even at room temperature for 72 hours in serum[9]. Degrades rapidly at room temperature at DT_{50} of 45.2 h[10], worst in EDTA compared to other matrices. Degrades at -25°C at DT_{50} of 23.6 years[10]
Vitamin B12 (cobalamin)	Required as a cofactor in the methylation of homocysteine by methyl tetrahydrofolate[11].	Not influenced by recent intake[12]	Decreases over pregnancy[4,5,13].	Stable for storage at -70°C[7,8] Stable in EDTA at room temperature for 8 days and at -25°C for 29 years[10]
Pyridoxal (PL), pyridoxal-5'-phosphate (PLP), and pyridoxic acid (PA).	These are 3 of the 7 vitamers that exist for vitamin B6[14]. In previous studies concentrations of these three vitamers have followed a very similar seasonal pattern[2]. PLP is required to reduce THF to methylene-THF, and is also a cofactor in the transsulfuration pathway converting homocysteine to cysteine[11].	Positive association[2]	Gradual decline in PLP throughout pregnancy[4], especially in 3rd trimester (although may not be due to haemodilution)[3].	PA is stable at room temperature for 8 days and at -25°C for 29 years[10]. PLP in serum degrades over time at -25°C at rate of DT_{50} = 14.7 years[10]. Note degraded PLP tends = recovery of PL, so PL+PLP relatively stable, and EDTA inhibits degradation[10]

Nutrient	Main role in one-carbon pathways	Correlation between plasma biomarker & intake data	Do plasma levels change in pregnancy?	Stability on storage and free-thaw cycles
Riboflavin (vitamin B2)	Required as a precursor to FAD, which is used as cofactor for MTHFR to reduce methylene-THF to methyl-THF[15].	Erythrocyte glutathione reductase (EGRAC) is more common than plasma B2. Plasma B2 does not seem to vary much with intake[3].	Minor changes only; slight increase in third trimester[4]	Accumulates 3-4% per hour for first 13 hours at room temperature in serum, but only marginally in EDTA[10]
Choline	Precursor to betaine (two-step oxidation reaction)[16]. Synthesised endogenously or obtained from diet (good sources: red meat, poultry, milk, eggs, fish[17]).	No correlation[2,18,19]	Increases over pregnancy from maternal hepatic stores[5,18], used for increased phosphatidylcholine synthesis in favour of betaine synthesis[20]. Concentration maintained to allow for active transport to foetus[16]	Free choline in serum accumulates 3-5% per hour at room temperature and 8.68% per year stored at -25°C but is stable in EDTA[10].
Betaine	Donation of methyl group to homocysteine via betaine-homocysteine methyl transferase (BHMT)[16]. Obtained in diet (good sources: wheat bran, wheat germ, spinach, beets[17]) or oxidation of choline. Eventually catabolised to glycine.	No correlation[18] Positive association[2]	Decreases over pregnancy to GW24[5,6,18,21] due to increased methyl donor demand in foetus[22]	Stable in EDTA at room temperature for 8 days and at -25°C for 29 years[10].
Dimethylglycine (DMG)	Product formed after the betaine has methylated homocysteine. Increased DMG levels then inhibit this reaction by reducing activity of BHMT[23]. Also a derivative of glycine[16].	N/A: an intermediary metabolite	Decreases in early pregnancy[5], increases after GW16[6,18].	Stable in EDTA at room temperature for 8 days and at -25°C for 29 years[10]

Nutrient	Main role in one-carbon pathways	Correlation between plasma biomarker & intake data	Do plasma levels change in pregnancy?	Stability on storage and free-thaw cycles
Methionine	Condensed with ATP to form SAM. It can be formed from the methylation of Hcy or gained directly from the diet. Increased methionine upregulates GNMT[24]	No correlation[2]	Increases over pregnancy[21]	Oxidised to methionine sulfoxide over time at -25°C at rate of DT_{50} = 5.5 years in serum, but more stable in EDTA[10]
Cysteine	An amino acid that can be formed from homocysteine, via cystathionine and PLP as a cofactor, through the transsulfuration pathway.	Not assessed in[2]	Undocumented	Stable in EDTA at room temperature for 8 days and at -25°C for 29 years[10]
Homocysteine	Held in equilibrium with SAH and is either degraded by the transsulfuration pathway or methylated by methyl-THF or betaine to form methionine.	N/A: an intermediary metabolite	Slight decline in first trimester[4,5]. Increase from GW 13-delivery[21].	Stable in EDTA at room temperature for 8 days and at -25°C for 29 years[10]. Particularly stable if plasma separated from cells quickly. Repeated freeze-thaw cycles may increase concentration[3].
S-adenosyl methionine (SAM)	Methylation of cytosine, facilitated by DNA methyl transferases (DNMTs), leading to epigenetic modifications. Elevated SAM inhibits MTHFR and BHMT[25]	N/A: an intermediary metabolite	Undocumented	Concentration after 2nd freeze-thaw cycle is 3.53% lower than after 1st free-thaw cycle (UBC unpublished data).
S-adenosyl homocysteine (SAH)	Product formed after SAM methylates an acceptor. Increased SAH reduces the SAM methylation reaction rates.	N/A: an intermediary metabolite	Undocumented	Concentration after 2nd freeze-thaw cycle is 16.7% higher than after 1st free-thaw cycle (UBC unpublished data).

References

1. Scott JM & Weir DG (1998) Folic acid, homocysteine and one-carbon metabolism: a review of the essential biochemistry. *J. Cardiovasc. Risk* **5**, 223–7.

2. Dominguez-Salas P, Moore SE, Cole D, et al. (2013) DNA methylation potential: dietary intake and blood concentrations of one-carbon metabolites and cofactors in rural African women. *Am. J. Clin. Nutr.* **97**, 1217–27.

3. Gibson R (2005) *Principles of Nutritional Assessment*. 2nd ed. Oxford University Press.

4. Cikot RJLM, Steegers-Theunissen RPM, Thomas CMG, et al. (2007) Longitudinal vitamin and homocysteine levels in normal pregnancy. *Br. J. Nutr.* **85**, 49.

5. Velzing-Aarts F V, Holm PI, Fokkema MR, et al. (2005) Plasma choline and betaine and their relation to plasma homocysteine in normal pregnancy. *Am. J. Clin. Nutr.* **81**, 1383–9.

6. Fernàndez-Roig S, Cavallé-Busquets P, Fernandez-Ballart JD, et al. (2013) Low folate status enhances pregnancy changes in plasma betaine and dimethylglycine concentrations and the association between betaine and homocysteine. *Am. J. Clin. Nutr.* **97**, 1252–9.

7. Brinc D, Chan MK, Venner AA, et al. (2012) Long-term stability of biochemical markers in pediatric serum specimens stored at -80 °C: a CALIPER Substudy. *Clin. Biochem.* **45**, 816–26.

8. Jansen EHJM, Beekhof PK, Cremers JWJM, et al. (2012) Long-term (in)stability of folate and vitamin B12 in human serum. *Clin. Chem. Lab. Med.* **50**, 1761–3.

9. Oddoze C, Lombard E & Portugal H (2012) Stability study of 81 analytes in human whole blood, in serum and in plasma. *Clin. Biochem.* **45**, 464–9.

10. Hustad S, Eussen S, Midttun Ø, et al. (2012) Kinetic modeling of storage effects on biomarkers related to B vitamin status and one-carbon metabolism. *Clin. Chem.* **58**, 402–10.

11. Selhub J (1999) Homocysteine metabolism. *Annu. Rev. Nutr.* **19**, 217–46.

12. Green R (2011) Indicators for assessing folate and vitamin B-12 status and for monitoring the efficacy of intervention strategies. *Am. J. Clin. Nutr.* **94**, 666S–72S.

13. Molloy AM, Kirke PN, Brody LC, et al. (2008) Effects of folate and vitamin B12 deficiencies during pregnancy on fetal, infant, and child development. *Food Nutr. Bull.* **29**, S101-11.

14. Ueland PM, Ulvik A, Rios-Avila L, et al. (2015) Direct and Functional Biomarkers of Vitamin B6 Status. *Annu. Rev. Nutr.* **35**, 33–70.

15. Mason JB (2003) Biomarkers of Nutrient Exposure and Status in One-Carbon (Methyl) Metabolism. *J. Nutr.* **133**, 941S–947.

16. Ueland PM (2011) Choline and betaine in health and disease. *J. Inherit. Metab. Dis.* **34**, 3–15.

17. Cho E, Zeisel SH, Jacques P, et al. (2006) Dietary choline and betaine assessed by food-frequency questionnaire in relation to plasma total homocysteine concentration in the Framingham Offspring Study. *Am. J. Clin. Nutr.* **83**, 905–11.

18. Visentin CE, Masih S, Plumptre L, et al. (2015) Maternal Choline Status, but Not Fetal Genotype, Influences Cord Plasma Choline Metabolite Concentrations. *J. Nutr.* **145**, 1491–7.

19. Abratte CM, Wang W, Li R, et al. (2009) Choline status is not a reliable indicator of moderate changes in dietary choline consumption in premenopausal women. *J. Nutr. Biochem.* **20**, 62–9.

20. Yan J, Jiang X, West AA, et al. (2013) Pregnancy alters choline dynamics: results of a randomized trial using stable isotope methodology in pregnant and nonpregnant women. *Am. J. Clin. Nutr.* **98**, 1459–67.

21. Wallace JM, Bonham MP, Strain J, et al. (2008) Homocysteine concentration, related B vitamins, and betaine in pregnant women recruited to the Seychelles Child Development Study. *Am. J. Clin. Nutr.* **87**, 391–7.

22. Friesen RW, Novak EM, Hasman D, et al. (2007) Relationship of dimethylglycine, choline, and betaine with oxoproline in plasma of pregnant women and their newborn infants. *J. Nutr.* **137**, 2641–6.

23. Finkelstein JD, Harris BJ & Kyle WE (1972) Methionine metabolism in mammals: kinetic study of betaine-homocysteine methyltransferase. *Arch. Biochem. Biophys.* **153**, 320–4.

24. Stipanuk MH (2004) Sulfur amino acid metabolism: pathways for production and removal of homocysteine and cysteine. *Annu. Rev. Nutr.* **24**, 539–77.

25. Obeid R (2013) The metabolic burden of methyl donor deficiency with focus on the betaine homocysteine methyltransferase pathway. *Nutrients* **5**, 3481–95.

ANNEX 4.4: Annotated 1-carbon pathway diagrams

Chapter 3 provides a basic insight into the complexities of one-carbon pathways, but the summary remains extremely simplified. One-carbon metabolism is governed by intricately controlled feedback loops, which help protect the flux of metabolites through key reactions over a range of nutrient and co-factor concentrations[1,2]. An example of these complex allosteric regulations come from considering what happens when levels of SAM increase, for example through increased methionine input. Elevated SAM a) increases SAH, which in turn inhibits methyl transferases, b) increases CBS activity (aiding removal of Hcy) and c) inhibits MTHFR and BHMT (decreasing the conversion of Hcy to methionine)[3]. Since methyl-THF binds to (and deactivates) GNMT the reduced levels of methyl-THF brought about by MTHFR inhibition in c) help speed up GNMT activity, decreasing levels of SAM and thus helping achieve SAM homeostasis[1]. Even with these mechanisms helping to buffer impacts on SAM levels, it should be noted that different methyltransferases (MTs) have different K_m values for SAM (K_m: Michaelis constant; the concentration of substrate (SAM) required at half the maximum MT reaction velocity). Since DNMTs have low K_m values for SAM their velocity reaches maximum at low SAM concentrations. Therefore higher SAM concentrations do not have as large as impact on DNMTs as for MTs with higher K_m values[1]. These few examples help illustrate why varying nutrient inputs may not impact the one-carbon metabolome in ways that might seem intuitive, and a mathematical modelling / systems biology approach may help to disentangle the complexities of human metabolism.

The following diagrams show how the interlinking metabolic paths of:

a) Transmethylation
b) Methionine and homocysteine metabolism
c) Folate metabolism
d) Transsulfation pathway

The subsequent diagrams then show where additional metabolites mentioned in Chapter 4 (amino acids, formate, decarboxylated SAM) fit within one-carbon pathways:

e) Decarboxylated SAM and the polyamine pathway
f) The role of formate and its relationship with amino acids

References

1. Reed MC, Gamble M V, Hall MN, et al. (2015) Mathematical analysis of the regulation of competing methyltransferases. *BMC Syst. Biol.* **9**, 69.
2. Nijhout HF, Best J & Reed MC (2014) Escape from homeostasis. *Math. Biosci.* **257**, 104–10.
3. Nijhout HF, Reed MC, Anderson DF, et al. (2006) Long-range allosteric interactions between the folate and methionine cycles stabilize DNA methylation reaction rate. *Epigenetics* **1**, 81–7.

Figure A: Transmethylation

Transmethylation

DNA methyltransferases

DNMT

Cytosine

SAM CH₃ → SAH

5' Methyl-cytosine

SAM:SAH ratio is often used as a proxy of methylation potential[1]

S-adenosyl methionine (SAM)

MTs
DNMT

Acceptor

Glycine

GNMT

Sarcosine

S-adenosyl homocysteine (SAH)

Acceptor – CH₃

The acceptors are C, N, O, S atoms of proteins, nucleic acids, lipids & small molecules. Methylation is used in transcription, translation, protein localization & signalling. *i.e.* SAM is the major source of substrate for post-translational modifications (histone, DNA & RNA methylation)[2,3].

Abbreviations: DNMT; DNA methyltransferase; GNMT, Glycine N-Methyltransferase; MT, Methyl Transferases

References: 1. Mason JB. J Nutr. 2003; 133:941S–947. 2. Lu SC, Mato JM. Physiol Rev. 2012;92:1515–42. 3. Škovierová H, *et al*. Int J Mol Sci. 2016;17.

Figure B: Methionine and homocysteine metabolism

Methionine & homocysteine metabolism

THF

Methylcobalamin is oxidised to a cobalamin cofactor, which renders MTR ineffective. It has to then be restored by another enzyme (MTRR)[1]

MAT

ATP

PPi+ Pi

Methionine

Dimethylglycine (DMG)

Vit B12

N^5-methyl-THF

BHMT

MTR

Increasing SAM inhibits BHMT and DMG[2]

S-adenosyl methionine

Betaine

Donation of methyl group to S atom of Hcy from betaine or methyl-THF

Homocysteine

Choline

Acceptor

MTs
DNMT

Glycine

GNMT

Main pathway from Hcy to methionine is the vitamin B12-dependent one . The betaine pathway is mostly in the liver and kidneys[3].

Sarcosine

AHCY

Acceptor – CH$_3$

S-adenosyl homocysteine

Under equilibrium SAH is favoured over Hcy[4]

Increasing SAH inhibits most MTs[5]

References: 1. Vaughn JD, et al. J Nutr. 2004;134:2985–90. **2.** Nijhout HF, et al. Epigenetics. 2006;1:81–7. **3.** Selhub J. Annu Rev Nutr. 1999;19:217–46. **4.** Finkelstein JD. J Nutr Biochem. 1990; 1:228–237. **5.** Kerr SJ. J Biol Chem. 1972; 247(13):4248-52.

Abbreviations: AHCY, adenosylhomocysteinase; BHMT, Betaine Homocysteine Methyltransferase; DNMT; DNA methyltransferase; GNMT, Glycine N-Methyltransferase; MAT, Methionine Adenosyltransferase; MT, Methyl Transferases; MTR, Methionine Synthase; MTRR, Methionine Synthase Reductase; THF, tetrahydrofolate

320

Figure C: Folate metabolism

Folate metabolism

Abbreviations: DHFR, Dihydrofolate Reductase; GNMT, Glycine N-Methyltransferase; MTHFR, Methylenetetrahydrofolate Reductase; SHMT, Serine Hydroxymethyltransferase;

Supplement form

Folic acid

DHFR

Dietary folates mostly enter here

Dihydrofolate (DHF)

DHFR

Tetrahydrofolate (THF)

Scaffold which activates & carries 1-carbon units[1]

Interconversion to THF via methenyl, formyl and formate forms

serine

glycine

Vit B6

SHMT

Methionine

ATP

PPi+ Pi

Dimethylglycine (DMG)

Vit B12

MTHFR

$N^{5,10}$-methylene-THF

Irreversible step

Vit B2 as FAD

S-adenosyl methionine (SAM)

Betaine

N^5-methyl-THF

'methyl trap' – can accumulate at expense of other THF forms[2]

Choline

Homocysteine

Acceptor

Glycine

GNMT

Sarcosine

Acceptor – CH$_3$

S-adenosyl homocysteine (SAH)

High levels of SAM inhibit MTHFR and therefore production of methyl-THF. In turn lower methyl-THF reduces the inhibition of GNMT, so that excess SAM can be metabolised through the GNMT reaction[3].

References: 1. Ragsdale SW. Vitam Horm. 2008;79:293–324. **2.** Finkelstein JL, et al. Adv Nutr. 2015;6(5):552–63. **3.** Beagle B, et al. J Nutr. 2005; 135:2780–5.

Figure D: Transsulfuration

Transsulfuration pathway

Abbreviations: CBS, Cystathionine-Beta-Synthase; CTH, Cystathionine Gamma-Lyase; Hcy, homocysteine

Tetrahydrofolate

serine

glycine

Vit B6

$N^{5,10}$-methylene-THF

Methionine

ATP

PPi+ Pi

Dimethylglycine

Vit B12

Vit B2 as FAD

N^5-methyl-THF

S-adenosyl methionine

Betaine

Homocysteine

Choline

serine

H_2O

Vit B6

CBS

Acceptor

Glycine

Sarcosine

A powerful tripeptide antioxidant. Used to scavenge & reduce reactive oxygen species (ROS) and maintain NADPH / NADP+ ratio[1]

Acceptor – CH$_3$

S-adenosyl homocysteine

Cystathionine

Glutathione

α-ketobutyrate

Vit B6

Carbon skeleton from Hcy

CTH

Ammonium

Amino group from Hcy

Cysteine

Y-Glu-cysteine

Contains the S atom from Hcy
Cysteine enters the sulphide, pyruvate and taurine pathways[2]

References: 1. Wu G. Amino Acids. 2009;37:1–17.
2. Stipanuk MH. Annu Rev Nutr. 2004;24:539–77.

Figure E: Polyamine pathway

The polyamine pathway

References: **1.** Scotti M, et al. Rev Syst Biol Med. 2013;5:343–65.
2. Heby O. Int J Dev Biol. 1995;39:737–57.

Figure F: Amino acids and formate

Role of amino acids:

- **Methyl donors and acceptors through serine / glycine[1]**
- **Production of formate, a key 1C donor[2,3]**

Folic acid

Dihydrofolate (DHF)

Formate

Used to make **purines** (adenine & guanine)

Tetrahydrofolate (THF)

Via Thymidylate synthase pathway

N^{10} formyl-THF

serine

glycine

Vit B6

$N^{5,10}$methenyl-THF

$N^{5,10}$-methylene-THF

Vit B2 as FAD

Via remethylation of Hcy to methionine

N^5-methyl-THF

S-adenosyl methionine (SAM)

Homocysteine (Hcy)

serine

H_2O

Formate is produced in mitochondria then released into cytosol - > 10-formyl-THF by condensation with THF -> purine synthesis

Acceptor

Glycine

Sarcosine

Vit B6

S-adenosyl homocysteine (SAH)

Cystathionine

Acceptor – CH_3

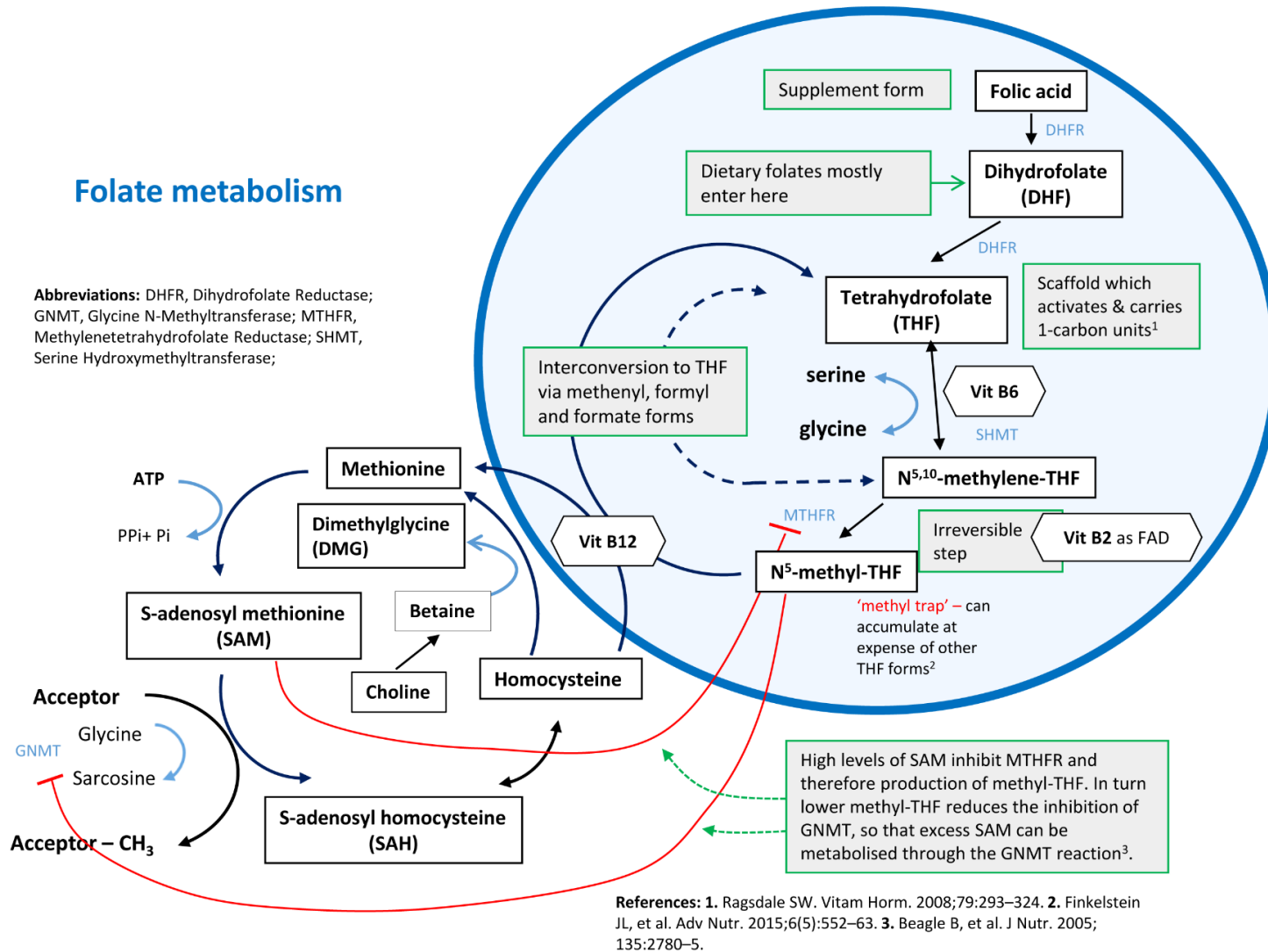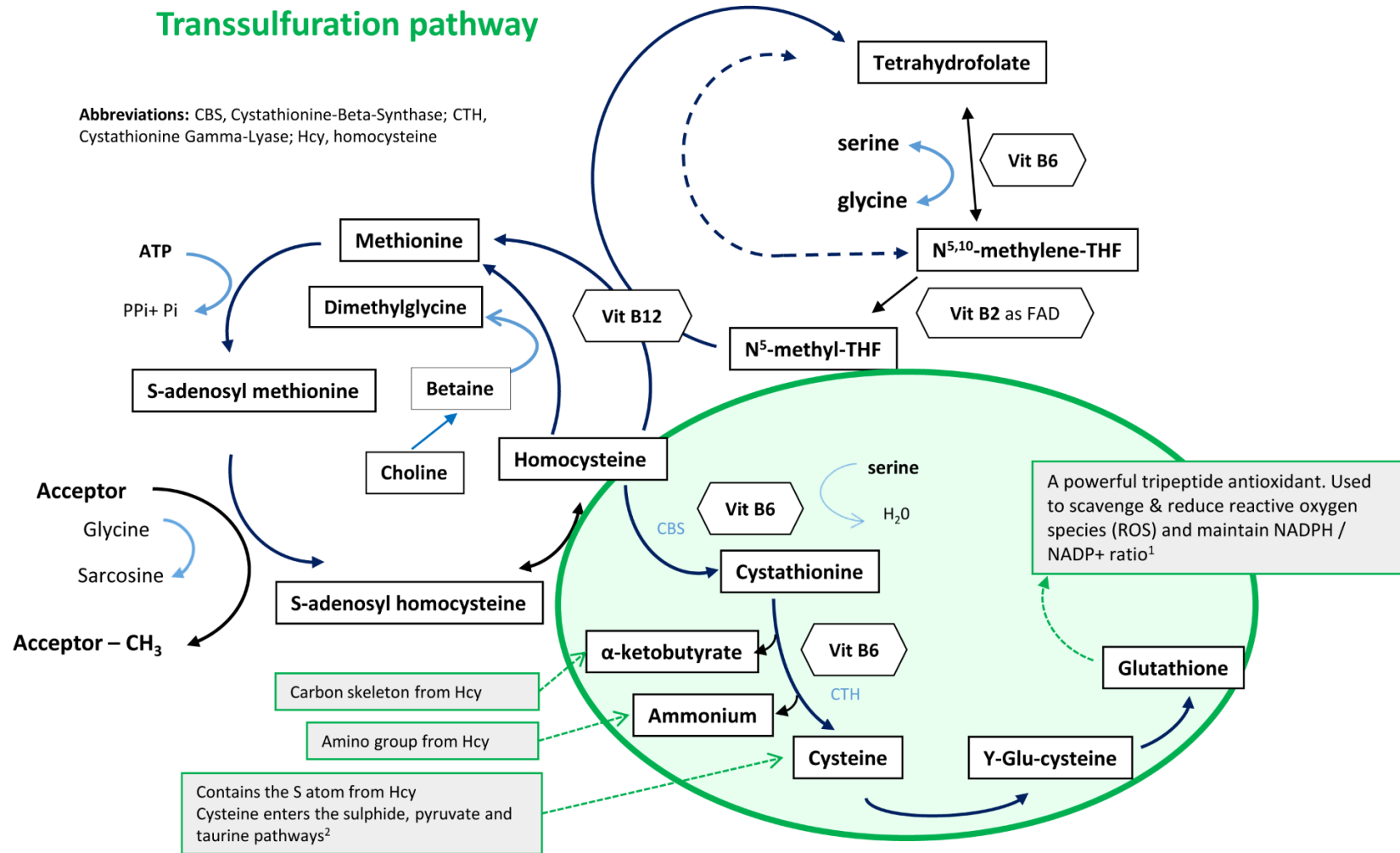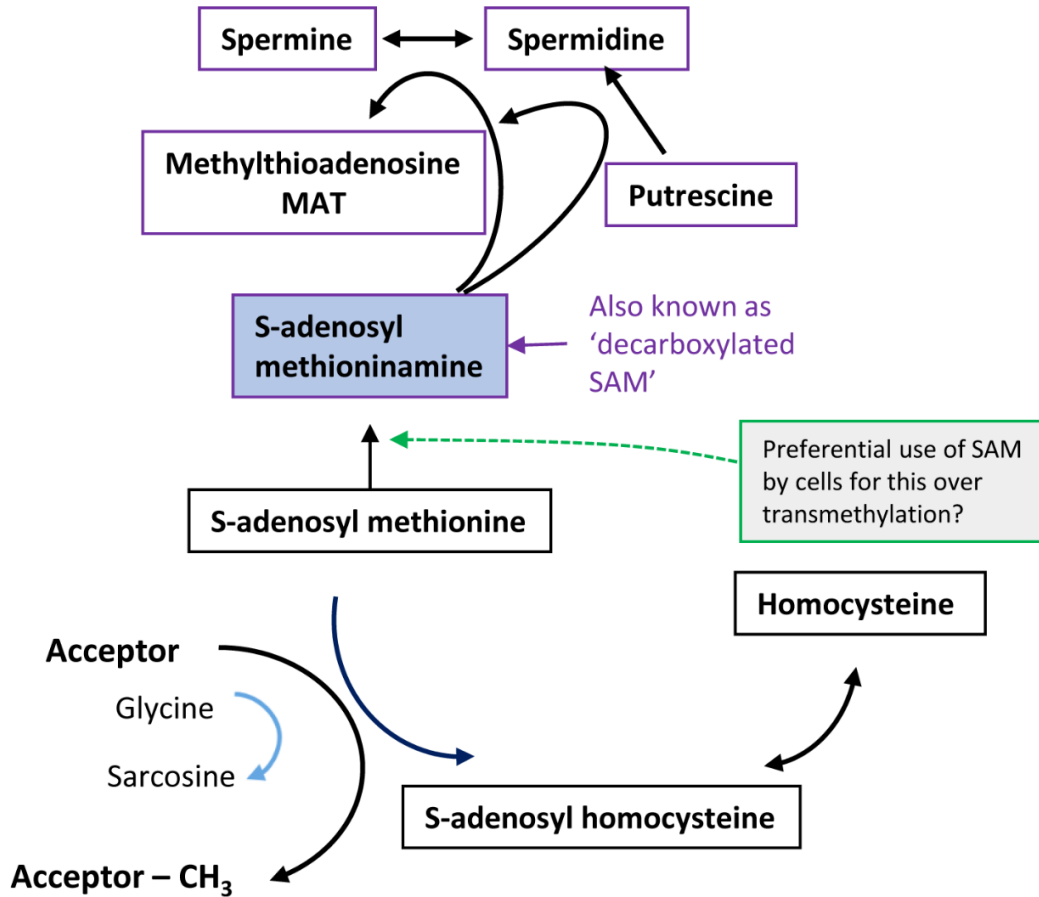α-ketobutyrate

Vit B6

Glutathione

Ammonium

Cysteine

Y-Glu-cysteine

•Formate is produced in cytoplasm by methanol oxidation, cholesterol synthesis, phytanic acid oxidation, catabolism of **serine, tryptophan** & **histidine**[2,3].
•Formate is the **primary source of one-carbons** for nuclear and cytoplasmic one carbon metabolism[2,3]

References: **1.** Davis SR, et al. Am J Physiol Endocrinol Metab. 2004 Feb;286(2):E272-9. **2.** Stover PJ, Field MS. Adv Nutr. 2011;2:325–31. **3.** Lamarre SG, et al. Clin Chem Lab Med. 2013;51:571–8.

324

ANNEX 7.1 Selection of B vitamin and folic acid trials detailing dosage and side effect information

Study	Design	Population	Micronutrients Provided	Outcome investigated	Results	Side effects
Bhutta Z, et al. (2009)[1]	RCT, daily supplementation in pregnancy. Fe-Fol vs. MMN	2,378 pregnant women, Pakistan	Fe, **folic acid (400µg)**, vitamins A, D, E, C, B1, niacin, **B2 (1.4mg)**, B6, **B12 (2.6µg)**, Zn, Cu, Se, I (UNIMMAP)	Birthweight	Increase of 70g birthweight in MMN arm.	Comparable rates of self-reported morbidities between arms. Minimal reported gastrointestinal side effects.
Brough L, et al. (2010)[2]	RCT, daily supplementation in pregnancy. MMN vs. placebo	402 pregnant women, UK	β-carotene, B1, **B2 (2mg)**, B6, **B12 (6µg)**, **folic acid (400µg)**, C, D, E, K, Fe, Zn, Mg, I, Cu	Maternal nutrient status, birth size	↑ markers of Fe, folate, B1 and vitamin D in MMN arm in third trimester. In compliant subset MMN arm had ↓SGA infants (borderline significance)	No difference in reported side effects.
Christian P, et al. (2003)[3]	Cluster RCT, daily supplementation in pregnancy. Vit A (control) vs. Vit A + FA vs. Vit A + FA + Fe vs. Vit A + FA + Fe + Zn vs. Vit A + Fe + MMN	4926 pregnant women, Nepal	Vitamin A, **folic acid (400µg)**, Fe, Zn, D, E, B1, **B2 (1.8mg)**, B6, **B12 (2.6µg)**, C, K, niacin, Cu, and Mg	Foetal loss and infant mortality	No effect of any arm on reduction of overall foetal loss. In stratified analyses the Vit A+FA arm and Vit A + FA + Fe arms ↓ 3-month mortality amongst pre-terms compared to control. The MMN arm showed ↑ risk of mortality in term infants (borderline significance).	No maternal side effects reported here.
Dawson S, et al. (2016)[4]	Meta-analysis of 21 studies supplementing with either Omega-3 or Omega-3 + FA + B6 + B12 in daily supplementation	Range of populations	**Folic acid (150 – 2500µg)**, omega-3, B6, **B12 (1.9 µg – 12mg)**	Plasma homocysteine	Trials combining omega-3 with FA and B vitamins reduced plasma Hcy the most.	Not reported.
Ebbing M, et al. (2009)[5]	RCT, daily supplementation for median 39 months. FA+B12+B6 vs.	6,837 participants with ischaemic heart	**Folic acid (800µg), B12 (400µg)**, B6	Cancer incidence, mortality	Increased lung cancer risk and cancer mortality in FA+B12 arm vs. control. No increased risks in B6 arms.	No minor side effects reported but see results section for possible serious adverse events.

Study	Design	Population	Micronutrients Provided	Outcome investigated	Results	Side effects
	FA+B12 vs. B6 vs. placebo	disease, Norway				
Fawzi W, et al. (2007)[6]	RCT, daily supplementation in pregnancy to 6 weeks post-partum. MMN + Fe-Fol vs. placebo + Fe-Fol	8,468 HIV-negative pregnant women, Tanzania	B1, **B2 (20mg)**, B6, niacin, **B12 (50 µg)**, C, E, **folic acid (800µg)**	Low birth weight, prematurity, foetal death	15% lower risk of LBW and 67g higher birthweight in MMN arm. No effect on preterm birth or foetal death.	No difference in maternal mortality or miscarriages by arm.
Friis H, et al. (2004)[7]	RCT, daily supplementation in pregnancy until deliver. Fe-Fol + MMN vs. Fe-Fol	1,106 pregnant women, Zimbabwe.	Folic acid, Fe, vitamins A, D, E, C, B1, niacin, **B2 (1.6mg),** B6, **B12 (4.0µg),** Zn, Cu, Se	Birth outcomes	46g increase in the MMN + Fe-Fol arm but borderline significance.	Maternal side effects not reported.
Galan P, et al. (2010)[8]	RCT, daily supplementation for median 4.7 years. Omega 3 vs. omega 3 + B vits vs. B vits vs. placebo	2,501 patients with a history of myocardial infarction, unstable angina, or ischaemic stroke, France	5-methyltetrahydrofolate, B6, **B12 (20 µg)**, omega-3 fatty acids	Prevention of cardiovascular events	B vitamins arm ↓ plasma homocysteine concentrations by 19% compared with placebo, but had no effect on major vascular events. No effect of omega-3 on vascular events compared to placebo.	2.1% of people experienced side effects (mainly gastrointestinal and nausea) in the B vitamins group, compared with 2.6% in the omega-3 fatty acids group, and 1.6% in the placebo group.
Hall M, et al. (2016)[9]	RCT, daily supplementation for 12 weeks. FA (800µg) vs. FA (400µg) vs. creatine vs. creatine + FA 400µg vs. placebo	622 arsenic-exposed adults, Bangladesh	**Folic acid (800µg, 400µg),** creatine	Plasma choline, betaine and dimethylglycine (DMG) concentrations	FA groups ↑plasma betaine, ↓DMG and prevented a ↓ in choline. No difference when adding creatine.	Not reported.

Study	Design	Population	Micronutrients Provided	Outcome investigated	Results	Side effects
Hininger I, et al. (2004)[10]	RCT, daily supplementation for two months, MMN vs. placebo	100 healthy pregnant women, France	**Folic acid (200µg),** vitamin C, E, B1, **B2 (1.6mg),** B6, **B12 (1µg),** pantothenic acid, β-carotene, Zn, Mg, Ca	Plasma micronutrient levels, oxidative stress parameters	↑Folic acid, vitamin C, E, B2, B6 and β-carotene concentration in MMN arm. No difference in oxidative stress parameters.	Not reported.
Liu J, et al. (2013)[11]	RCT, daily supplementation in pregnancy, FA (control) vs. FA + Fe vs. FA + Fe + MMNs	18,775 pregnant women, China	Fe, **folic acid (400µg),** vitamins A, D, E, C, B1, niacin, **B2 (1.4mg),** B6, **B12 (2.6µg),** Zn, Cu, Se, I (UNIMMAP)	Perinatal mortality, neonatal mortality, birth size, maternal nutrient status	No effect of supplements on perinatal mortality. Supplement arms ↓ maternal anaemia in third trimester (~28% risk reduction).	No serious adverse effects reported. 6% reported gastrointestinal discomfort in MMN group vs. only 2.2% in FA controls.
Kaestel P, et al. (2005)[12]	RCT, daily supplementation in pregnancy, Fe-Fol (control), vs. 15 micronutrients at 1xRDA (MMN1) vs. 15 micronutrients at 2xRDA (MMN2)	2,100 pregnant women, Guinea Bissau	Fe, **folic acid (400µg, 800µg),** vitamins A, D, E, C, B1, niacin, **B2 (1.4mg, 2.8mg),** B6, **B12 (2.6µg, 5.2µg),** Zn, Cu, Se, I (UNIMMAP)	Birthweight and birth outcomes	MMN1 and MMN2 ↑ birthweight compared to control, 53g and 95g respectively. No difference in proportion of LBW. Larger effect with MMN2 in anaemic women (↑ birthweight by 218g and ↓LBW risk by 31%).	Lower risk of miscarriage in MMN2 group (but not MMN1 group) compared to control. No difference in maternal mortality between arms.
Kirke P, et al. (1992)[13]	RCT, daily supplementation pre-conception for at least two months. FA vs. MMN vs. FA+MMN	354 women, Ireland	FA: **Folic acid (0.36mg),** MMN: Vitamins A, D, B1, **B2 (1.5mg),** B6, niacin, C, Ca, Fe	Neural tube defects	No statistical difference in prevention of NTDs per arm.	No statistically significant differences in pregnancy outcome information by trial arm.
Kumwenda N, et al. (2002)[14]	RCT, daily supplementation in pregnancy. Fe-Fol vs. Fe-Fol + Vit A	697 HIV-positive pregnant women, Malawi	**Folic acid (400µg),** Fe, Vit A	Birthweight	Fe-Fol plus Vit A arm ↑birthweight (90g).	Not reported.

Study	Design	Population	Micronutrients Provided	Outcome investigated	Results	Side effects
Melse-Boonstra A, et al. (2005)[15]	RCT, daily supplementation. 6 incremental doses of FA vs. placebo	308 adults (50-75y), The Netherlands	**Folic acid (50-800µg),**	Plasma betaine and homocysteine (Hcy)	Plasma concentration of Hcy ↓ as ↑folate and ↑ betaine. Plasma betaine ↑ as FA dose ↑.	Not reported.
Moore SE, et al. (2012)[16]	RCT, daily supplementation in pregnancy. Protein-energy (PE) + MMN, PE, MMN, Fe-Fol	620 pregnant women, The Gambia	Fe, **folic acid (400µg),** vitamins A, D, E, C, B1, niacin, **B2 (2.8mg),** B6, **B12 (5.2µg),** Zn, Cu, Se, I, protein, lipids	Thymic index	Not yet reported for primary outcome. Secondary outcome analysis suggests no influence of supplementation on foetal growth[17].	Not yet reported
Wald N, et al. (1991)[18]	RCT, daily supplementation pre-conception until week 12 of pregnancy. FA vs. FA + MMN vs. MMN vs. placebo	1,817 women in 7 countries	**Folic acid (4mg),** MMN: Vit A, D, B₁, **B2 (1.5mg),** B6, C, nicotinamide	Neural tube defects	FA groups had ↓ risk of NTDs. No protection from MMNs without FA.	No difference in self-reported side effects between arms.
Osrin D, et al. (2005)[19]	RCT, daily supplementation in pregnancy. Fe-Fol (control) vs. MMN	1,200 pregnant women, Nepal	Fe, **folic acid (400µg),** vitamins A, D, E, C, B1, niacin, **B2 (1.4mg),** B6, **B12 (2.6µg),** Zn, Cu, Se, I (UNIMMAP)	Birthweight, gestational length	MMN arm ↑ birthweight by 77g and ↓ LBW by 25%.	Fewer miscarriages in the MMN arm (5 vs. 2). No differences in gastrointestinal complaints.
Ramakrishnan U, et al. (2003)[20]	Double-blind RCT, 6 days/week for at least two months, MMN vs. Fe-only	873 pregnant women, Mexico	Vitamin A, B1, D, E, **B2 (1.87mg),** niacin, **folic acid (215µg),** B6, **B12 (2.04µg),** C, Zn, Fe, Mg	Birth size	No effect of multiple micronutrients versus iron alone	None reported.
Ramakrishnan U, et al. (2016)[21]	RCT, daily supplementation pre-conception to conception (mean 33 weeks). FA vs. FA + Fe vs. FA + FE + MMNs	1,813 women, Vietnam	**Folic acid (2.8mg),** Fe, A, D, E, C, B1, niacin, **B2 (1.4mg),** B6, **B12 (2.6µg),** Zn, Cu, Se, I	Birth outcomes	No difference in birthweight, gestational age, preterm delivery between arms.	No difference in reported side effects by arm.

Study	Design	Population	Micronutrients Provided	Outcome investigated	Results	Side effects
Roberfroid D, et al. (2008)[22]	RCT, daily supplementation in pregnancy. Fe-Fol (control) vs. MMN	1,426 pregnant women, Burkina Faso	Fe, **folic acid (400μg)**, vitamins A, D, E, C, B1, niacin, **B2 (1.4mg)**, B6, **B12 (2.6μg)**, Zn, Cu, Se, I (UNIMMAP)	Fetal growth	MMN arm ↑ birthweight by 52g, no difference in risk of LBW	No difference in miscarriage between arms. No other maternal side effects reported.
Rumiris D, et al. (2006)[23]	RCT, daily supplementation in pregnancy. MMN vs. Fe-Fol (control)	Pregnant women with low antioxidant status	Vitamins A, B6, **B12 (2.2 μg)**, C, E, **folic acid (400 μg)**, N-acetylcysteine, Cu, Zn, Mn, Fe, Ca, Se	Perinatal outcomes	↓Preeclampsia in the MMN group compared to control	Not reported.
Rush D, et al. (1980)[24]	RCT, daily supplementation in pregnancy, protein-energy (PE) + MMN vs. MMN vs. standard care.	1,051 pregnant women, USA	Ca, Mg, Fe, Zn, Cu, I, Vitamins A, D, E, C, B1, **B2 (15mg)**, niacin, B6, pantothenic acid, biotin, **folic acid (350μg), B12 (8μg)**, protein, carbohydrate, fat	Birthweight, post-natal development	PE+MMN prevented low birthweight in mothers who smoked heavily.	Higher neonatal deaths in supplement group although not statistically significant (and note this was in the arm containing additional protein-energy).
Shankar AH, et al. (2008)[25]	RCT, daily supplementation in pregnancy, Fe-Fol vs. MMN	31,290 pregnant women, Indonesia	Fe, **folic acid (400μg)**, vitamins A, D, E, C, B1, niacin, **B2 (1.4mg)**, B6, **B12 (2.6μg)**, Zn, Cu, Se, I (UNIMMAP)	Infant mortality, foetal loss, birthweight	↓Early infant mortality (18% reduction) in the MMN group compared to control. Combined foetal loss and neonatal mortality ↓ by 11% in the MMN arm. LBW ↓ by 14% in MMN arm.	Not reported.
Tofail F, et al. (2008)[26]	RCT, daily supplementation week 14 gestation to delivery. MMN (and early food) vs. Fe (30mg)-fol (and early food) vs. Fe (60mg)-fol (and early food) vs.	2,853 pregnant mothers, Bangladesh	Fe, **folic acid (400μg)**, vitamins A, D, E, C, B1, niacin, **B2 (1.4mg)**, B6, **B12 (2.6μg)**, Zn, Cu, Se, I (UNIMMAP)	Infant development	No intervention effects on whole group, but varying effects of the 6 different groups for low BMI mothers.	Not reported.

Study	Design	Population	Micronutrients Provided	Outcome investigated	Results	Side effects
	MMN (normal food) vs. Fe (30mg)-fol (normal food) vs. Fe (60mg)-fol (normal food)					
Toole J, et al. (2004)[27]	RCT, daily supplementation for mean 20 months, high dose vs. low dose B6, B12, FA.	3,680 adults with non-disabling cerebral infarction, USA, Canada Scotland	High dose: B6, **B12 (0.4mg), folic acid (2.5mg)** Low dose: B6, **B12 (6µg), folic acid (20 µg)**	Plasma Homocysteine, stroke	Plasma Hcy ↓ by 2µmol/L in high dose group compared to low dose group. No effect on stroke or death.	No significant differences between groups for itching, skin rash, or gastrointestinal complaints.
West K, et al. (2014)[28]	RCT, daily supplementation in pregnancy to 12 weeks postpartum, Fe-Fol vs. MMN.	44,567 pregnant women, Bangladesh	Fe, **folic acid (400µg),** vitamins A, D, E, C, B1, niacin, **B2 (1.4mg),** B6, **B12 (2.6µg),** Zn, Cu, Se, I (UNIMMAP)	Birth outcomes, 6-month infant mortality	No effect on 6-month infant mortality. MMN arm ↓ pre-term birth by 15% and ↓LBW by 12%.	No differences in morbidities between arms.
Zagré N, et al. (2007)[29]	RCT, daily supplementation in pregnancy, Fe-Fol vs. MMN.	3,670 pregnant women, Niger	Fe, **folic acid (400µg),** vitamins A, D, E, C, B1, niacin, **B2 (1.4mg),** B6, **B12 (2.6µg),** Zn, Cu, Se, I (UNIMMAP)	Birthweight	Birthweight 67g higher in MMN arm compared to Fe-Fol arm.	Miscarriages and maternal deaths were comparable in both groups
Zeng, L et al. (2008)[30]	Cluster RCT, daily supplementation in pregnancy. FA (control) vs. Fe-Fol vs. MMN	5,828 pregnant women in China	Fe, folic acid **(400 µg)**, Zn, Cu, Se, I, Vitamins A, B1, **B2 (1.4 mg), B12 (2.6µg)**, D, C, E, niacin.	Birth size	Birthweight 42g higher in MMN group compared to FA group. Fe-Fol reduced risk of pre-term delivery.	No differences in perinatal mortality between groups.

Abbreviations: BMI, body mass index; FA, folic acid; Fe-Fol, iron-folic acid supplement; Hcy, homocysteine; LBW, low birthweight; MMN, multiple micronutrient supplement; NTD, neural tube defect; RCT, randomised controlled trial; RDA, recommended daily allowance; UNIMMAP, United Nations International Multiple Micronutrient Preparation

References

1.   Bhutta ZA, Rizvi A, Raza F, et al. (2009) A comparative evaluation of multiple micronutrient and iron-folic acid supplementation during pregnancy in Pakistan: impact on pregnancy outcomes. *Food Nutr. Bull.* **30**, S496-505.

2.   Brough L, Rees GA, Crawford MA, et al. (2010) Effect of multiple-micronutrient supplementation on maternal nutrient status, infant birth weight and gestational age at birth in a low-income, multi-ethnic population. *Br. J. Nutr.* **104**, 437–45.

3.   Christian P, West KP, Khatry SK, et al. (2003) Effects of maternal micronutrient supplementation on fetal loss and infant mortality: a cluster-randomized trial in Nepal. *Am. J. Clin. Nutr.* **78**, 1194–202.

4.   Dawson SL, Bowe SJ & Crowe TC (2016) A combination of omega-3 fatty acids, folic acid and B-group vitamins is superior at lowering homocysteine than omega-3 alone: A meta-analysis. *Nutr. Res.* **36**, 499–508.

5.   Ebbing M, Bønaa KH, Nygård O, et al. (2009) Cancer incidence and mortality after treatment with folic acid and vitamin B12. *JAMA* **302**, 2119–26.

6.   Fawzi WW, Msamanga GI, Urassa W, et al. (2007) Vitamins and perinatal outcomes among HIV-negative women in Tanzania. *N. Engl. J. Med.* **356**, 1423–31.

7.   Friis H, Gomo E, Nyazema N, et al. (2004) Effect of multimicronutrient supplementation on gestational length and birth size: a randomized, placebo-controlled, double-blind effectiveness trial in Zimbabwe. *Am. J. Clin. Nutr.* **80**, 178–84.

8.   Galan P, Kesse-Guyot E, Czernichow S, et al. (2010) Effects of B vitamins and omega 3 fatty acids on cardiovascular diseases: a randomised placebo controlled trial. *BMJ* **341**, c6273.

9.   Hall MN, Howe CG, Liu X, et al. (2016) Supplementation with Folic Acid, but Not Creatine, Increases Plasma Betaine, Decreases Plasma Dimethylglycine, and Prevents a Decrease in Plasma Choline in Arsenic-Exposed Bangladeshi Adults. *J. Nutr.* **146**, 1062–7.

10.   Hininger I, Favier M, Arnaud J, et al. (2004) Effects of a combined micronutrient supplementation on maternal biological status and newborn anthropometrics measurements: a randomized double-blind, placebo-controlled trial in apparently healthy pregnant women. *Eur. J. Clin. Nutr.* **58**, 52–9.

11.   Liu J, Mei Z, Ye R, et al. (2013) Micronutrient Supplementation and Pregnancy Outcomes. *JAMA Intern. Med.* **173**, 276.

12.   Kaestel P, Michaelsen KF, Aaby P, et al. (2005) Effects of prenatal multimicronutrient supplements on birth weight and perinatal mortality: a randomised, controlled trial in Guinea-Bissau. *Eur. J. Clin. Nutr.* **59**, 1081–9.

13.   Kirke PN, Daly LE & Elwood JH (1992) A randomised trial of low dose folic acid to

prevent neural tube defects. The Irish Vitamin Study Group. *Arch. Dis. Child.* **67**, 1442–6.

14.   Kumwenda N, Miotti PG, Taha TE, et al. (2002) Antenatal vitamin A supplementation increases birth weight and decreases anemia among infants born to human immunodeficiency virus-infected women in Malawi. *Clin. Infect. Dis.* **35**, 618–24.

15.   Melse-Boonstra A, Holm PI, Ueland PM, et al. (2005) Betaine concentration as a determinant of fasting total homocysteine concentrations and the effect of folic acid supplementation on betaine concentrations. *Am. J. Clin. Nutr.* **81**, 1378–82.

16.   Moore SE, Fulford AJ, Darboe MK, et al. (2012) A randomized trial to investigate the effects of pre-natal and infant nutritional supplementation on infant immune development in rural Gambia: the ENID trial: Early Nutrition and Immune Development. *BMC Pregnancy Childbirth* **12**, 107.

17.   Johnson W, Darboe MK, Sosseh F, et al. (2017) Association of prenatal lipid-based nutritional supplementation with fetal growth in rural Gambia. *Matern. Child Nutr.* **13:** e 12367.

18.   MRC Vitamin Study Research Group. (1991) Prevention of neural tube defects: results of the Medical Research Council Vitamin Study. *Lancet* **338**, 131–7.

19.   Osrin D, Vaidya A, Shrestha Y, et al. (2005) Effects of antenatal multiple micronutrient supplementation on birthweight and gestational duration in Nepal: double-blind, randomised controlled trial. *Lancet* **365**, 955–62.

20.   Ramakrishnan U, González-Cossío T, Neufeld LM, et al. (2003) Multiple micronutrient supplementation during pregnancy does not lead to greater infant birth size than does iron-only supplementation: a randomized controlled trial in a semirural community in Mexico. *Am. J. Clin. Nutr.* **77**, 720–5.

21.   Ramakrishnan U, Nguyen PH, Gonzalez-Casanova I, et al. (2016) Neither Preconceptional Weekly Multiple Micronutrient nor Iron-Folic Acid Supplements Affect Birth Size and Gestational Age Compared with a Folic Acid Supplement Alone in Rural Vietnamese Women: A Randomized Controlled Trial. *J. Nutr.* **146,**1445S-52S.

22.   Roberfroid D, Huybregts L, Lanou H, et al. (2008) Effects of maternal multiple micronutrient supplementation on fetal growth: a double-blind randomized controlled trial in rural Burkina Faso. *Am. J. Clin. Nutr.* **88**, 1330–40.

23.   Rumiris D, Purwosunu Y, Wibowo N, et al. (2006) Lower rate of preeclampsia after antioxidant supplementation in pregnant women with low antioxidant status. *Hypertens. pregnancy* **25**, 241–53.

24.   Rush D, Stein Z & Susser M (1980) A randomized controlled trial of prenatal nutritional supplementation in New York City. *Pediatrics* **65**, 683–97.

25.   Supplementation with Multiple Micronutrients Intervention Trial (SUMMIT) Study Group, Shankar AH, Jahari AB, et al. (2008) Effect of maternal multiple micronutrient supplementation on fetal loss and infant death in Indonesia: a double-blind cluster-randomised trial. *Lancet* **371**, 215–27.

26. Tofail F, Persson LA, El Arifeen S, et al. (2008) Effects of prenatal food and micronutrient supplementation on infant development: a randomized trial from the Maternal and Infant Nutrition Interventions, Matlab (MINIMat) study. *Am. J. Clin. Nutr.* **87**, 704–11.

27. Toole JF, Malinow MR, Chambless LE, et al. (2004) Lowering Homocysteine in Patients With Ischemic Stroke to Prevent Recurrent Stroke, Myocardial Infarction, and Death. *JAMA* **291**, 565.

28. West KP, Shamim AA, Mehra S, et al. (2014) Effect of maternal multiple micronutrient vs iron-folic acid supplementation on infant mortality and adverse birth outcomes in rural Bangladesh: the JiVitA-3 randomized trial. *JAMA* **312**, 2649–58.

29. Zagré NM, Desplats G, Adou P, et al. (2007) Prenatal multiple micronutrient supplementation has greater impact on birthweight than supplementation with iron and folic acid: a cluster-randomized, double-blind, controlled programmatic study in rural Niger. *Food Nutr. Bull.* **28**, 317–27.

30. Zeng L, Dibley MJ, Cheng Y, et al. (2008) Impact of micronutrient supplementation during pregnancy on birth weight, duration of gestation, and perinatal mortality in rural western China: double blind cluster randomised controlled trial. *BMJ* **337**, a2001.

ANNEX 9.1 Candidate exposures to consider in future epigenetic research in The Gambia

Microbiome

It is now well documented that the gut microflora influence a multitude of metabolic processes in the host[1], and that some of these processes are associated with a range of outcomes such as obesity, insulin resistance and glucose homeostasis[2,3]. There is evidence that diet can affect microbiota composition[4–6]. The intake of polysaccharides, oligosaccharides and other non-digestible carbohydrates are of particular interest, since these are the forms of carbohydrate that pass into the colon and can be fermented by the microbiota. In turn, the microbiota produce a number of metabolites that could influence one-carbon methyl donor concentrations[7] and possibly also histone acetylation processes[8,9]. The most documented metabolites include short-chain fatty acids, choline and bile acid derivatives[10]. However, some species of bacteria are also thought to be able to produce riboflavin (e.g. *L. lactis, L. fermentum, L. plantarum) and folate (multiple strains of Lactobacillales)[7]*. There are several reviews that draw together animal and *in vitro* experimental data to explore how the metabolites produced by the microbiome could influence epigenetic processes in the host[11,12]. However, there is a dearth of human *in vivo* evidence linking the microbiome with host epigenetic processes, although this research gap has been acknowledged and there are calls for more integrative research to address this[13]. For example, there is preliminary evidence from a pilot study in humans that gut microflora composition may be associated with DNA methylation[14]. However, this research has not been in the intergenerational setting, and further investigation is required as to whether the maternal microbiome can influence offspring epigenetics.

Other nutritional influences

Whilst a metabolomics approach outlined above would cover a wide breadth of metabolic networks, there are a two other specific nutrients of interest to consider: polyunsaturated fatty acids (PUFAs) and vitamin C. Several studies indicate that increased consumption of ω-3 PUFAs is associated with reduced homocysteine levels[15] and there is increasing interest as to whether epigenetic mechanisms are affected. In rodent models, providing dams a diet with excess folic acid and restricted vitamin B12 (reflecting the nutritional status of many rural Indian mothers) is associated with increased maternal oxidative stress, lower offspring

placental and brain docosahexaenoic acid (DHA) concentration, and decreased placental global methylation[16]. However, maternal dietary supplementation with PUFAs then partially ameliorates the disrupted one-carbon metabolism associated with this diet[17]. One potential mechanism is through PUFAs up-regulating enzymes responsible for the methylation of homocysteine to methionine[18], which in theory could increase the SAM:SAH ratio and help overcome the impact of reduced *MTHFR* activity brought about by the restricted vitamin B12. An additional hypothesis describes how a decreased maternal PUFA intake may increase the availability of methyl groups for DNA methylation. This is because there would be reduced PUFAs available for phosphatidylcholine (PC) synthesis through the phosphatidylethanolamine N-methyltransferase (PEMT) pathway, a process that requires three methyl groups from SAM for each molecule of PC formed. The disturbance to methyl balance brought about by low PUFA levels is therefore hypothesised to alter epigenetic programming of placental genes, increasing the risk of aberrant methylation in offspring and adverse pregnancy outcome[19]. Human evidence at the intergenerational level, however, is currently limited to a study by Lee *et al.* (2013). Here 261 mothers were given either 400mg of ω-3 PUFA or placebo daily from approximately 18 weeks gestation to birth. Amongst mothers supplemented with PUFAs there was increased global methylation (measured at LINE-1 repetitive elements) in offspring of mothers who smoked[20] and increased offspring *IGF2* promoter 3 methylation in pre-term infants[21].

Vitamin C may be influential during the process of active erasure of methylation marks (demethylation), predominantly in the paternal genome immediately after fertilization between the zygote and blastocyst stage[22]. In demethylation, 5-methylcytosine (5mC) can be sequentially oxidised to different states by ten-eleven translocation (TET) dioxygenases that use vitamin C (ascorbate) as a co-factor[23]. 5mC is first oxidised to 5-hydroxymethylcytosine (5hmC), and then in turn to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). Via a base excision repair mechanism 5fC or 5caC can be converted to an unmodified cytosine[23]. Recent studies, however, have indicated that there may be several different mechanisms for demethylation, and that TET-mediated pathways may only account for 25% of paternal genome demethylation[22]. Despite this, there is some preliminary evidence that vitamin C remains a potentially important contributing factor to consider in epigenetic processes during early embryogenesis. *In vitro* studies, for example, indicate that adding vitamin C to mouse or human embryonic stem cells increases activity of TET enzymes, with active demethylation seen in the germline and associated gene expression changes[24,25]. *In vivo* studies will be required to validate these observations.

Oxidative stress

Menezo *et al.* (2016)[26] make a strong case for considering the role of oxidative stress when considering exposures affecting DNA methylation. They summarise evidence suggesting that increased oxidative stress is associated with dysregulated DNA methylation. This may be linked to the one-carbon cycle given that two of the most powerful antioxidants are glutathione and hypotaurine, products of the transsulfuration pathway, both of which are abundant in the genital tract. The authors suggest that it is particularly at the stage of maintaining the parental imprints post-fertilization when oxidative stress could affect methylation. Potential mechanisms for the association between oxidative stress and methylation are still being investigated, but one hypothesis is that guanine is the nucleotide most susceptible to oxidation, and therefore if not repaired could affect the methylation that would occur at a CpG site. Free-radical mediated oxidation of cytosine also occurs, forming 5hmC, which as described above, can lead to demethylation. Together such alterations at the CpG level could affect the affinity of transcription factors with implications for chromatin organisation. Another potential mechanism the authors present suggests the presence of reactive oxygen species (ROS) may directly affect the expression of DNA methyltransferase.

The role of homocysteine in all this is complex. Homocysteine is required as a precursor along with serine to form cystathioinine, cysteine and then the anti-oxidant glutathionine in the transsulfuration pathway. However, excess homocysteine may increase oxidative stress, possibly through nitric oxide formation[27] or through several other routes to ROS production resulting from the oxidation of its thiol group[28,29]. Despite the fact the mechanisms remain to be elucidated fully, collecting information on glutathione and species of ROS may help address the question of whether some of the methylation disturbances are due to oxidative stress mechanisms, and the role that imbalances in one-carbon metabolites may play in dysregulating homocysteine recycling[30].

Inflammation and infection

Claycombe *et al.* (2015)[31] review some of the intricate ways that infection, inflammation, nutrition and epigenetics can be interlinked. They describe existing animal and *in vitro* evidence suggesting the presence of certain pathogens can influence T cell differentiation, macrophage phenotype and cytokine production through mechanisms thought to involve DNA methylation and histone modifications. They hypothesise that inflammation during pregnancy may alter epigenetic regulation of immune function, which may contribute

towards inflammatory components of chronic diseases such as cardiovascular disease, obesity and asthma. Nutrition can modulate these relationships due to direct effects on inflammation, for example, the pro-inflammatory effects of a high saturated fat intake and the potential anti-inflammatory effects of vitamin D and folic acid. The authors highlight the fact that human *in vivo* evidence for causal relationships between maternal infection, offspring epigenetic effects and adverse phenotypes is scarce.

Recent animal models, however, do suggest there is rationale for exploring these factors further. A mouse study subjected pups to bacterial infection *in utero* and found a number of methylation differences at genetic loci related to the toll-like receptor (TLR) signalling pathway, which correlated with gene expression[32]. The TLR pathway is involved in innate immunity, promoting a pro-inflammatory response which could in turn be associated with phenotypes such as pre-term birth[32]. There is also growing evidence to suggest that pups exposed to various viruses and bacteria *in utero* go on to develop neurodevelopmental disorders, potentially mediated by methylation at genes such as *MECP2*, reviewed in Weber-Stadlbauer (2017)[33].

Infection could also influence epigenetic mechanisms via oxidative stress (see above), since the presence of pathogens can trigger an inflammatory response in the host that affects redox balance, increasing oxidative stress[34,35].

Toxin exposure

There is a developing literature on maternal exposure to toxins and offspring DNA methylation effects, especially linked to arsenic exposure, tobacco smoke, air pollutants, phthalates and bisphenol A[36]. In The Gambia maternal plasma markers of aflatoxin exposure divided into high and low exposure groups have been associated with offspring differential methylation at 71 CpGs, including genes associated with inflammatory response[37]. Of the toxins that may demonstrate a seasonal trend in The Gambia pesticide exposure may be the first choice to focus on, given it appears to affect DNA methylation not only in the person directly exposed[38] but potentially intergenerationally as well[39].

Paternal influence

There is growing interest in the potential for paternal exposures to affect the infant epigenome transgenerationally via sperm methylation profiles[40,41]. Illum *et al.* (2018)[42] have reviewed mouse experiments altering paternal diet and detecting changes in either

sperm methylation or offspring methylation, which are also associated with offspring phenotype. To date there is preliminary, although conflicting, evidence to suggest that both paternal high-fat or low-protein diets may affect sperm methylation, and that these changes could contribute to adverse phenotypes such obesity and insulin resistance seen in the offspring[42]. Despite the fact mechanisms and consistent patterns remain to be determined there is already a call for researchers to consider joint parental contribution in intergenerational and transgenerational epigenetics studies[43,44].

References

1.   Cani PD (2014) Metabolism in 2013: The gut microbiota manages host metabolism. *Nat. Rev. Endocrinol.* **10**, 74–6.

2.   Shen J, Obin MS & Zhao L (2013) The gut microbiota, obesity and insulin resistance. *Mol. Aspects Med.* **34**, 39–58.

3.   Greiner T & Bäckhed F (2011) Effects of the gut microbiota on obesity and glucose homeostasis. *Trends Endocrinol. Metab.* **22**, 117–123.

4.   Koenig JE, Spor A, Scalfone N, et al. (2011) Succession of microbial consortia in the developing infant gut microbiome. *Proc. Natl. Acad. Sci. U. S. A.* **108 Suppl**, 4578–85.

5.   Duncan SH, Belenguer A, Holtrop G, et al. (2007) Reduced Dietary Intake of Carbohydrates by Obese Subjects Results in Decreased Concentrations of Butyrate and Butyrate-Producing Bacteria in Feces. *Appl. Environ. Microbiol.* **73**, 1073–1078.

6.   Sheflin AM, Melby CL, Carbonero F, et al. (2017) Linking dietary patterns with gut microbial composition and function. *Gut Microbes* **8**, 113–129.

7.   LeBlanc JG, Chain F, Martín R, et al. (2017) Beneficial effects on host energy metabolism of short-chain fatty acids and vitamins produced by commensal and probiotic bacteria. *Microb. Cell Fact.* **16**, 79.

8.   Krautkramer KA, Kreznar JH, Romano KA, et al. (2016) Diet-Microbiota Interactions Mediate Global Epigenetic Programming in Multiple Host Tissues. *Mol. Cell* **64**, 982–992.

9.   Mischke M & Plösch T (2016) The Gut Microbiota and their Metabolites: Potential Implications for the Host Epigenome. *Adv. Exp. Med. Biol.* **902**, 33–44.

10.   Nicholson JK, Holmes E, Kinross J, et al. (2012) Host-gut microbiota metabolic interactions. *Science* **336**, 1262–7.

11.   Bhat MI & Kapila R (2017) Dietary metabolites derived from gut microbiota: critical modulators of epigenetic changes in mammals. *Nutr. Rev.* **75**, 374–389.

12.   Ye J, Wu W, Li Y, et al. (2017) Influences of the Gut Microbiota on DNA Methylation and Histone Modification. *Dig. Dis. Sci.* **62**, 1155–1164.

13. Carbonero F (2017) Human epigenetics and microbiome: the potential for a revolution in both research areas by integrative studies. *Futur. Sci.* **3**, FSO207.

14. Kumar H, Lund R, Laiho A, et al. (2014) Gut microbiota as an epigenetic regulator: pilot study based on whole-genome methylation analysis. *MBio* **5**, e02113-14.

15. Huang T, Zheng J, Chen Y, et al. (2011) High consumption of Ω-3 polyunsaturated fatty acids decrease plasma homocysteine: a meta-analysis of randomized, placebo-controlled trials. *Nutrition* **27**, 863–7.

16. Roy S, Kale A, Dangat K, et al. (2012) Maternal micronutrients (folic acid and vitamin B(12)) and omega 3 fatty acids: implications for neurodevelopmental risk in the rat offspring. *Brain Dev.* **34**, 64–71.

17. Khot V, Kale A, Joshi A, et al. (2014) Expression of genes encoding enzymes involved in the one carbon cycle in rat placenta is determined by maternal micronutrients (folic acid, vitamin B12) and omega-3 fatty acids. *Biomed Res. Int.* **2014**, 613078.

18. Huang T, Hu X, Khan N, et al. (2013) Effect of Polyunsaturated Fatty Acids on Homocysteine Metabolism through Regulating the Gene Expressions Involved in Methionine Metabolism. *Sci. World J.* **2013**, 1–8.

19. Khot V, Chavan-Gautam P & Joshi S (2015) Proposing interactions between maternal phospholipids and the one carbon cycle: A novel mechanism influencing the risk for cardiovascular diseases in the offspring in later life. *Life Sci.* **129**, 16–21.

20. Lee H-S, Barraza-Villarreal A, Hernandez-Vargas H, et al. (2013) Modulation of DNA methylation states and infant immune system by dietary supplementation with ω-3 PUFA during pregnancy in an intervention study. *Am. J. Clin. Nutr.* **98**, 480–7.

21. Lee H-S, Barraza-Villarreal A, Biessy C, et al. (2014) Dietary supplementation with polyunsaturated fatty acid during pregnancy modulates DNA methylation at IGF2/H19 imprinted genes and growth of infants. *Physiol. Genomics* **46**, 851–7.

22. Eckersley-Maslin MA, Alda-Catalinas C & Reik W (2018) Dynamics of the epigenetic landscape during the maternal-to-zygotic transition. *Nat. Rev. Mol. Cell Biol.* **19**, 436–450.

23. Young JI, Züchner S & Wang G (2015) Regulation of the Epigenome by Vitamin C. *Annu. Rev. Nutr.* **35**, 545–564.

24. Chung T-L, Brena RM, Kolle G, et al. (2010) Vitamin C promotes widespread yet specific DNA demethylation of the epigenome in human embryonic stem cells. *Stem Cells* **28**, 1848–55.

25. Blaschke K, Ebata KT, Karimi MM, et al. (2013) Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature* **500**, 222–226.

26. Menezo YJR, Silvestris E, Dale B, et al. (2016) Oxidative stress and alterations in DNA methylation: two sides of the same coin in reproduction. *Reprod. Biomed. Online* **33**, 668–683.

27. Tyagi N, Sedoris KC, Steed M, et al. (2005) Mechanisms of homocysteine-induced

oxidative stress. *Am. J. Physiol. Circ. Physiol.* **289**, H2649–H2656.

28. Škovierová H, Vidomanová E, Mahmood S, et al. (2016) The Molecular and Cellular Effect of Homocysteine Metabolism Imbalance on Human Health. *Int. J. Mol. Sci.* **17**, 1733.

29. Hoffman M (2011) Hypothesis: Hyperhomocysteinemia is an indicator of oxidant stress. *Med. Hypotheses* **77**, 1088–1093.

30. Menezo YJR, Elder K & Dale B (2015) Link Between Increased Prevalence of Autism Spectrum Disorder Syndromes and Oxidative Stress, DNA Methylation, and Imprinting. *JAMA Pediatr.* **169**, 1066.

31. Claycombe KJ, Brissette CA & Ghribi O (2015) Epigenetics of inflammation, maternal infection, and nutrition. *J. Nutr.* **145**, 1109S–1115S.

32. Lu L & Claud EC (2018) Intrauterine Inflammation, Epigenetics, and Microbiome Influences on Preterm Infant Health. *Curr. Pathobiol. Rep.* **6**, 15–21.

33. Weber-Stadlbauer U (2017) Epigenetic and transgenerational mechanisms in infection-mediated neurodevelopmental disorders. *Transl. Psychiatry* **7**, e1113.

34. Pohanka M (2013) Role of oxidative stress in infectious diseases. A review. *Folia Microbiol.* **58**, 503–513.

35. Sies H, Berndt C & Jones DP (2017) Oxidative Stress. *Annu. Rev. Biochem.* **86**, 715–748.

36. Perera F & Herbstman J (2011) Prenatal environmental exposures, epigenetics, and disease. *Reprod. Toxicol.* **31**, 363–73.

37. Hernandez-Vargas H, Castelino J, Silver MJ, et al. (2015) Exposure to aflatoxin B1 in utero is associated with DNA methylation in white blood cells of infants in The Gambia. *Int. J. Epidemiol.* **44**, 1238–48.

38. Rusiecki JA, Beane Freeman LE, Bonner MR, et al. (2017) High pesticide exposure events and DNA methylation among pesticide applicators in the agricultural health study. *Environ. Mol. Mutagen.* **58**, 19–29.

39. Schmidt RJ, Schroeder DI, Crary-Dooley FK, et al. (2016) Self-reported pregnancy exposures and placental DNA methylation in the MARBLES prospective autism sibling study. *Environ. Epigenetics* **2**, dvw024.

40. Soubry A (2015) Epigenetic inheritance and evolution: A paternal perspective on dietary influences. *Prog. Biophys. Mol. Biol.* **118**, 79–85.

41. Radford EJ, Ito M, Shi H, et al. (2014) In utero effects. In utero undernourishment perturbs the adult sperm methylome and intergenerational metabolism. *Science* **345**, 1255903.

42. Illum LRH, Bak ST, Lund S, et al. (2018) DNA methylation in epigenetic inheritance of metabolic diseases through the male germ line. *J. Mol. Endocrinol.* **60**, R39–R56.

43. Steegers-Theunissen RPM, Twigt J, Pestinger V, et al. (2013) The periconceptional period, reproduction and long-term health of offspring: the importance of one-carbon metabolism. *Hum. Reprod. Update* **19**, 640–55.

44. Fleming TP, Watkins AJ, Velazquez MA, et al. (2018) Origins of lifetime health around the time of conception: causes and consequences. *Lancet* **391**, 1842–1852.